CHAPTER ONE

INTRODUCTION

The world population has grown tremendously over the past two hundred years since it reaches one billion in numbers since 1804. United Nation predicted that on October 31st 2011, the 7 billion human beings will be born to this world. The seven billion figures are eye-catching, but the reality behind it is worrying. Irrespective of where the baby is born, it will add to the 9 billion people who will inhabit earth in 2050 (UNFPA, 2011). Food production will have to increase by 70% to feed the extra 2 billion, says the UN's Food and Agriculture Organisation. Failure to do so will result in global hunger and malnutrition.

Marine and aquatic species are one of the cheapest sources of animal protein that can be considered to replace meat and poultry in the human diet as they can be harboured limitless from the sea and the freshwater. However the demand of a growing population now, far out strip the sustainable yield from the sea and freshwater. At the same time, fishing has become more industrialised and the wild stock of all seafood increasingly be depleted (FAO, 2010). Global consumption of marine and aquatic species has increased from 40 million tonnes in 1970 to 86 million tonnes in 1998 (FAO, 2000), reaching 145.1 million tonnes by 2010 (FAO, 2010).

As fisheries become depleted, many fishermen and governments respond by expanding aquaculture sectors. The total aquaculture production has increased from 10 million tonnes of fish in 1984 to 38 million tonnes in 1998 (FAO, 2000) with a growth rate at 11% per year. However, some of the non-government organizations and environmental groups proposed that

the rapid growth of aquaculture has led to environmental problems and conflicts over limited resources. One of the problems is the loss of mangrove forests (Naylor *et al*., 2000) that claimed to be important in maintaining the coastal ecosystems. Another problem that has been raised is "Biological pollution", a term used to describe the potential effects of genetically or ecologically introduced aquaculture species in natural population.

Aquaculture in Malaysia has been developed tremendously over the years right after the Government formulated the first National Agriculture Policy (NAP) in the early 1980's. Department of Fisheries (DoF) Malaysia is one of the government agencies that is responsible for enhancing food security which lead to an increase in aquaculture production and contribution, and increase in export value like one of the policies in third NAP (1998-2010). As a result, export value had increased from RM 2,112.5 million in 2008 to RM 2,328 million in 2009, and the main export was the prawns (51%) based on a report from DoF. *Macrobrachium rosenbergii* is one of the freshwater prawn species that were exported. In comparison to shrimps, the export quantity of this species is relatively small, but the high price of prawns makes it one of the income generators for many fishermen as most of the supply are consumed locally.

The giant river prawn, *Macrobrachium rosenbergii*, or locally known as "Udang galah" in Malaysia (Uno *et al.*, 1969) is ranked as the sixth largest aquaculture species in Asia based on volume (Mather *et al.,* 2008). For the past fifteen years, *Macrobrachium rosenbergii* has been the focus of aquaculture studies upon realizing its commercial value (See *et* al., 2008). Extensive studies have been done on *Macrobrachium rosenbergii* in order to enhance their stock, increase their productivity, and improve the prawn traits which will help to ensure long

term sustainability, prevention from disease and increase the genetic gain rate. These traits will also lead to lower cost production (Thanh *et al.,* 2009).

The demand for *Macrobrachium rosenbergii* has increased over the years due its popularity. Production from the wild stocks had increased from 5,246 tonnes in 1984 to approximately 130,000 tonnes in 1998 (FAO, 2000), with Asia accounting for more than 98% of global production. The culture of this species has expanded rapidly not only within Asia but also in regions far remote from the natural distribution of the species. These prawns are now cultured in at least 43 countries across five continents and have become the top 31 most captured freshwater species (FAO, 2000). However, the wild population that lives naturally cannot fulfill the commercial demand. Overharvesting of the prawns has caused some depletion of this species in all the river line systems all over the world (New *et al*., 2000). This kind of phenomenon has been seen in the river line system in Malaysia. One of the rivers affected by this problem is Sungai Timun in Negeri Sembilan, Malaysia. This river used to have an abundant supply the freshwater prawn, *Macrobrachium rosenbergii*. However, a report written by the Department of Fisheries (DoF) in 2007 showed a decline of this species (Zulkafli *Unpublished data*, 2007). It had been estimated that only around 1800 kg of freshwater prawn were caught in 2006 (Zulkafli *Unpublished data*, 2007). In addition to the issue of overharvesting, pollution is thought to be another key factor that is causing the decrease of prawn landing. It's important for DoF to take action in tackling this problem before it becomes worse due to the economic value of freshwater prawns and the responsibility of having to fulfill government's policy. A restocking program of this species is vital and should be done as soon as possible in order to fulfill the demand for the prawns. This program is aimed at supplementing the dwindling wild prawn population with hatchery-produced offsprings. However, the genetic diversity of both populations will need to be determined before the

