CHAPTER 1

INTRODUCTION

Freshwater aquaculture is growing alongside with the marine penaeid prawn which is the major crustacean group in culture. Giant freshwater prawn, *Macrobrachium rosenbergii*, is the largest species being cultured in the freshwater industry. There are many countries involved in established inland prawn farming such as Thailand, Brazil, India, Ecuador, China, Israel, Mexico, Peru and United States (Boyd, 2002). Aquaculture being one of the vast expanding industries in Malaysia holds a great potential to revolutionize the rural freshwater prawn farming by generating high profit for the farmers.

Though the farming of freshwater prawns forms only a small percentage of the global aquaculture production, it has a considerable importance in the growing appetite of customers around the world. Farmed giant freshwater prawns, *Macrobrachium rosenbergii*, has invaded the international market soon after it hit the domestic markets. Asia contributes around 80% of the total world production (Mather, 2008; Mather and Bruyn, 2003). This rapid demand for the giant freshwater prawn has stimulated and intensified interest in this species culture around the world. Malaysia is well known for its aquaculture activities which have been established even before independence and is still one of the key holders of aquaculture products in the South East Asia where freshwater prawn production is among one of the industry. Due to the increasing economical importance of this giant freshwater prawn, more attention, money and energy are channeled into the hatchery and grow-out phase.

The large giant freshwater prawn, *M. rosenbergii*, which originated from Malaysia, has been introduced into many other countries, as it has the potential to grow fast and tolerate moderate temperature and salinity changes. Fujimura and Okamoto (1972) stated that in the year 1965-1966, *M. rosenbergii* was imported to Hawaii where Takuji Fujimura and his team developed a mass rearing technology for larval production. Grow-out experiments have established commercialized farming techniques. D'Abramo and Brunson (1996) quoted that Malaysia was the first to develop the production techniques in 1950s and research was also being done in Israel, Hawaii and several States of America (South Carolina, Florida, Texas, Louisiana and Kentucky).

There are many different species of *Macrobrachium* in the world, but *M. rosenbergii* is the most commercially cultured species due to its large size, less aggressive nature under cultured conditions and aquaculture potential (Ahmed *et al*., 2008). *M. rosenbergii* is mostly found in the tropical and subtropical waters in the Indo-Pacific region in Malaysia, Thailand, the Philippines, India, Sri Lanka, Bangladesh, Myanmar, Indonesia and Vietnam (Tayamen, 2001). The tropical and sub-tropical climate enables the prawn to survive and breed throughout the year where it normally takes approximately six months to grow into an adult from newly hatched larvae. This adaptation has built constraints for its culture in other temperate countries where they can only survive during summer and will die during winter. Therefore, these countries can only run one cycle in a year.

According to Nandlal and Pickering (2005), giant freshwater prawn farming is not capital intensive and is accessible to small scale farmers compared to marine shrimp farming. The hatchery and farm operation does not require high technical knowledge. Since

much emphasis is given to the farming of this species, there are a few general genetics issues that should be addressed. During domestication it is a common practice that only a small founder population and a small number of broodstock is chosen for breeding causing mating between close relatives. Inbreeding and genetic drift can result in loss of genetic variation, in turn eliminating the potential for future genetic improvement.

Many factors have been studied and practiced in commercial production of this species to increase profitability and sustainability of production enterprises. There is now a growing importance of genetics in aquaculture. Freshwater prawn farming has been latent over the past years due to a few drawbacks such as infections, lack of understanding of prawn reproductive biology, poor hatchery and farm management and limited availability of wild brood stock and postlarvae (Jerry *et al*., 2004). Major gains from nutrition and husbandry management have already been achieved. They also added that inbreeding, negative selection and unwanted hybridization has led to an increased awareness of poor genetic management. The technologies in genetic improvement include exploiting additive genetic variance (classic selective breeding and enhanced selective breeding using molecular markers), exploiting non-additive genetic variance (hybridization, crossbreeding) and genetic manipulation (chromosome set manipulation, sex control, transgenes).

Molecular markers show significant promise for aquaculture applications. Microsatellites are repeats of 1-6bp nucleotide motifs. Microsatellite analysis reveals more detailed information on genetic diversities than other molecular markers. It also has become instrumental as genetic markers in population genetics, parentage analysis and genetic mapping (Li *et al*., 2003). These markers are able to provide valuable information such as comparison of hatchery and wild stocks, genetic identification and discrimination of

hatchery stocks, monitoring inbreeding or other changes in the genetic variation and assignment of progeny to parents through genetic tags (Ciftci and Okumus, 2003; Ferguson, 1995). The genetic variability among domesticated and their wild donor population has evolved due to domestication selection or the mating design.

At the current moment, population structure information is not sufficient for the Malaysian giant fresh water prawn species. Thus, this study can contribute to estimate the population structure. The objectives of the present study were:

1. To carry out a population structure study of *M. rosenbergii* using microsatellite markers to identify the base population for genetic improvement program,
2. To investigate the likelihood of homogeneity and divergence among populations of the same location and between populations in Malaysia and Thailand, and
3. To study the factors involved in backyard hatchery operation including the larval stages.

CHAPTER 2

LITERATURE REVIEW

2.1 *Macrobrachium rosenbergii*

2.1.1 Classification and taxonomy

Kingdom: Animalia

Phylum: Arthropoda

Subphylum: Crustacea

Class: Malacostraca

Subclass: Eumalacostraca

Superorder: Eucarida

Order: Decapada

Suborder: Pleocyemata

Infraorder: Caridea (sometimes called as Natantia)

Superfamily: Plaemonoidea

Family: Palaemonidae

Subfamily: Palaemoninae

Genus: *Macrobrachium*

Species: *rosenbergii* ((De Man, 1879) -giant river prawn

Subspecies: *Macrobrachium rosenbergii rosenbergii* (De Man, 1879)
Macrobrachium rosenbergii dacqueti (Sunnier, 1925) previously known as *Macrobrachium rosenbergii schenkeli* (Johnson, 1973)

(source:http://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=all&search_value=Macrobrachium+rosenbergii&search_kingdom=every&search_span=exactly_for&categories=All&source=html&search_credRating=All)

All of the cultured freshwater prawns were from the genus *Macrobrachium* which is the largest genus of the family Palaemonidae (Anon, 2002). Common names for this species are giant freshwater prawn, freshwater river prawn, Malaysian freshwater prawn, ‘udang galah’ and ‘scampi’ in India.

Macrobrachium rosenbergii rosenbergii (De Man, 1879) is referred to as the eastern form which is found in the Philippines, Indonesia (Sulawesi and Irian Jaya), Northern Australia and Papua New Guinea. The western form is known as *M. rosenbergii dacqueti* (Sunnier, 1925) and found in the waters of the east coast of India, Malaysia, Thailand, Bangladesh and Indonesia (Sumatra, Java, Kalimantan). This statement has been further supported by Mather and Bruyn (2003), whereby they utilized 16S rRNA mtDNA gene to substantiate the divergence between the eastern and the western forms. These two subspecies is said to be separated about 5 to 10 million years ago. This has clearly placed the Malaysian giant freshwater prawn into the western form within Huxleys line (Bruyn *et al*., 2004; Mather and Bruyn, 2003).

2.1.2 Biology of *Macrobrachium rosenbergii*

2.1.2.1 Habitat and life cycle

As a freshwater river prawn, this species normally lives in freshwater environment in the tropical and subtropical waters (Ismael and New, 2000). It inhabits lakes, rivers, swamps, irrigation ditches, canals, ponds and also estuarine regions. Although it is known as freshwater prawn, a part of its life cycle is also influenced by brackish water regions where the larval development takes place (Ling and Merican, 1961).

Gravid females migrate to low salinity estuarine regions for spawning purposes where the eggs hatch as free-swimming larvae and once metamorphosis is complete into post larvae (PL) where they migrate back to freshwater for growth and maturity (Kutty *et al.,* 2000). The free-swimming larvae go through eleven zoeal stages before they reach PL. During the larval stages, they are planktonic and actively swim tail first with the ventral side facing the surface (Ismael and New, 2000; Nandlal and Pickering, 2005). In their natural environment, the larval diet consists of zooplankton and also the larval stages of other crustaceans.

On the other hand, once the larvae metamorphose into PL, they resemble miniature adult prawns and will start swimming forward with the dorsal side uppermost (New and Singholka, 1985). Now, they become benthic and are able to walk on substratum. According to Scudder *et al*., (1981), these juvenile prawns exhibit characteristics of the adults where they become active at night (nocturnal) and omnivorous. There are signs of cannibalism when there is a shortage of food and also when the habitat is overpopulated.

M. rosenbergii can maintain its sexual maturation and gonadal development throughout the year better than the other species of this genus (Ra'anan and Cohen, 1983). When the adult prawns are matured, mating occurs continuously or periodically. Prior to mating, ripe females undergo pre-spawning molt which usually occur at night and become soft-shelled while the males do not go through any changes. Once the females molt, males begin to insert a sperm mass into the ventral thoracic region of the female. As eggs extrude from the female, fertilization occurs externally. The fertilized eggs then adhere to the pleopods of the females. (Damrongphol and Jaroensastraraks, 2000)

A

B

Figure 2.1: A. Male *Macrobrachium rosenbergii* [Photo: Thamayanthi Nada Raja]

B. Berried *M. rosenbergii* females [Photo: Takuji Fujimura]

2.1.2.2 Morphology

Macrobrachium rosenbergii is the largest species in the genus, whereby the males can grow up to 32 cm in total length while adult females can measure up to 25 cm at similar age. Their body is usually greenish to brownish grey in color, where brown, grey and whitish streaks patterns appear longitudinally on the dorsal side of the body (New and Valenti, 2000).