program can be implemented. This will ensure that both populations have little genetic variation to avoid domestication of cultured stock, causing the loss of genetic variation.

Several studies have been done to calculate the genetic differences of the prawns between all of the river line systems in Malaysia. A study using Random Amplified Polymorphic DNA Marker (RAPD) showed that there are only 0.0343 genetic distance between Sungai Timun, Negeri Sembilan and Sungai Perak, Perak (See *et al.*, 2008). This genetic distance value is negligible. Other studies done by using Random Amplified Microsatellites (RAMS) and Long Primer Random Amplified Polymorphic DNA Marker (LP-RAPD) also showed the low genetic distance between these two populations (Bhassu *et al.*, 2007). A value of 0.0743 was shown as the genetic distance between these two populations showed by a study of wild stock population by using cross amplified Microsatellite primers (See *et al.*, 2007). From these results, the Department of Fisheries Malaysia decided to restock the Sungai Timun population with the Sungai Perak population. The strategy to restock was aimed at combating the depletion of aquatic resources known as stock enhancement. Stock enhancement can be defined as supplementation of depleted wild fish and invertebrate stocks with individuals reared in aquaculture facilities or captured from other populations (McMillen-Jackson *et al*., 2005). However to make it more specific Ihssen *et al*. (1981) defines a stock as "an intraspecific group of randomly mating individuals with temporal or spatial integrity". Other researchers such as Larkin (1981) defined a stock as "a population of organisms which share a common gene pool, is sufficiently discrete to warrant consideration as a self-perpetuating system which can be managed". On the other hand, enhancements may be defined as supply-side interventions in fisheries which aim to directly increase fisheries production through the stocking of hatchery fish. However, the overall yield thus remains strongly dependent on natural processes beyond management control.

Although stock enhancement has been practiced on a large scale since the mid-nineteenth century, it has faced controversies on its effectiveness and possible adverse impacts on wild stocks. Domestication or displacement of the wild population by the broodstock is one of the problems that raises some concerns. Domestication can lead to divergence from the wild population and may produce two different subspecies (John E.Joyce *et al*., 2004). The hatchery environment, particularly during the earliest development stage of the species is very different from the wild environment and the selection pressure may also differ. For example, the specific local adaptation of certain temperature and chemicals is an important fitness with a genetic basis that will cause the loss of adapting genes. Domestication could also result in foraging behaviour and predator avoidance as the hatchery and wild environment differ greatly in the availability of food and the presence or absence of predators. Therefore, it is not possible that domestication can produce changes in morphological and behavioural traits important for survival.

As a result, researchers have to ensure that their broodstock should be from the same local population into which the offspring will be released or have very little genetic variation between both populations. This can be done through the genetic analysis using genetic markers as tools. Genetic analysis can detect variation as a result of differences in either DNA sequences or specific genes or modifying factors. A number of genetic markers have proven to be useful. These include microsatellites and mitochondrial DNA that was used in this study. The application of molecular markers has allowed the detection of genetic variation in species, population, and within the population. The genetic variation at the species level can show the differences between the species, while the variation at the population level can help determine genetic classes, genetic diversity and their evolutionary relationship with their ancestor. The genetic variability within the population is extremely useful to gather information on individual

identity, breeding patterns, degree of relatedness of genetic variation among them (Schierwater *et al.,* 1994).

Mitochondrial DNA was chosen as a genetic marker as it allows trading of broodstock populations lineages via maternal inheritance. On the other hand, microsatellites or simple sequence repeat (SSR) is a powerful tool with unique characteristics such as distributed abundant in genome, high level of polymorphisms, locus-specific co-dominant inheritance and transferability across related species and repeatability. Both of these markers are based on Polymerase Chain Reaction (PCR).

This thesis focused on stock enhancement program at Sungai Timun, Negeri Sembilan, and Sungai Perak, Kampung Acheh, Perak, Malaysia by looking at genetically and ecologically aspect. The genetic variation between the two populations was assessed using the microsatellites (Type I and Type II) and mitochondrial DNA marker. This study will assess the genetic structure of the sample, and whether or not the restocking program between the two different populations is successful. In addition, this study will document the changes and the effects of ecological factors such as water quality and benthic study to the development of the prawns in this river. It is aimed to evaluate the effectiveness of the stock enhancement program that has been done in the area.

1.1 RESEARCH QUESTIONS

1.1.1 Is the stock enhancement programme of the *Macrobrachium rosenbergii* of Sg.Timun Negeri Sembilan with Kampung Acheh Perak successful?

1.1.2 If successful, how much in percentage?

1.2 OBJECTIVES

1.2.1 To isolate and characterize new microsatellite markers for *M. rosenbergii.*

1.2.2 To apply the polymorphic microsatellites markers on the parental stocks, F1 stocks, and wild stocks, and to test the level of genetic variability and follow the Hardy Weinberg expectation.

1.2.3 To use benthic data to infer the carrying capacity of Sungai Timun, Negeri Sembilan.

1.2.4 To determine the productivity success of restocking program using mitochondria marker.

1.3 HYPOTHESIS

1.3.1 Stock enhancement of *Macrobrachium rosenbergii* in Sg. Timun have shown postive impact on restocking initiative.

CHAPTER 2

LITERATURE REVIEW

2.1 *Macrobrachium rosenbergii*

Macrobrachium rosenbergii (Figure 2.1) is a decapods crustacean with a hard shell. It belongs to a group loosely called "shrimp" or "prawn". For example, animals classified under genus *Macrobrachium* are generally referred to as freshwater prawns in Australia and freshwater shrimp in the United States of America (USA). In its statistical data, FAO refers to the genus *Macrobrachium* as freshwater prawns but also uses the word "prawn" for many species of marine shrimp, including the banana prawn (*Fenneropenaeus merguiensis*), the giant tiger prawn (*Penaeus monodon*) and the kuruma prawn (*Marsupenaeus japonicus*) (FAO, 2001).

Figure 2.1: *Macrobrachium* rosenbergii (Source: Private Picture)

2.1.1 Taxonomic classification

The classification for giant freshwater prawn:

Kingdom:	Animalia
Phylum:	Arthropoda
Subphylum:	Crustacea
Class:	Malacostraca
Order:	Decapoda
Suborder:	Pleocyemata
Infraorder:	Caridea
Family:	Palaemonidae
Genus:	*Macrobrachium*
Species:	*Macrobrachium rosenbergii*

Source from Dore and Frimodt (1987)

Some taxonomists recognize a western sub-species (found in the waters of the east coast of India, Bay of Bengal, Gulf of Thailand, Malaysia, and the Indonesian regions of Sumatra, Java and Kalimantan) and an eastern sub species (inhabiting the Philippines, the Indonesian regions of Sulawesi and Irian Jaya, Papua New Guinea and northern Australia). The sub-species of the west is known as *Macrobrachium rosenbergii dacqueti* (Sunier, 1925) whereas sub-species of the east is known as *Macrobrachium rosenbergii* (De Man, 1879) for the eastern

form. A study done by Mather and Bruyn, 2003 using 16S ribosomal RNA and mitochondrial DNA (mtDNA) has supported the fact that the Malaysian *Macrobrachium rosenbergii* is the western sub species.