The body of this river prawn is divided into two parts which are the cephalothorax (head) and the abdomen (tail). The former is covered by a protective shield called carapace which is smooth and hard. It has eight thoracic appendages where the first three appendages known as maxillipeds is utilized for feeding. Of the eight, the last five appendages are known as pereopods where the first two pairs carry pincers known as chelipeds, utilized for food capturing while the last three pairs are typical walking legs. The abdomen bears six abdominal appendages where the first five pairs are well-developed pleopods also known as swimmerets adapted for swimming. The last pair called the uropods form a very effective propulsive tail fan for reverse movement. (Ismael and New, 2000)

At the same age, male prawns are larger compared to females. An obvious observation on the male indicates that its carapace is proportionally larger than the abdomen. The female chelipeds are longer and larger than the female’s. The common method to identify the sexes is the protrusion between the first and the second swimmeret pairs which signifies a male (Ismael and New, 2000). Karplus *et al*., (2000), stated that the variation among male prawn size portrays a population structure where it consists of three mature male morphophytes. The three male morphotypes; blue claw male, orange claw

male and small male describes sexually matured males. Of the three forms, the blue claw (BC) male, possesses extremely large claw and is said to be most sexually active. The orange claw (OC) male has orange claw which in length is shorter than those of blue claw males. The small male has small claws and is also known as runt (Fujimura and Okamoto, 1972).

Nandlal and Pickering (2005), explained that there are three types of females, virgin females, berried females (egg carrying females) and spent females (open brood chamber). Females have smaller head, slender claws and the first three pairs of pleopods are longer and broader, forming an egg incubating chamber.

2.2 Molecular markers

Molecular markers show significant promise for aquaculture applications. According to Brooker (2005), molecular marker is a segment of DNA found at specific sites of the genome and has properties that enable it to be uniquely recognized using gel electrophoresis. The advancements in molecular genetics has now made possible for restriction endonucleases and Polymerase Chain Reaction (PCR) technology to detect polymorphism directly at the DNA level.

Molecular markers can be used for a wide range of purposes such as germplasm characterization, genetic diagnostics, characterization of transformants, study of genome organization, phylogenetic analysis, marker-assisted selection, mapping quantitative trait loci (Nalini *et al*., 2003), cytogenetic and physical mapping studies (Brooker, 2005), gene disease association studies (Lench *et al*., 1996) and study of genetic diversity (Gupta & Varshney, 1999).

Molecular markers could be clustered into three classes. They are allozymes, mitochondrial DNA and nuclear DNA. Allozymes produced by a single gene locus are allelic variants of proteins. According to Hillis *et al*., 1996, starch gel electrophoresis of allozymes was commonly used in fisheries molecular genetics where it possesses Mendelians' co-dominant characteristic. Allozymes have been utilized in aquaculture genetics for parentage analysis, tracking inbreeding, stock identification (Liu and Cordes, 2004) and linkage mapping (McGoldrick and Hedgecock, 1997). However, Liu and Cordes (2004), also added that one of the drawbacks using allozymes is heterozygote deficiencies due to null allele. In addition to that, it is also said that only a small proportion of enzyme

loci are polymorphic in many organisms (Magoulas, A, 1998). A few advantages of allozyme usage include ease of use, low cost and are also a co-dominant marker (Liu and Cordes, 2004).

Secondly, mitochondrial DNA became popular in phylogeny and population structure studies during the past few decades. Mitochondrial DNA can be distinguished from nuclear DNA through several properties: it is maternally inherited (transmitted uniparently), it occurs in multiple copies in each cell, and it does not recombine. Uniparental transmission has a disadvantage where the effective population size for mtDNA is smaller than that of the nuclear DNA; therefore, mtDNA is a good indicator for bottleneck and hybridization. So, divergence among populations can be higher in mtDNA than in nuclear DNA (Magoulas, A., 1998). On the contrary, although mtDNA is maternally transmitted, gene flow can also be contributed paternally. Therefore, the disadvantage of mtDNA is, it is unable to trace genetic mixing through paternal contributions.

Finally, nuclear DNA markers are the latest addition in aquaculture genetics. They are widely accepted due to their ability to detect high levels of polymorphism, their abundance in the eukaryotic nuclear genome and Mendelian segregation (Jerry *et al.*, 2006). According to O'Brien (1991), nuclear DNA can be grouped into two types, type I and type II, based on their function. A DNA segment which is a part of an expressed DNA is not very polymorphic but show high evolutionary conservation ability. Single nucleotide polymorphism (SNP) and restriction fragment length polymorphism (RFLP) are two popular type I nuclear DNAs. On the other hand, type II markers are known as the non-functional markers without biological functions but able to show high levels of

polymorphism. Microsatellites, minisatellites, randomly amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) are the highly utilized nuclear DNA markers used in aquaculture studies to date.

The choice of molecular marker solely depends on the objective of a genome research. Table 2.1 gives an insight and a comparative analysis of some of the markers being widely utilized.

Table 2.1: Types of DNA markers, their characteristics and potential applications.

Marker type	Acronym or alias	Requires prior molecular information?	Mode of inheritance	Type	Locus under investigation	Likely allele numbers	Polymorphism or power	Major applications
Allozyme		Yes	Mendelian, Codominant	Type I	Single	2-6	Low	Linkage mapping, population studies
Mitochondrial DNA	mtDNA	No[a]	Maternal inheritance	–		Multiple haplotypes		Maternal lineage
Restriction fragment length polymorphism	RFLP	Yes	Mendelian, Codominant	Type I or type II	Single	2	Low	Linkage mapping
Random amplified polymorphic DNA	RAPD, AP-PCR	No	Mendelian, Dominant	Type II	Multiple	2	Intermediate	Fingerprinting for population studies, hybrid identification
Amplified fragment length polymorphism	AFLP	No	Mendelian, Dominant	Type II	Multiple	2	High	Linkage mapping, population studies
Microsatellites	SSR	Yes	Mendelian, Codominant	Mostly Type II	Single	Multiple	High	Linkage mapping, population studies
Expressed sequence tags	EST	Yes	Mendelian, Codominant	Type I	Single	2	Low	Linkage mapping, physical mapping
Single nucleotide polymorphism	SNP	Yes	Mendelian, Codominant	Type I or type II	Single	2, but up to 4	High	Linkage mapping, population studies?
Insertions/ deletions	Indels	Yes	Mendelian, Codominant	Type I or type II	Single	2	Low	Linkage mapping

[a] Conserved PCR primers can be adopted from sequence information from a related species.

Modified from Liu and Cordes (2004), Aquaculture 238: 1-37

2.2.1 Microsatellites

Microsatellites which are also known as short tandem repeats (STRs) or simple sequence repeats (SSRs) range in size from 1 to 6 base pairs (Tautz, 1989; Litt and Luty, 1989). A large number of repeats in a microsatellite are more polymorphic, where the repeat motifs include mono (one), di (two), tri (three), tetra (four), and penta (five) nucleotides (Karsi *et al*., 2002b). In addition, Presti *et al*., (2009), stated that microsatellites are co-dominant and are also present abundantly throughout the genome of an organism. Microsatellites are generally found in non-coding regions (Jerry et al., 2006), but, they have also been found in coding regions and non-gene sequences (Liu and Cordes, 2004).

Perfect microsatellite consists of a single repeat motif and is not interrupted at any location by a base which does not match the repeat motif. Disrupted microsatellites can be divided into three groups. Firstly, imperfect microsatellite with one or more repeats having a base pair that does not fit the repeat pattern. Interrupted microsatellite, on the other hand, has a small number of base pair insertions. Finally, compound microsatellite consists of two or more adjacent microsatellites with different repeat patterns. (Estoup and Cornuet, 1999)

The high allelic diversity in microsatellites creates the potential to determine pedigree relationship among individuals. However, there are certain constraints such as presence of mutations, null (non-amplifying) alleles and stuttering which can cause errors in likelihood analyses (Hoffman and Amos, 2005).

The microsatellite DNA markers can be directly amplified through PCR by primers that have been synthesized using unique sequence. An example of a forward and reverse

primer pair flanking both sides of a microsatellite region is shown in Figure 2.2. Homozygous individuals will show a single band for a microsatellite locus while a heterozygous individual would show two bands. This proves Mendelian inheritance theory that microsatellites are co-dominantly inherited. High mutation rates in organisms result in high polymorphisms of microsatellites.

Figure 2.2: Forward and reverse primer flanking microsatellite region.

Although microsatellites are said to be the best for PCR amplification process, they do have their advantages and disadvantages. The advantage of microsatellite as a molecular marker includes the co-dominance inheritance, easy scoring of alleles, reproducibility, and accessibility to laboratories lacking highly sophisticated analysis equipments (Paniego *et al*., 2002). On the contrary, the disadvantage of microsatellites include non-amplified alleles which can produce null homozygotes, cost and time to isolate microsatellite is expensive and time consuming respectively (Korpelainen *et al*., 2007).

The high mutation rates in microsatellites can be due to two mechanisms; (1) recombination between DNA molecules by gene conversion or by unequal crossing over (Smith, 1976, Jeffreys *et al*., 1994) and (2) slipped-strand mispairing during DNA replication (Levinson and Gutman, 1987a; 1987b). The latter seemed to be the main

mechanism generating length mutation in microsatellites (Levinson and Gutman, 1987b). This slippage mispairing mechanism mostly causes small changes in repeat numbers where alleles of similar lengths are closely related than that of very different sizes. It can generate size homoplasy depending on mutation rate of the locus and also divergence time of two populations.

Microsatellite markers require development of such markers prior to the utilization. If, time and resources are provided to develop these microsatellite markers, it would be worthwhile for genome mapping. Microsatellites can be applied as DNA marker in species identification, strain identification, hybrid identification, paternity determination, diversity analysis, genetic mapping and comparative mapping. (Liu and Cordes, 2004)

CHAPTER 3

METHODOLOGY

3.1 Breeding experiment

Prior to the base population structure studies, basal experiments were conducted in a private backyard hatchery located in Port Dickson, Negeri Sembilan beginning 12th May 2009 to 24th July 2009. The experiments were conducted in 6 tanks (1A, 1B, 2A, 2B, 3A and 3B) which had a flow-through system while two more tanks tested had a semi-closed system (4A and 5A). Berried females were brought in from grow-out farms operated at Pantai and Kuala Klawang, Negeri Sembilan. They were stored in freshwater tanks until they were ripe for egg hatching.

During the three month experiment, management necessities such as parameter tests, tank hygiene maintenance, water treatment using industrial grade chlorine, broodstock disinfection using formalin and alternate formalin treatment for the larvae were carried out to avoid any possible mishap. Formalin treatment works for bacterial infection up to a certain level but does not seem to affect viral infections. In a hatchery operation, water quality is critical. Water quality is given utmost importance, as; increase in ammonia level can be fatal for the larvae. Therefore, excessive or residual feed always has to be siphoned out to circumvent increase in ammonia, nitrite, nitrate and phosphorus.