2.1.2 Genetic content

On the genetic part, based on the cytological analysis studies done by (Das & Lakra, 1998), it revealed a modal diploid chromosomes number of 2n = 118 in both female and male individuals. The karyotype comprised 29 metacentric, 5 submetacentric, 3 subtelocentric, and 22 acrocentric pairs in males and 37 metacentric, 6 submetacentri, 1 subtelocentric and 15 acrocentric pairs in females.

2.1.3 Distribution

Macrobrachium rosenbergii is indigenous to the whole of the South and Southeast Asian area as well as in northern Oceania and in the western Pacific islands. The species requires brackish water in the initial stages of their life cycle (larval stage), and they are found in water that is directly or indirectly connected to the sea although some complete their life cycle in inland saline and freshwater lakes.

Since its import into Hawaii from Malaysia in 1965, where the pioneering work of Ling (1969) was translated into a method for the mass production of post larvae by Fujimura and Okamoto (1972), it has been introduced into almost every continent for farming purposes. *Macrobrachium rosenbergii* is now farmed in many countries; the major producers are Bangladesh, Brazil, China, Ecuador, India, Malaysia, Taiwan Province of China, and Thailand

(FAO, 2002). More than thirty other countries reported production of this species in the year 2000. In addition, there are also valuable captured fisheries for *M. rosenbergii*, for example in Bangladesh, India, and several countries in Southeast Asia (FAO, 2002).

2.1.4 Morphology

Freshwater prawn eggs of this species are slightly elliptical, and they are bright orange in colour until 2-3 days before hatching when they become grey-black. This colour change occurs as the embryos utilize their food reserves. The larvae go through 11 distinct stages (Uno and Kwon, 1969) before metamorphosis. Larvae swim upside down by using their thoracic appendages, and are positively attracted to light. Newly metamorphosed post larvae are generally translucent, and have a light orange pink head area.

Adults usually have a distinctive blue-green color, although sometimes they may take on a brownish hue. Color is usually the result of the quality and type of diet. Adult males are larger than the females, and the sexes are easily distinguishable (Figure 2.3). The second walking legs or claws (chela) and the head region of males are larger than those of the females. Three types of males have been identified, based upon external characteristics. Blue-claw (BC) males are easily distinguishable and are characterized by long, spiny blue claws. Two other classes of non-blue-claw males exist, orange claw (OC) and strong orange claw (SOC) males. The OC and SOC serve as a transitional stage before the male prawn reaches BC stage. Smaller OC males (< 10g) grow slowly but are more reproductively mature than are other OC males. However, males can reach a body size of 32 cm; females grow to 25 cm (New and Singholka, 1985).