The hatchlings were grown in separate tanks with the density of approximately 60,000 to 100,000 larvae per tank to estimate the best survival rate density.

3.2 Marker selection

Freshwater prawns in Southeast Asia are categorized into the 'western' form. Thailand scientists have isolated and characterized loci Mbr1, Mbr3, Mbr4, Mbr5, Mbr7, Mbr8, Mbr10A, Mbr10B and Mbr11 (Charoentawee *et al*., 2006). Locus Mr78 was isolated from the 'eastern' form of the giant freshwater prawn and had been tested in the 'western' form by Chand *et al*., (2005). The list of primers is shown in Table 3.1.

3.3 Sample collection

M. rosenbergii were collected from four different geographical locations in Peninsular Malaysia. Absolute ethanol was used to store the body tissues collected at site which prevents tissue degradation. The populations involved in this study consist of Pantai in Negeri Sembilan, Kampung Acheh in Perak, Teluk Kumbar, Penang and Sungai Perak in Perak (Table 3.2), with a sampling size of 32 individuals per location. The government funded hatchery run in Kg Acheh, have reported that they normally use 150 females and 50 males to ensure variability. The information on eight other locations which were sampled in 2006/07 by our previous research team is also provided in Table 3.2. The sampling locations are shown in Figure 3.1.

All the populations were mixed and matched to conduct separate analysis. The data analysis conducted for each populations/groups is as follows:

1. Population Teluk Kumbar, Kg Acheh, Pantai and Sg Perak 2009 (All analyses)
2. Population Kg Acheh and the grouped populations of Teluk Kumbar, Pantai and Sg Perak 2009 (AMOVA)
3. All 12 populations collected in the 2009 and 2006/07 (PHYLIP v3.67, AMOVA and STRUCTURE)
4. All 12 populations were clustered into two separate groups based on the year of collection, 2009 and 2006/07 (AMOVA)

Table 3.1: Microsatellite markers used for the study

Locus	Primer sequence 5'-3'	Annealing temperature(°C)	Repeat sequence	GenBank Accession no.
Mbr-1	F:CCCACCATCAATTCTCACTTACC	62	(GA)24	DQ019863
	R:TCCTTTTCACATCGTTTCCAGTC			
Mbr-3	F:CAACTCTATGTTTCGGCATTTGG	62	(AG)14	DQ019865
	R:GGGGAATTTTACCGATGTTTCTG			
Mbr-4	F:CCACCTACCGTACATTCCCAAAC	62	(GT)29	DQ019866
	R:CGGGGCGACTTTTAGTATCGAC			
Mbr-5	F:CAAGGCTCGTGTCTCTTGTTTC	62	(AG)25	DQ019867
	R:GCTTGTACTTGTTCAGCTTTTGC			
Mbr-7	F:ATAAAAGAGTCGCCAAATGAGCA	62	(TGC)16	DQ019869
	R:ATTGGGAATTGTTGACCTCCAAG			
Mbr-8	F:AACCAGCCGACTTAGACTGTGC	62	(AGC)6(AG)5AA(AG)4	DQ019870
	R:CGCCATTTGCGTCTATCTCTTAC			
Mbr-10A	F:ATGACGATGATGAGGAATGAAGC	60	(ATG)3A(ATG)4	DQ019871
	R:TTTCAGGCTATATCAAGCAACAG			
Mbr-10B	F:ATGACGATGATGAGGAATGAAGC	60	(ATG)3A(ATG)4	DQ019871
	R: TTTCAGGCTATATCAAGCAACAG			

Table 3.1: Microsatellite markers used for the study (continued)

Locus	Primer sequence 5’-3’	Annealing temperature(°C)	Repeat sequence	GenBank Accession no.
Mbr-11	F:GTATTGAGAACAAAGGCGCACAG	63	(AG)31	DQ019872
	R:ATCTCTTTCCAAAACAGGGCACA			
Mr-78	F:GGACAAAACAAGCAGAAAAG	60	(GA)31	AY791967
	R:CAGGCACAGTGATAACCAA			

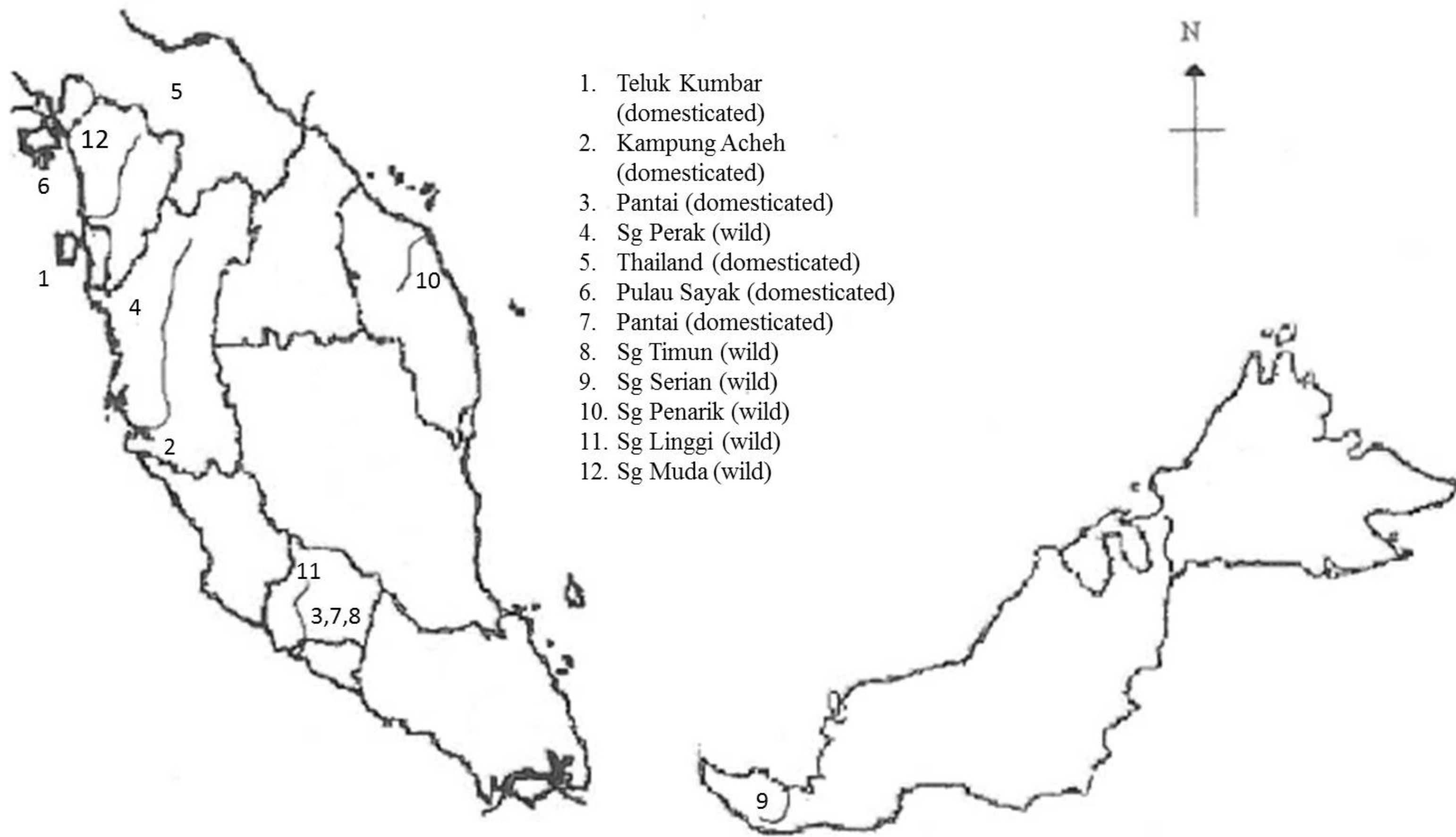

Figure 3.1: Sampling locations.

Table 3.2: Description of the location of sample collection.

	Location	**Longitude/ latitude**	**Date of sampling**	**No. of samples**
P1-P32	Teluk Kumbar, Penang	5°17'18.27" N 100°14'20.82" E	5th August 2009	32
K1-K32	Kampung Acheh, Perak	4°14'41.58" N 100°40'08.17" E	5th August 2009	32
M1-M32	Pantai, Negeri Sembilan	2°48'11.82" N 101°59'30.76" E	5th August 2009	32
S1-S32	Sg Perak, Perak	5°06'19.33" N 100°29'07.4" E	17th September 2009	32

Table 3.2: (continue) Samples collected in 2006/07.

	Location	Longitude/ latitude	Date of sampling	No. of samples
T1-30	Thailand	(the location of the hatchery was not revealed)	2006/07	30
D1-30	Pulau Sayak, Kedah	5º40’00.00”N 100º20’00.00”E	2006/07	30
Mu1-30	Pantai, Negeri Sembilan	2°48’11.82” N 101°59’30.76” E	2006/07	30
W1-30	Sg Timun, Negeri Sembilan	2º56’49.75”N 102º05’07.63”E	2009	30
Sk1-30	Sg Serian, Sarawak	1º50’05.64”N 113º54’06.27”E	2006/07	30
TE1-30	Sg Penarik, Terengganu	5º37’48”N 102º48’36”E	2006/07	30
N1-24	Sg Linggi, Negeri Sembilan	2º28’29.23”N 102º02’05.04”E	2006/07	24
Ke1-24	Sg Muda, Kedah	5º43’01.32”N 100º31’46.60”E	2006/07	24

3.4 Genomic DNA preparation

3.4.1 DNA isolation

The genomic DNA isolation was conducted using a GF-1 Tissue Extraction Kit (Vivantis) on the *M. rosenbergii* tissues according to the instructions with slight modifications.

20mg of tissue sample was cut into small pieces using a clean scalpel and were placed into a 1.5 ml micro centrifuge tube. This was followed by addition of 250μl of TL Buffer and 20μl of Proteinase K solution (20mg/ml) and then vortexed to obtain a homogenous mixture. Then, 12μl of Lysis Enhancer was added and mixed well quickly. The tubes were incubated for 1-3 hour at 65°C in a rotating water-bath and occasionally mixed to ensure an overall digestion of the sample. Then, another 20μl of RNase A was added to the sample and mixed thoroughly. The tube was incubated at 37°C for 10 minutes.

Next, 600μl of TE Buffer was added to the sample and vortexed, then incubated for another 10 minutes at 65°C. After incubation, 200μl of absolute ethanol was added to the sample and mixed well (to prevent uneven precipitation of nucleic acid). Then, approximately 600μl of the mixture solution was transferred gently into a column assembled in a clean collection tube and centrifuged at 5000 ×g for 1 minute. The remainder of the original solution was kept and the flow-through was discarded. The column was then washed with 750μl of Wash buffer and centrifuged again at 5000 ×g for 1 minute. The flow-through was discarded. This washing step was repeated once again. Next,

the column was centrifuged at 10000 ×g for 1 minute to remove all traces of ethanol.

Finally, DNA elution was carried out by placing the column into a clean micro centrifuge tube, followed by the addition of 200µl of preheated Elution Buffer directly onto the column membrane and incubated under the room temperature for 2 minutes. Next, the solution was centrifuged at 5000 ×g for 1 minute to elute the DNA. The extracted DNA product was then stored at -20°C.

3.4.2 Measurement of DNA purity

The quality and purity of the isolated DNA was quantified using a UV-spectrophotometer and agarose gel electrophoresis. The DNA sample was diluted 1:20 by dissolving 5µl of DNA in 95µl of autoclaved distilled water. The absorbency was measured at 260nm and 280nm. The formula for quality calculation is as below:

$$\text{DNA concentration} = \text{OD 260 value} \times 50 \times \text{dilution factor}$$

50 represent the concentration of double stranded DNA at OD1.

DNA with ratios ranging from 1.8 to 2.0 was considered as pure and can be subjected to electrophoresis for quality check. In this study, the average ratio between 20 samples was 1.77.

3.4.3 PCR Amplification

PCR optimization was done for $MgCl_2$ depending on the microsatellite locus. The cycle number and several annealing temperatures (modified from Charoentawee *et al*., 2006) were tested for each primer set. Markers which successfully amplified were chosen to screen all the samples of the four prawn populations.

2.0µl of template DNA, 1µl of 1.5mM $MgCl_2$, 3µl of 1X Promega reaction buffer, 0.25µl of each dATP, dTTP, dGTP and dCTP were the contents of the PCR amplification reaction. 0.3µl of Promega *Taq* Polymerase, 0.5µl of each reverse and forward primer and 1.7µl of double distilled water were added into the 10µl reaction mixture. Normally, a master mix containing all the mixture components except for DNA and de-ionized distilled water will be prepared for all the cycles for each microsatellite primer pair. The master mix was prepared for two extra samples to compensate for pipetting errors. Then, the PCR tube was gently tapped and thoroughly mixed by vortexing. Next, 8µl of the mixture was added into 2µl of DNA sample in a 0.5 ml tube. PCR amplification was then performed using an Eppendorf Mastercycler gradient thermocycler.

Table 3.3: PCR protocol

Step	**Temperature**	**Time**
Initial denaturation	94°C	5 minutes
1. Denaturation	94°C	40 seconds
2. Annealing	60-63°C	40 seconds
3. Extension	72°C	40 seconds
Step 1 to 3 repeated for 35 cycles		
Final extension	72°C	7 minutes

3.5 Gel electrophoresis

The PCR product was run through electrophoresis on a 4% Metaphor agarose gel. This gel was prepared using 2g of Metaphor agarose which was added into 50ml of 1X TBE buffer in a 500ml conical flask using a magnetic stirrer for a homogenous mixture (heated in microwave oven until completely dissolved). The mixture was then poured into the gel tray and allowed to solidify for about an hour. Then, the solidified gel was recovered and placed into the electrophoresis tank. The gel was completely submerged with 1X TBE buffer. Next, 5μl of PCR product was loaded into one of the gel wells. 2.0μl of 100bp ladder was used to determine the size of the PCR product. Electrophoresis was carried out for 2 - 2.30 hours at 75V. Then, the gel was removed and submerged in a container with ethidium bromide (Etbr) and soaked in distilled water for 10 minutes before being photographed using Alpha Imager Gel Documentation System TM (2200) at 365nm. These results were used to determine amplification.

3.6 **Fragment Analysis** (Applied Biosystems, 2004)

1µl of each PCR product was transferred into a new tube and was diluted with 9µl of double distilled water (ddH_2O). Then, 1µl of the diluted mixture was transferred to another PCR tube. Later, 10µl of dye (Hidi) was added to the PCR tube followed by 0.2µl of GenScan 500 Liz ladder. The mixture was pipetted a few times to mix the solution. The mixture was then vortexed and spun down using Denville, scientific Inc. Brushless microcentrifuge.

Next, the mixture was heated for 5 minutes at 95°C in the ABI Prism® 3130*xl* Genetic Analyzer (Applied Biosystems, USA). Immediately after heating, the mixture was placed on ice for exactly 5 minutes. Then, the mixture was transferred carefully into the analysis plate in order to avoid bubbles. The plate was tapped a few times to ensure that there are no bubbles in it. Finally, the plate was placed into the Genetic Analyzer where the results were obtained 6 hours later for every 96 samples loaded.

3.7 Genetic Data Analysis and Statistical Analysis

3.7.1 Heterozygosity

Heterozygosity is widely used to measure the genetic variation of polymorphic loci. The value for heterozygosity can range from 0 to 1. It can be divided into observed heterozygosity (H_o) and expected heterozygosity (H_e). The former is defined as the proportion of individuals in a population which are heterozygous at a particular locus. Low

observed heterozygosity may be caused by inbreeding and genotypic vigor. While, the latter means, the estimated fraction of all possible heterozygous individuals based on a known allele frequency (Nei, 1978). Therefore, deviation of the H_o from the H_e can be used in population dynamics. GDA v1.0 (Lewis & Zaykin, 2001) was used to estimate the heterozygosity value.

$$H_o = n\,(A_iA_j) / N$$

$$H_e = N\,(1-\sum Pi^2) / N-1$$

Where, $n\,(A_iA_j)$ = number of individuals with genotype A_iA_j, $i \neq j$

N = sample size

A_iA_j = alleles at the locus

Pi = frequency of the *i*th allele

3.7.2 Hardy-Weinberg Equilibrium (HWE)

HWE was developed by Hardy and Weinberg (1908) based on Mendel's principles of inheritance which is now known as Hardy-Weinberg principle. According to the principle, in a large, randomly breeding (diploid) population, allele frequency will not change from one generation to the next with the assumption that no mutation, gene migration, genetic drift or selection and genotypic frequencies are related to gene frequencies.

A deviation from HWE proportions indicates selection, population mixing or non-random mating and its detection is one of the first steps in the study of population structure (Rousset and Raymond, 1995) where such deviation can be tested using Fisher's exact test (Raymond and Rousset, 1995).

When a population meets all of the Hardy-Weinberg conditions, it is said to be in HWE. If the population was not in HWE, it shows that the population may be affected by migration, mutation, genetic drift or selection which alters the gene and genotypic frequencies. HWE was also estimated using GENPOP v4.0 (Raymond & Rousset, 1995).

3.7.3 Micro-checker analysis

When microsatellite allele data is being interpreted using this software, the following errors can be detected:

i. Null allele when one or more alleles fail to amplify during PCR,

ii. Stuttering when slight changes occur in the allele size during PCR, and

iii. Large allele dropout when large alleles do not amplify as efficiently as small alleles.

HWE theory is also used to calculate the expected allele frequencies and the frequency of any null alleles detected.

3.7.4 Genetic distance

Genetic distance refers to the genetic relationship between any two populations in allele frequencies as a function of differences between them. In short, it is a method to measure the dissimilarity of the genetic material between individuals of different populations. Nei's distance is one of the most frequently used measures of genetic distance for molecular data (Nei, 1978). UPGMA (unweighted pair group method arithmetic) dendogram was constructed based on Nei's (1978) unbiased genetic distance. POPGEN v1.31 was used to estimate the genetic distance between populations.

3.7.5 Consensus tree

Dendogram of genetic relationships among populations was constructed using Phylogeny Inference Package (PHYLIP) version 3.67.

3.7.6 Analysis of Molecular Variance (AMOVA)

ARLEQUIN software was used to carry out the AMOVA analysis to determine the genetic structures among populations. F_{is} which is known as inbreeding coefficient indicates the variance within individuals relative to their populations. Pairwise genetic distance is based on F_{ST} for a short divergence period.

3.7.7 Histogram

STRUCTURE v2.2 (Pritchard, Stephens & Peter, 2000) was used to investigate the genetic structure of the populations regardless of their geographical locations.

CHAPTER 4

RESULT

4.1 Microchecker analysis

Ten primers were chosen to study the genetic diversity of all the populations. Therefore, after the DNA was extracted, PCR optimization was done to standardize the protocol. The scored results were then run through several data analysis softwares to interpret the data.

During microsatellite amplification through polymerase chain reaction (PCR), one of the errors which could happen is the presence of null allele. This occurs due to the failure of one or more alleles to amplify during PCR. The result indicated that null allele is absent in four loci namely Mbr 7, Mbr 8, Mbr 10A and Mbr 10B.

Table 4.1: The results of Microchecker analysis for the presence of null allele.

No.	**Locus**	**Presence of null allele**
1	Mbr 1	Yes
2	Mbr 3	Yes
3	Mbr 4	Yes
4	Mbr 5	Yes
5	Mbr 7	No
6	Mbr 8	No
7	Mbr 10A	No
8	Mbr 10B	No
9	Mbr 11	Yes
10	Mr 78	Yes

4.2 Allele frequencies

Allele frequencies at 10 different microsatellite loci in *M. rosenbergii* samples from four different localities are presented in Appendix B. No population-specific alleles were observed for eight of the loci except for Mbr 11 and Mr 78 for population Kg Acheh. The private allele size observed to occur among the samples for Mbr 11 and Mr 78 are 363bp and 117bp. The specific allele sizes are marked in the graphs shown in Appendix B and C. More samples have to be collected from Kg Acheh in order to confirm the loci as a population specific marker as only 30 samples were used during this study. The common allele size for Mbr 11 for population Teluk Kumbar, Pantai and Sg Perak is 277bp while for Mr 78 is 147bp.