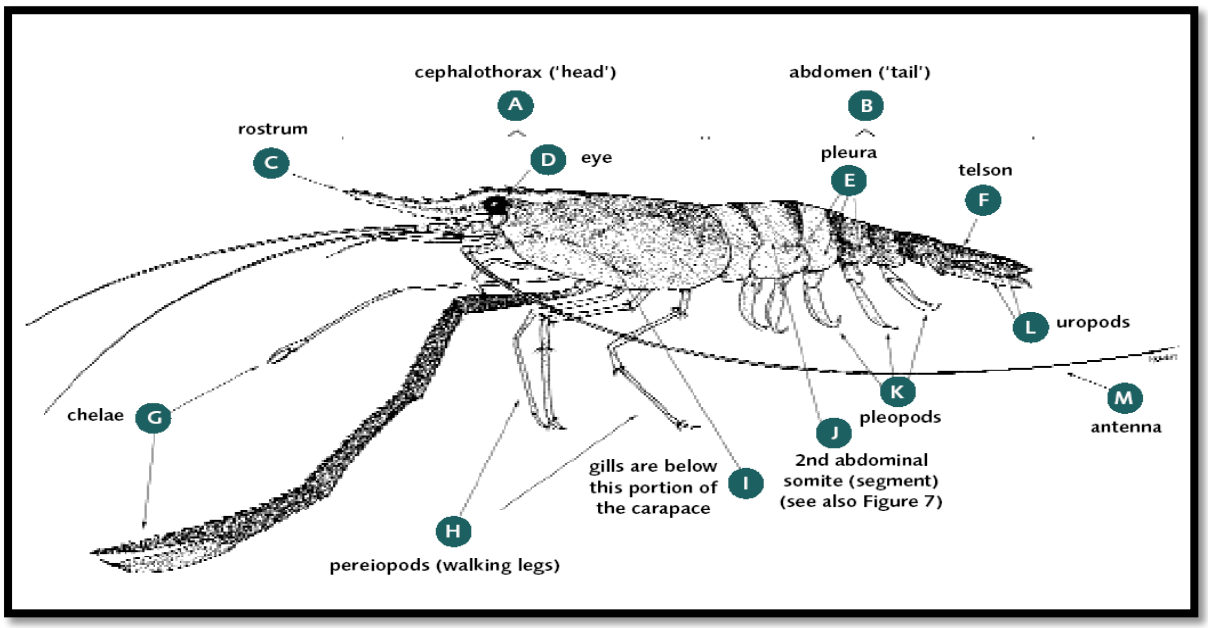


Figure 2.2: Morphology of *Macrobrachium* rosenbergii (Source: Emanuela D'Antoni)

Figure 2.3: Different morphology of *Macrobrachium* rosenbergii between male (big size) and female (small size) (Source: Private Picture)

2.1.5 Life Cycle

There are four distinct phases in the life cycle of the freshwater prawns, namely eggs, larvae, post larvae and adults. The mating process begin when the male deposits sperm into a gelatinous mass that is held underneath the body of the female, between her fourth pair of walking legs. Eggs are laid within a few hours after mating and are fertilized by the sperm contained in the gelatinous mass attached to the outside of the female's body. Eggs hatch approximately 20 to 21 day after spawning. Females carrying eggs are termed "berried females" (Figure 2.4). The colour of egg change from bright-orange color to orange, then brown, and finally gray about 2 to 3 days before hatching.

After hatching, the larvae cannot survive in freshwater beyond approximately 48 hours and migrate to brackish water with a salinity of 10 to 14 parts per thousand (ppt). The larvae will undergo 11 molts, each representing a different stage of metamorphosis. Following the last molt, larvae transform into post larvae. Transformation from newly hatched larvae to post larvae requires 15 to 40 days, depending upon food quantity and quality, temperature, and a variety of other water quality variables. Post larvae resemble miniature adult prawns. Post larvae can tolerate a range of salinities and migrate to freshwater upon transformation. Post larvae are translucent and may have a light-orange-pink head. As they change to the juvenile stage, they take on the bluish to brownish color of the adult stage. Juveniles (1-2g) are intermediate in size between post larvae and adults; however, no standard definition for the juvenile stage exists. After juveniles, the prawn will then transform to the adult stage (New, 2000)

Figure 2.4: "Berried" female of *Macrobrachium* rosenbergii (Source: Takuji Fujimura, obtained from Coreira *et al.*, 2000)

2.1.6 Nutrition

Prawns are omnivorous because it eats small fishes, aquatic worms, aquatic insects (benthos organisms), grains, plankton and algae. However, when prawns are sufficiently hungry, they can be cannibalistic. Nowadays, food has been formulated in a pellet form to feed the prawns reared in pond and aquarium.

2.1.7 Problems

The main issue that has been raised in cultured stocks is the loss of genetic diversity. This problem begins when the broodstocks from Malaysia were introduced to the Anuenue Fisheries Research Center in Hawaii in 1965. At that time, the broodstocks consisted of only

12 individuals (Hedgecock *et al*., 1979). Later, the broodstocks from Hawaii and SE Asia were introduced into many regions where *M.rosenbergii* was not indigenous, including North and South America, Africa, Europe, and parts of Asia (most notably China and Taiwan) to initiate culture industries (New, 2000). Thus, genetic attributes of these cultured stocks had cause a decline in the genetic diversity among the species. However, wild stocks can still provide an immediate source for solving this problem.

It's important to point out that some wild stocks of *M. rosenbergii* have been seen to decline rapidly in recent years, largely due to over-harvesting, habitat-loss, and increased pollution particularly in SE Asia. The decline had also been recorded in Bangladesh, India, Indonesia, the Philippines and Thailand (New, 2000), and the species is now believed to be extinct in Singapore, largely because of pollution and loss of natural habitat (Ng, 1997). To this end, the patterns and extent of genetic diversity that are present in wild stocks need to be adequately documented so that we can identify which stocks may carry unique genetic attributes, and prioritize conservation efforts.

2.2 Molecular markers

Molecular markers can be described as any polymorphic segments of DNA sequences, which are used by biologist as distinctive landmarks along the chromosomes. These markers give biologists signs or hints of the changes that occurred inside the cell due to environmental issues, mutation, etc. The resolving power of genetic markers is determined by the level of polymorphisms detected, so with DNA markers, it is theoretically possible to observe and exploit genetic variations in the entire genome. Genetic variation is an important factor in the process of evolution in natural populations, and it arises through mutations in conjunction with

the forces of migration, genetic drift, and various types of natural selection (Gunderina, 2003). There are three classes of molecular markers that are known to the geneticist community. The first one is Allozyme, a genetic variation at the level of proteins, which is directly encoded by DNA.

The second class is Nuclear DNA marker, which is quite common in our times. The popularity of this class of marker is being helped by its stability. In addition, the marker is technically easy to assay with the invention of Polymerase Chain Reaction (PCR) by Mullis in 1983 (Saiki, *et al.,* 1988). Nuclear DNA markers that may be applied in genome analysis and assignment of genetic variation can be generally grouped into two types based on their association functionality (O' Brien, 1991). Type I or functional markers are DNA segments, which are part of expressed DNA sequences. These markers are typically not very polymorphic because any variation can be a lethal aberration eliminating the survival of the carrier. The markers, therefore, show high evolutionary conservation. Examples of this type of marker are EST SSR, restriction fragment length polymorphism (RFLP), and single nucleotide polymorphism (SNP). Type II or nonfunctional markers usually do not have identifiable biological functions, and seem to show higher levels of polymorphism. Randomly Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), minisatellite, and microsatellite usually represent type II markers.

Mitochondrial DNA markers (mtDNA markers), the last class of marker is a piece of non-nuclear DNA located within organelles in the cytoplasm called mitochondria or known as a “power house” of the cell. This molecule (of which 93% is coding) encodes 13 proteins, 2 rRNA and 22 tRNA (Kochzius 2009). Mitochondria are almost independent from other neighboring organelles because they have their own DNA, and are replicating autonomously

from nuclear DNA. However, both nuclear DNA and mitochondrial DNA complement each other as most of the mitochondria proteins are encoded by the nuclear DNA. It is found that some nuclear DNA will not function without mitochondrial subunits (Brown, 1983).

Since all of these molecular markers have their strengths and weaknesses, choosing the perfect molecular marker for any research will have to be based on genomic abundance, level of polymorphism detected, locus specificity, reproducibility, technical requirements and financial investment. As a result, two classes of markers have been chosen for this study, Nuclear DNA (Microsatellites Type I and Type II Marker) and Non-Nuclear DNA Marker (Mitochondria DNA).

2.2.1 Microsatellites

Microsatellite are rapidly replacing RFLPs and RAPDS in most applications in populations biology, from identifying relatives to inferring demographic parameters (Jame & Lagoda, 1996) as its technically simple to operate. Microsatellite, also known as simple sequence repeats (SSR), consists of 1-6 bp tandem repeats, usually less than a 100bps long, and flanked by unique sequences. The primers are designed to anneal to these sequences, amplifying the repeat region in between (Tautz, 1989). Microsatellites tend to be evenly distributed in the genome on all chromosomes and all regions of the chromosome. They have been found inside gene coding regions (Liu et al., 2001), introns, and in the non-gene sequences.