Loci Mbr7 possessed similar allele sizes where the most frequently occurring allele are 262, 274 and 275 at all four localities, as shown in Appendix B and C. Referring to graphs in Appendix B, all loci were found to possess rare alleles at low frequencies (less than 0.05) at mostly all populations. This pattern may be because all the microsatellite primers used in this study were isolated from populations in other countries. The allelic variance among the populations may be the factor causing the difference when used on Malaysian populations. To further analyze the allelic differences several other analysis were carried out.

Referring to Appendix B, Mbr 4 exhibited a high random fluctuation of allele frequency with high number of unique alleles. Most of the frequencies were lower than 50%. This may be due to the small number of individuals where inbreeding effects is evident. This has to be proven with more individuals from the same population to validate

the present data.

4.3 Heterozygosity, F_{is} values and HWE

Table 4.2: Level of heterozygosity, number of alleles, HWE probability values and Fis values.

		Teluk Kumbar	Kg Acheh	Pantai	Sg Perak
Mbr1	H_o	0.3750	0.5937	0.4375	0.4375
	H_e	0.9260	0.9682	0.9250	0.9573
	P	**0.0000**	**0.0000**	**0.0000**	**0.0000**
	n_e	8.7149	11.0108	4.2227	3.7236
	n_a	15	18	11	9
	F_{is}	0.5989	0.3905	0.5310	0.5469
Mbr3	H_o	0.5937	0.3125	0.4687	0.6562
	H_e	0.8745	0.9667	0.9330	0.9241
	P	**0.0000**	**0.0000**	**0.0000**	**0.0000**
	n_e	6.1687	13.7450	11.0108	6.8267
	n_a	16	20	20	14
	F_{is}	0.3245	0.6802	0.5016	0.2931
Mbr4	H_o	0.4687	0.6250	0.6562	0.4062
	H_e	0.9389	0.9806	0.9196	0.8353
	P	**0.0000**	**0.0000**	**0.0000**	**0.0000**
	n_e	5.4180	8.8276	2.9767	3.7996
	n_a	17	16	11	10
	F_{is}	0.5047	0.3663	0.2896	0.5176
Mbr5	H_o	0.7187	0.5312	0.5937	0.2500
	H_e	0.9533	0.9523	0.9171	0.8531
	P	**0.0000**	**0.0000**	**0.0000**	**0.0000**
	n_e	6.0592	9.9902	7.1860	4.4619
	n_a	14	16	12	12
	F_{is}	0.2490	0.4416	0.3562	0.7102

Mbr7	H_o	0.8438	0.7812	0.9062	0.8125
	H_e	0.8715	0.8263	0.7698	0.8859
	P	**0.0023**	**0.0000**	**0.0000**	**0.0000**
	n_e	7.0378	5.5956	4.1290	7.8168
	n_a	15	13	12	15
	F_{is}	0.0323	0.0554	-0.1806	0.0841
Mbr8	H_o	0.8750	0.9688	0.9375	0.6875
	H_e	0.8105	0.6949	0.7733	0.8258
	P	**0.0000**	**0.0000**	**0.0000**	**0.0000**
	n_e	4.9469	3.3032	4.1881	3.6377
	n_a	14	11	11	16
	F_{is}	-0.0809	-0.4029	-0.2165	0.1698
Mbr10A	H_o	0.7188	0.8125	0.6562	0.6250
	H_e	0.8318	0.8244	0.8730	0.8358
	P	0.2483	0.6353	**0.0016**	**0.0000**
	n_e	5.5202	5.5351	3.0983	2.4468
	n_a	14	17	8	12
	F_{is}	0.1378	0.0146	0.2512	0.2552
Mbr10B	H_o	0.8438	0.7188	0.7500	0.7812
	H_e	0.7535	0.7137	0.7019	0.6101
	P	**0.0309**	0.6747	0.2895	0.3864
	n_e	3.8715	3.4478	3.2354	2.5037
	n_a	13	9	13	10
	F_{is}	-0.1219	-0.0070	-0.06973	-0.2863
Mbr11	H_o	0.2812	0.5000	0.5000	0.3750
	H_e	0.9444	0.9052	0.9533	0.9533
	P	**0.0000**	**0.0000**	**0.0000**	**0.0000**
	n_e	9.4815	7.3935	8.4280	7.8467
	n_a	14	16	22	18
	F_{is}	0.7055	0.4516	0.4795	0.5907
Mr78	H_o	0.4687	0.5000	0.3750	0.4687

H_e	0.8640	0.9538	0.8120	0.8779
P	**0.0000**	**0.0000**	**0.0000**	**0.0000**
n_e	3.6120	6.5016	2.4885	3.8209
n_a	11	18	9	11
F_{is}	0.4614	0.4798	0.5421	0.4700
Mean statistics (by population)				
H_o	0.6187	0.6343	0.6281	0.5500
St. Dev	0.2123	0.1898	0.1923	0.1881
H_e	0.8768	0.8786	0.8578	0.8513
St. Dev	0.0654	0.1078	0.0870	0.0948
A_n	14.30	15.40	12.90	12.70
St. Dev	1.636	3.406	4.533	2.946

H_o: observed heterozygosity, H_e: expected heterozygosity, $n_{a:}$ observed number of alleles, $n_{e:}$ effective number of alleles, F_{is}: Inbreeding coefficient, $A_{n:}$ mean number of alleles, St. Dev: standard deviation, P: Probability values

Table 4.3: Mean observed and expected heterozygosity by locus

Locus	H_e	H_o
Mbr1	0.9651	0.4609
Mbr3	0.9524	0.5078
Mbr4	0.9600	0.5390
Mbr5	0.9481	0.5234
Mbr7	0.8723	0.8359
Mbr8	**0.8649**	**0.8671**
Mbr10A	0.8972	0.7031
Mbr10B	**0.7473**	**0.7734**
Mbr11	0.9633	0.4140
Mr78	0.9329	0.4531

Table 4.4 Overall population $F_{is.}$ Values are calculated using AMOVA and shows population specific F_{is} indices (1023 permutations).

Population	F_{is}
Teluk Kumbar	0.2977
Kg Acheh	0.2812
Pantai	0.2709
Sg Perak	0.3576

P-value= 0.00000 for all

Referring to Table 4.2, mean observed heterozygosity across all loci (0.5500 to 0.6343) for each population was lower than the mean expected heterozygosity (0.8513 to 0.88786). Highest observed heterozygosity (H_o= 0.9688) was observed at locus Mbr8 for the Teluk Kumbar population while the lowest observed heterozygosity (H_o= 0.2500) was exhibited at locus Mbr5 for Sg Perak population. Only two loci (Mbr8 and Mbr10B) out of 10, showed higher observed heterozygosity than the expected values, where for the rest of the loci observed heterozygosity was lower than expected (Table 4.3). The mean heterozygosity exceeds 50% which means most individuals are experiencing mild heterozygosity excess.

The inbreeding coefficient (F_{is}) is associated with heterozygosity. Populations exhibiting higher observed heterozygosity than expected at several loci show signs of out-breeding (Locus Mbr7 for Pantai; Mbr8 and Mbr10B for all population). F_{is} values lesser than zero (-ve) indicating excess of heterozygotes was evident in these cases (Table 4.2). Nevertheless when looking at mean F_{is} (AMOVA) values according to population Table

4.4, the locations stated above do not seem to be undergoing out-breeding. All four locations are exhibiting in-breeding effects (F_{is} values greater than zero) hence indicating deficiency of heterozygotes.

Referring to Table 4.2, the P-values of HWE for each loci and population under study, is shown with significant values in bold. Eight loci (Mbr1, Mbr3, Mbr4, Mbr5, Mbr7, Mbr8, Mbr11 and Mr78) exhibited consistent significance from HWE expectations ($P<0.05$) at all locations. Significant deviation at locus Mbr10A was found only in the Pantai and Sg Perak samples; at locus Mbr10B only in Teluk Kumbar samples. However, two populations, namely Teluk Kumbar and Kg Acheh were found not to have deviated from HWE at locus Mbr10A while population Kg Acheh, Pantai and Sg Perak did not deviate from HWE at locus Mbr10B.

Looking into allele frequencies, number of alleles per locus ranged from 8 to 22. The highest observed number of alleles (n_a) was found for the Pantai population at locus Mbr 11 with 22 alleles, while the lowest value was also found for Pantai population at locus Mbr10A with 8 alleles. The highest effective allele number (n_e) was 13.7450 in population Kampung Acheh and the lowest was 2.4468 in population Sg Perak. The values of effective number of alleles across all loci were more than 2 indicating existence of allele variation.

Population Kg Acheh seemed to be having the highest allele variability, with a mean number of alleles, 15.40. While, population Pantai and Sg Perak exhibited the lowest mean number of alleles, 12.90 and 12.70 respectively.

Looking at the overall result discussed so far, apparently Mean observed

heterozygosity value which is lower than expected substantiates the mean positive F_{is} values that in-breeding and correlates strongly with homozygosity is evident. Therefore, to confirm the genetic relationships among *M.rosenbergii* from different localities below stated statistical analysis result will be useful.

4.4 Genetic distance and cluster analysis

To further analyze the genetic variance among the four populations in order to assist the genetic improvement program, genetic distance and cluster analysis was carried out.

Table 4.5: The values of genetic identity (above diagonal) and genetic distance (below diagonal) between four populations of *M.rosenbergii.* (Nei's, 1978)

Population	Teluk Kumbar	Kg Acheh	Pantai	Sg Perak
Teluk Kumbar	****	0.3363	0.4664	0.5727
Kg Acheh	1.0897	****	0.2994	0.2741
Pantai	0.7627	1.2059	****	0.5407
Sg Perak	0.5574	1.2944	0.6149	****

Teluk Kumbar and Sg Perak are the most closely related populations with the highest value of genetic identity, 0.5727 and with the lowest distance value of 0.5574. These two populations were more genetically related. The highest genetic distance value was spotted between population Kampung Acheh and Sg Perak with the value of 1.2944.

4.5 Consensus tree

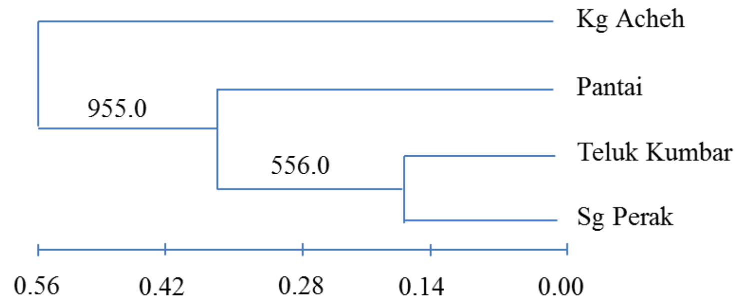

Figure 4.1: Consensus tree.

Referring to Table 4.5, the genetically identical populations which are location Teluk Kumbar and Sg Perak were clustered together in the same branch which showed 55.60% bootstrap value out of 1000 iterations. The consensus tree also shows close relatedness between population Pantai, Teluk Kumbar and Sg Perak. On the contrary, Kampung Acheh formed the out-group which in other hand, concludes that this population is most divergent from the others.

4.6 AMOVA

Table 4.6: AMOVA design and results (average over 10 loci)

Source of variation	Sum of squares	Variance components	Percentage of variance
Among populations	69.355	0.29356	6.35
Within populations	1091.391	4.33092	93.65
Total	1160.746	4.62447	
Fixation index	F_{ST}: 0.06348		

P-value = 0.00000
No. of Permutations = 1023

The results obtained from AMOVA showed that there is high variation among individuals of the same species. Based on the 10 polymorphic loci investigated, 6.35% of the variation was among populations and 93.65% of variation was within individuals of the populations.

4.7 Population pairwise F_{ST}

Table 4.7: Population pairwise F_{ST} distance

	Tk Kumbar	Kg Acheh	Pantai	Sg Perak
Tk Kumbar	0.00000			
Kg Acheh	0.07216	0.00000		
Pantai	0.04017	0.08409	0.00000	
Sg Perak	0.03902	0.09197	0.05033	0.00000

Generally, F_{ST} values were low throughout. The highest population pairwise F_{ST} value was 0.09197 between population Kampung Acheh and population Sg Perak which supports the genetic distance result and the consensus tree, while population Teluk Kumbar and population Sg Perak were most identical (0.03902). F_{ST} estimates in the range of 0-0.05

indicate very little genetic differentiation among populations (Hartl and Clark, 1997). Overall, the low F_{ST} values indicate that there is a mild genetic differentiation among the populations.

4.8 STRUCTURE

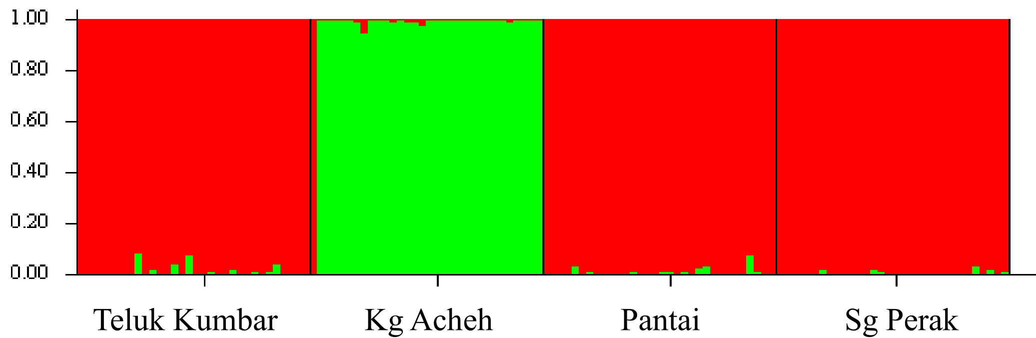

Figure 4.2: STRUCTURE histogram for the four base populations.

This schematic figure is plotted based on allele frequencies from what is assumed to be potentially unlinked loci, and uses these figures to identify which set of genes are decedents of which population. Kampung Acheh is more genetically divergent compared to other populations. Two distinctive population structuring is observed from Figure 4.2.

4.9 AMOVA between population Kampung Acheh and the grouped populations of Teluk Kumbar, Pantai and Sg Perak.

Since, population Kampung Acheh was clearly shown as an out-group compared to the others, the three subpopulations (Teluk Kumbar, Pantai and Sg Perak) were grouped as one. For the analysis, there were only two groups.

Table 4.8: AMOVA analysis between two distinct populations.

Source of variation	Sum of squares	Variance components	Percentage variation
Among groups	69.355	0.27272	5.90
Among individuals within groups	702.391	1.31269	28.39
Within groups	389.000	3.03906	65.72
Total	1160.7466	4.62447	

Fixation indices
F_{IS} : 0.30165
F_{ST} : 0.05897
P-value 0.0000

Previous results have illustrated population Kampung Acheh to be variant compared to population Teluk Kumbar, Pantai and Sg Perak. The latter populations were clustered together as one population as they belong to the same group (shown in STRUCTURE result Figure 4.2). There was evidence of differentiation among the two distinct *M. rosenbergii* subpopulations with a 5.90% variation among them. AMOVA analysis indicates a significantly weak genetic differentiation among the sampled populations in Peninsular Malaysia. With the significant value of 0.0000 ($P<0.05$), a higher percentage of variation was observed within individuals (65.72%).

Comparison of status between samples collected in 2009 and 2006/07.

4.10 Consensus tree (12 populations)

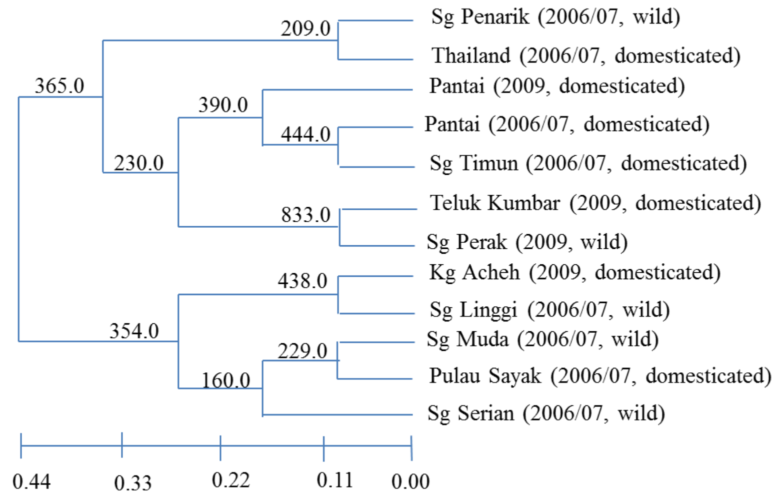

Figure 4.3: Comparison of status between samples collected in 2009 and 2006/07.

All the 12 populations are clustered into 2 main groups where populations from the same state grouped separately. The bootstrap values are very low.

The Sg Penarik and Thailand populations were unexpectedly clustered together. These two populations are very much similar, although the Thailand population is a hatchery strain which clustered together with the wild Terengganu strain.

As for population Pantai (2009), Pantai (2006/07) and Sg Timun (2009), there was a significant differentiation as the sample Pantai (2009) was different compared to the other two populations collected in 2006/07 and 2009. Yet, there was a significant genetic differentiation ($P<0.05$) between these domesticated and the wild Negeri Sembilan

populations. On the contrary, population Teluk Kumbar (2009) and Sg Perak (2009) were clustered together with population Pantai (2009), Pantai (2006/07) and Sg Timun (2009). Besides, population Kg Acheh (2009) and Sg Linggi (2006/07) are clustered together along with population Pulau Sayak (2006/07), Sg Serian (2006/07) and Sg Muda (2006/07).

4.11 AMOVA (12 populations)

Table 4.9: AMOVA design and results (average over 8 loci)

Source of variation	Sum of squares	Variance components	Percentage of variance
Among populations	368.807	0.5138	14.36
Within populations	2145.43	3.0649	85.64
Total	2514.237	3.5787	

F_{ST}: 0.14357

P-value = 0.00000

No. of Permutations = 1000

The results obtained from all 12 populations based on 8 loci indicate that the variation was higher within populations; 85.64%, while the variation among population was 14.36%.

Table 4.10: Population pairwise F_{ST} (12 populations)

Table 4.10: Population pair wise F_{ST} distance

		2009				2006/07							
		Teluk Kumbar	**Kg Acheh**	**Pantai**	**Sg Perak**	Thailand	Pulau Sayak	Pantai	Sg Timun	Sg Serian	Sg Penarik	Sg Linggi	Sg Muda
2009	**Teluk Kumbar**	0.00000											
	Kg Acheh	0.09642	0.00000										
	Pantai	0.10401	0.12593	0.00000									
	Sg Perak	0.07962	0.14606	0.13664	0.00000								
2006/07	Thailand	0.12285	0.13425	0.13837	0.16527	0.00000							
	Pulau Sayak	0.17876	0.10233	0.17494	0.22176	0.16890	0.00000						
	Pantai	0.13389	0.17688	0.13504	0.16861	0.16513	0.21380	0.00000					
	Sg Timun	0.12285	0.15012	0.10708	0.14951	0.11368	0.17641	0.11359	0.00000				
	Sg Serian	0.15527	0.13373	0.19513	0.17740	0.16888	0.14449	0.18828	0.17090	0.00000			
	Sg Penarik	0.12625	0.15512	0.13524	0.1387	0.10645	0.15069	0.16806	0.10494	0.13246	0.00000		
	Sg Linggi	0.09982	0.07218	0.14427	0.15237	0.14167	0.12919	0.20394	0.16449	0.14542	0.14890	0.00000	
	Sg Muda	0.14553	0.08647	0.12613	0.1783	0.14658	0.10519	0.18863	0.13022	0.12796	0.10913	0.11624	0.00000

The value of divergence is greatest between population Sg Perak (2009) and Pulau Sayak (2006/07), with the value of 0.22176. While population Kg Acheh (2009) and population Sg Linggi (wild, 2006/07) are least divergent, 0.07218.

4.13 STRUCTURE (12 populations)

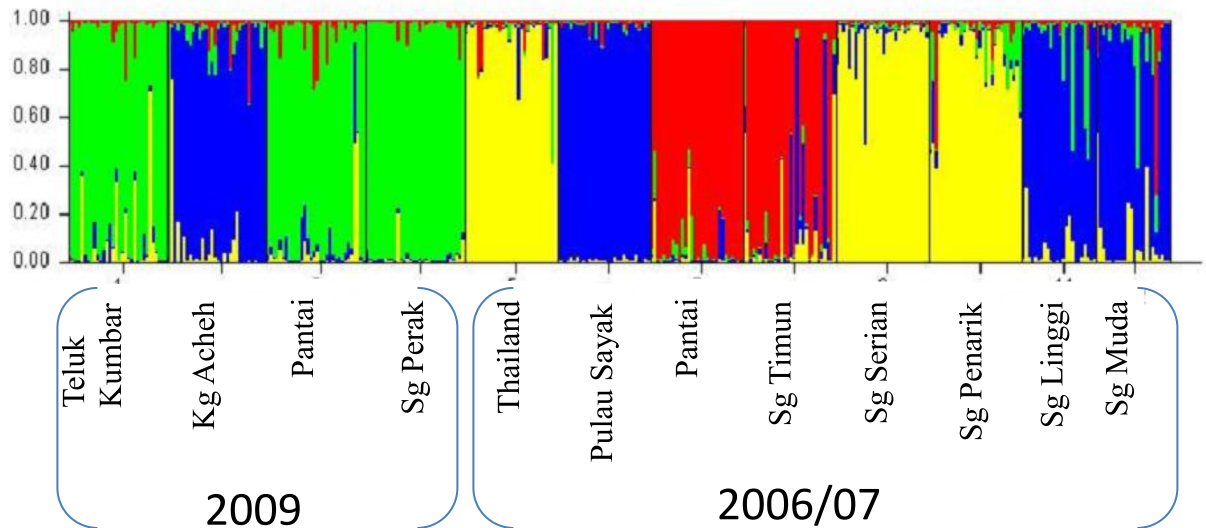

Figure 4.4: STRUCTURE analysis for all 12 populations.

There is minimal gene flow among the populations as they are divided into four separate groups. Population Sg Timun (2009) shows evidence of mild mixed lineage with population Kampung Acheh (2009). Population Sg Penarik (2006/07), Thailand (2006/07) and Sg Serian (2006/07) show very similar patterns and are clustered as one group. These locations are geographically separated, giving rise to the possibility of human transfer. Population Teluk Kumbar, Pantai and Sg Perak all collected in 2009 indicate the same genetic pattern. Population Kampung Acheh (2009) had similarity with population Sg Linggi (2006/07), Pulau Sayak (2006/07) and Sg Muda (2006/07).

4.14 AMOVA between samples collected in 2009 and 2006/07 (all populations clustered into two separate groups.)

Table 4.11: AMOVA analysis between populations fragmented into two groups (2009 and 2006/07).

Source of variation	Sum of squares	Variance components	Percentage variation
Among groups	49.649	0.13729	3.80106
Among individuals within groups	1639.588	1.15709	32.03653
Within groups	825	2.31742	64.1624
Total	2514.237	3.6118	

F_{IS} : 0.33302
F_{ST} : 0.03801
P-value: 0.0000

According to Table 4.8, the percentage of variation is only 3.80 between the samples collected in 2009 and 2006/07. While a higher percentage of variation was observed within individuals of a group regardless of the population (64.16%). The variation among individuals within groups was 32.04%.

CHAPTER 5

DISCUSSION

5.1 Hatchery management (Experiment and observation)

The first batch of berried females was brought in from Kuala Klawang farm where they originated from the Sungai Timun wild population. A complete cycle from a hatchling up to a juvenile takes about 30 days (larval stage 1 to stage 11 and post-larvae). The hatchlings growth was tremendously good where tank 5A reached post larvae (PL) stage on day 20. The other tanks were doing well until there was sudden mortality in all tanks at post larvae stage. The cause of rapid mortality was suspected to be *Vibrio* infection but was not confirmed as the samples could not be sent for further analysis on time.

The second batch of berried females was screened for disease and were tested positive with Infective Hypodermal and Hematopoeitic Necrosis Virus (IHHNV) positive. According to resources from professional researchers from the private sector, IHHNV spreads through the female. Also stated was that IHHNV is a marine shrimp viral disease. Its presence in a fresh water species was volatile where further diagnosis and study is being carried out by National Fish Health Research Centre (NAFISH). The possibility of the virus to mutate and inhabit freshwater species is still an untested fact. In these cases, salinity is always seen expected to be a limiting factor for these marine pathogens might not survive in fresh water. IHHNV can be severe and fatal to the shrimp when it is infected at its very early stages as an embryo or larvae. Healthy broodstock will produce offspring with normal growth even in an IHHNV infected environment, provided that they are stocked as juveniles in the ponds (CTSA, 1996). Since the year of IHHNV emergence in 1981 to

2001, the total product loss is estimated to be around 0.5 to 1.0 billion dollars (Lightner, 2003). As a normal practice, ponds are always disinfected with lime after harvest in order to kill any existing pathogen in the system; high pH is used to kill viruses.

In the episode of IHHNV, hatchlings crashed before they could reach the mid cycle period. Therefore, the cycle was repeated a few more times with more caution.

Based on the outcome of optimal tank density experiment, overpopulated tanks showed symptoms of cannibalism when the feed provided are not sufficient for their growing appetite. 60,000 larvae per tank was seen best for survival rate. Also the jumpers (larvae hatched during the first day and are most active compared to the late hatchlings) growth is more rapid where their size increased tremendously within a few weeks.

Furthermore, microscopic observation was made from day one of the stage 1 larvae. The morphological characteristics and their feeding behavior were studied every day after the feeding times. This observation enabled us to compare and contrast the feed intake of the larvae. There were two types of diet involved in this hatchery management; artemia and egg custard with special formulation. For the first few days, larvae were fed with artemia as their eye was still under development. Artemia which collides with larvae enable them to capture and feed on them. The larvae were fed four times with artemia and egg custard. The ingredients involved in the making of egg custard is egg, trace minerals, fish oil, spirulina, calcium carbonate, baby milk powder and water.

Apart from feed information, their metamorphosis stages were observed first hand. Throughout the hatchery period, larvae need a certain amount of sunlight which is crucial

for molting. The transformation from larvae with sessile eyes to a juvenile resembling an adult is very significant. Once they reach juvenile stage, they are grown in nursery before being released into the earthen ponds. New and Valenti (2000), have stated that molting which is also known as ecdysis is a cyclic and continuous process where the individuals increase their size. Factors affecting their growth and survival have been deliberated during this three month intense study. The characteristics of this species were understood and taken into consideration before carrying out the base population structure studies.

5.2 Base population structure studies

The main reason for genetic improvement program to be given importance in fresh water prawn industry in Malaysia is poor growth performance among captivated populations. Broodstocks which has been kept in captivity for generations show significant difference in size when compared to the wild populations. Due to inbreeding depression the size of marketed prawn has tremendously decreased. Therefore, farmers harvest, yield per every season is badly affected. This has increased concern over the growth performance of domesticated stocks where the availability wild catch for broodstock enhancement for the farmers is reducing rapidly. Genetic improvement program is seen as the essential method to boost the genetic variance thus enhances the performance among this species enabling it to restore.

During PCR amplification, the possible complications which might occur are the long stretches of repeated units in the DNA strand could create a hitch to DNA polymerase compared to shorter stretches. Longer stretches are capable of producing misalignments in the nascent strand (Christiakov *et al.*, 2006). Therefore, it is thought that trinucleotide

repeats and above, reduces slippage during polymerase chain reaction. Fragment analysis was carried out to avoid any feasible mis-scoring of microsatellite alleles due to stutter band effects. There were only minimal modifications throughout the methodology to ensure the quality of the result. The outcome of data analysis could be further interpreted to understand the base populations' genetic variation.

Freshwater species show high level of genetic population differences than marine species due to higher migration and gene flow in the latter group (Carvalho and Hauser, 1995). Less genetic drift is seen amongst marine species due to the vast open sea. Genetic difference between subpopulations will evolve in time although; there is little or no gene flow between them (Chakraborty and Leimar, 1987). Isolation by distance is one of the principles of population structure resulting in differentiation of genetic patterns. It explains that organisms exchange genes with geographically close areas or local interchange only (Baverstock and Moritz, 1996).

When conducting a genetic improvement program with prawn or other highly fecund species, it is important to obtain information on genetic divergence between the populations which will be crossed (Jerry *et al*., 2004). The mean number of alleles was seen to be highest within individuals of population Kg Acheh with the value of 15.40 (Table 4.2) while the other three populations showed a rather similar mean value.

A lower observed heterozygosity indicates the proportion of individuals in a population which are heterozygous at a locus has decreased. Heterozygote deficit may also be due to the mixing of genetically undetected divergent stocks with the present population (Gopalakrishnan *et al*., 2006). According to See *et al*., (2007), lower observed

heterozygosity could also be due to founder effect and presence of null alleles. It could be due to methodology bias (Jarne and Lagode, 1996). Based on Table 4.3, the observed heterozygosity for all populations over 8 loci is exhibited to be lower than expected. This is also substantiated by the presence of null alleles among 6 loci (Mbr1, Mbr3, Mbr4, Mbr5, Mbr11 and Mr78) in Table 4.1. Micro-checker results did indicate that null alleles maybe present in the fragment analysis data. It is rare for expected and observed heterozygosity to be equal, as it is only possible if the frequency of each allele is identical (McDonald, 2008). A comparison of the observed and expected heterozygosity over 8 loci deviated from HWE ($P<0.05$), which may have resulted from the small sample size (30 individuals) analyzed per population.

Referring to Table 4.5, Population Kg Acheh when paired with population Teluk Kumbar, Pantai and Sg Perak exhibited limited genetic identity with the value ranging from 0.2741 to 0.3363. Population Pantai, Sg Perak and Teluk Kumbar were clustered together as seen Figure 4.1. Domesticated samples from Pantai and Penang have been confirmed to be originated from some parts of Kedah and Perak. Despite the fact, approximately 6.35% variance (Table 4.6) was seen among populations which are geographically close. Therefore, it can be said that geographical distance does not really contribute towards genetic variance among populations. Crouau-Roy (1989), did state in his study that the genetic variance which occurred among populations within a close proximity may have resulted from low dispersal rates between populations. Therefore, the best pairs to be crossed for genetic improvement are population Kg Acheh and Sg Perak with the highest value for genetic distance, 1.2944, and Population Kg Acheh and Pantai (genetic distance =1.2059). On the other hand, as shown by the STRUCTURE result in Figure 4.2, all the other three populations were exhibiting a similar population structuring compared to

Kampung Acheh samples which was completely divergent. The genetic identity has substantiated the STRUCTURE result. So, individuals from Kampung Acheh can therefore be the core population to be crossed with the other three populations to produce a significant hybrid vigor. Also, the STRUCTURE result showed the four populations being grouped into two. Referring to Table 4.8, the genetic variation among the two groups is 5.90% which indicates mild genetic differentiation among the populations of both groups.

Information on genetic divergence between populations allows for a strategy to maximize the genetic response for traits of performance, while minimizing the detrimental effects of inbreeding (Jerry *et al*., 2004). Referring to Table 4.3, the genetic distance was reported to be the highest among population Kg Acheh and Sg Perak at a pairwise F_{ST} value of 1.2914. Meanwhile the other pairs showed a relatively low genetic distance compared to the above mentioned pair. High levels of gene flow and migration increases the similarity between the populations (Neigel, 1997).

Based on the mean F_{is} values over 10 loci (ranging from 0.2709 to 0.3576) given in Table 4.4, the positive values indicate more inbreeding effect than expected within a population than gene flow. Referring to Table 4.2, there are negative F_{is} values seen in loci Mbr7, Mbr8 and Mbr10B indicating low inbreeding effect. The difference between the values above may be due to the small population size (30 individuals). Besides, potential adaptability of a population decreases due to inbreeding when heterozygosity is lost (Ferguson *et al*., 1995). This may be due to the decrease in the population size where females have limited number of males to choose for spawning. Many wild populations are over-fished and exploited by fishermen and villagers as a source of income as they are highly priced and as a source of protein, respectively. Such harvest has an impending effect

on causing three types of genetic change such as alteration of population subdivision, loss of genetic variation and selective genetic changes (Allendorf *et al*., 2008; Kuparinen and Merila, 2007). They also added that genetic has to be considered into the management as it is fundamental to sustain productivity of harvested populations.

5.3 Comparison of status between samples collected in 2009 and 2006/07.

All 12 populations were divided into two clusters as shown in Figure 4.3 with very low bootstrap values. Consequently, the percentage of variance among the 12 populations was only 14.36% (Table 4.9). But, the greater variation among individuals within the populations with 85.64% supports the difference in the populations structuring as shown in Figure 4.4. If a larger population size has been used in this experiment, the figure presented by Voris H.K. in his paper published in the year 2000 can be put forth for questioning. He stated that, four large rivers of west coast of Peninsular Malaysia (Sungai Perak, Sungai Muar, Sungai Lenek and Sungai Bernam) and major Sumatran rivers (Sungai Simpang Kanan, Sungai Panai, Sunagi Rokan and Sungai Siak) originated from a river running south. This affirms that the rivers were formed form the same origin thus holds similar ecosystem pattern. Peninsular Malaysia, Java and Sumatra were connected as a large land mass estimated around 17,000 years ago (Appendix D) before the formation of Straits of Malacca. All of the existing rivers in west coast of Peninsular Malaysia now would have originated from the four large rivers of Peninsular Malaysia and Sumatra. Apart from the Pleistocene sea levels in Southeast Asia illustrating the impact of geological time over significantly changing sea levels, Malaysia's location explains most of the flood occurrences due to cyclical monsoons which occurs during the local tropical wet season. During floods, there is a possibility of giant freshwater prawn migration where they can

flow into a new river system. Therefore, the similarity among population of west coast of Peninsular Malaysia used in this study is inevitable.

The giant freshwater prawn industry which has boomed in Malaysia since 1980's, may have experienced population admixtures for many generations as the wild stock throughout Peninsular Malaysia has been depleting due to over-harvesting and pollution where the latter is an apparent scenario in Malaysia and the former as a good protein source and income for poor villagers. Many farmers may have obtained broodstocks from the fishermen and reared them in captivity for a very long duration. However the introduction of different set of populations to enhance the new generation may have contributed towards the allele diversity and identity among the populations. But if the broodstocks have been maintained for many generations (approximately more than 30 generation), inbreeding effects are to be expected. More wild or genetically different hatchery populations in Malaysia should be exploited or taken into consideration for genetic improvement program to enrich the allelic richness of the existing stocks for an improved giant fresh water prawn production. However, if broodstock samples could be obtained from countries to have the 'western' form of *M. rosenbergii*, it would be of a great break-through for the genetic improvement program as the genetic pool can be enriched for diversification.

Referring to Table 4.11, the higher variation among individuals within populations (64%) may be due to mixing of gene pool as different populations are being released and introduced into the river system through stock enhancement, human transfer, escapees from the hatchery and also due to natural selection to survive in an evolving environment or they could all be from a fairly homogenous stock.

All the 12 populations which were divided into two groups according to the time of collection only exhibited a low interpopulation variation, 3.80% (Table 4.8). Also, referring to Table 4.7, most of the pairs exhibit a high pairwise F_{ST} value whereby reducing the possibility of genetic drift and increasing the likelihood of homogeneity among the populations from different time frame ranging from 0.07218 to 0.22176. There are other factors causing increased genetic identity among populations. Human activities such as introduction of hatchery reared strains, over exploitation and transplantation of exotic stocks do molt the genetic variation (Ferguson, 1995). Besides, larval drift or adult migration (Ungfors *et al*., 2009) can be possible where the life cycle of the prawn involves brackish water systems. *Macrobrachium rosenbergii* is also known to survive in salt water up to certain salinity.

Apart from the above stated possibilities, there is information provided by the farmers regarding, a few parties restocking the Malaysian rivers with juvenile *M. rosenbergii* which can directly affect the genetic pool of the populations. The origin of these restocked individuals is unknown, as, these juveniles could have been fished from a different location in the country, obtained from a hatchery where the source of brood stock may have come from neighboring countries or a mixture of domesticated adults obtained from different localities. Several setbacks caused by the release of reared organisms into natural environment are interbreeding and competitive exclusions (Iguchi *et al*., 1999; Hindar *et al*., 1991).

An example would be the genetic identity between Sg Penarik and Thailand strain collected in 2006/07 showed relatedness with the F_{ST} value of 0.10645. STRUCTURE result indicated resemblance between the Sg Penarik, Thailand and Sg Serian strains. A

farmer through personal communication expressed his interest in bringing in the wild Sg Serian strains for genetic manipulation. This is because their body is bigger than the carapace which is vice versa among strains in Peninsular Malaysia. On the other hand, price of the post larvae in the current Malaysian market is RM0.06. Some resources have revealed that the post larvae from Thailand is only sold for about 2 to 4 cents each where it has attracted several farm operators.

Although the availability of cheap larvae benefits farmers, there are unseen disadvantages. Importing diseased post larvae is the hidden risk in such deals where it can destruct the freshwater prawn industry in Peninsular Malaysia. Therefore, screening the stocks prior to spawning can help overcome such incidents. Some hatchery operators who are currently working hand in hand with institutional experts do screen their broodstock for diseases as suggested before any genetic improvement program is carried out. The invention of a disease test-kit for *Macrobrachium rosenbergii* would definitely set a milestone in the Malaysian giant freshwater prawn industry.

CHAPTER 6

CONCLUSION

Freshwater prawn is seen as a cash crop in Peninsular Malaysia. Apart from management skills in freshwater prawn hatchery and grow-out, genetic aspects should be taken into consideration in improving the performance of these aquaculture species. Developing populations with broad genetic bases through improved crossing methods will permit continuous accumulation of favorable alleles. In conclusion, population Kampung Acheh can be one of the base populations to be crossed with population Teluk Kumbar, Pantai and Sg Perak, as there is significant genetic divergence between these two groups. The Kg Acheh hatchery practices a high number of founder parent population for breeding purposes where they use 150 females and 50 males. This significantly increases the genetic variability among the fries. On the other hand, there is low genetic variability among population Teluk Kumbar, Pantai and Sg Perak due to small number of broodstocks. The allelic similarity was observed among population Sg Perak and Pantai. Therefore, monitoring the genetic variability among wild and hatchery reared stocks is crucial as it is to determine and maintain the variation throughout the process (Primmer *et al*., 1999).

Coefficient of parentage on the basis of line pedigree is useful to evaluate genetic relatedness. Therefore, parentage assignment can be conducted further to enrich the genetic diversity study among Malaysian *Macrobrachium rosenbergii* species. In general, genetic variation between all the populations can be due to migration, drift and selection Excessive loss of genetic variability through inbreeding should be avoided (Ciftci and Okumus, 2002). Wild stock could be an important source for genetic improvement of culture stocks in the future (NACA, 2003). Information was obtained for genetic improvement and status

comparison between 2006/07 and 2009 samples for this study using 9 sets of primers extracted from the ‘western’ form of *Macrobrachium rosenbergii* and 1 from the ‘eastern form’. In a nutshell, the level of homogeneity and divergence among all 12 populations collected three years apart can be due to human transfer where it also involves introduction of new population into a different ecosystem.

Further studies involving greater numbers of populations from various geographical locations within East and West Malaysia would be necessary to construct an overall population structure in indigenous *M. rosenbergii* populations. In addition to that, additional microsatellite markers should be isolated and characterized from Malaysian giant freshwater prawn population and should be compared with existing markers by carrying out amplification on indigenous and other ‘western’ form of *Macrobrachium rosenbergii* to investigate the specificity of different localities. This is crucial as the allele frequency shown in Appendix B and C, indicated no population-specific alleles except for loci Mbr 11 and Mr 78 for population Kampung Acheh with allele sizes 363 and 117 respectively. Therefore, more samples from Kampung Acheh need to be retrieved and tested with these two loci to confirm whether the loci can be accepted as a population marker.