Most microsatellite loci are relatively small, ranging from a few to a few hundred repeats. The relatively small size of microsatellite loci is important for PCR-facilitated

genotyping. Microsatellites containing a larger number of repeats are more polymorphic, though polymorphism has been observed in microsatellites with as few as five repeats (Karsi et al., 2002). Microsatellites are markers that require prior genomic library, thus it is very species-specific markers. Therefore, most research will focus on the isolation of polymorphic microsatellite markers. So far, microsatellites markers have been isolated in fresh water prawn from the various research groups (Bhassu et al., 2008, Divu et al., 2007, Chand et al., 2005, Chareontawee et al., 2006, Chareontawee et al., 2007, Sudsuk & Sodsuk, 1998). The development of highly polymorphic genetic markers such as microsatellites has provided an essential tool for identifying parentage relationships among individuals and obtaining pedigree information in aquaculture selection programs (O'Reilly &Wright, 1995; Ferguson & Danzmann, 1998).

There has always been an ongoing debate as to whether microsatellites are neutral or subjected to selection. If those loci adjacent or within genes that are expressed most probably would be under selection. Others that occurs in introns can vary depending on its location and neighbor genes, but most studies showed that microsatellites lie far from coding region, branding them as selectively neutral (Scribner & Pearce, 2000).

The high mutation rate of microsatellite loci gives them a far greater resolution than the more traditional molecular marker such as allozymes. Mutation rates can go anywhere from 0 to 8x10-3(Weber & Wong, 1993) and it is thought to be found at least once every 100 kbps in eukaryotes (Tautz, 1989). There are two mutation models that are thought to have given rise to microsatellites. The first one is Allele Mutation Model (IAM) (Kimura & Crow, 1964) which states that new alleles arise regardless of the numbers of repeats different. While the second model, Stepwise Mutation Model (SMM) (Ohta & Kimura, 1973), stated that a single repeat

unit is added or deleted for every single mutation making it possible that the same alleles might arise again like the one shown in one of the research in human (Valdes, Slatkin, & Freimer, 1993). Due to higher mutation rates, microsatellites are said to have high polymorphism property or high number of alleles so only a small number (10-15) of loci is usually required to distinguish between individuals. This allows the study of relationship between very closely related populations and individual.

Microsatellites are analyzed through the amplification of loci with locus specific primers in a PCR. The amplified fragments can be analyzed using high-resolution gel electrophoresis. In the past, the locus specific primers were radioactively labeled and visualized on photo-receptive sheets. Nowadays, the primers are labeled with a fluorescent dye, and analyzed in automated DNA analyzers. These analyzers can assess differences in fragment length of a single base pair, allowing very accurate genotyping. From the fragment lengths, the number of repeats at each locus can be determined. In addition, small amount of DNA are needed since no Southern Blot is involved.

A number of applications of this marker will depend on the codominant expression of microsatellites alleles. Codominant markers, in contrast to dominant markers, express all their alleles which allow the identification of homozygote and heterozygote. In microsatellites all alleles can be detected after PCR amplification and fragment sizing. Scoring both homozygote and heterozygote in populations allows the calculations of allele frequencies which lead to various analyses that are based on Hardy & Weinberg Equilibrium.

The codominant property has contributed to identifying organisms and their origins (Hansens, Kenchington, & Nielsen, 2001). Three distinct forms of *M. rosenbergii* have been

recognized. They are an eastern, a western, and an Australian form, based on morphology, allozymes microsatellites, and mitochondrial DNA (De Bruyn et al., 2004). It also helps in observing changes in genetic variability due to breeding program (Norris, Bradley, &Cunningham, 1999). Studies have been done to assist the selection of individuals in prawn-breeding programs, which would facilitate the domestication of the species. Since microsatellite are single locus marker, many population studies have been done using species such as Rainbow Trouts (Nielsen, et al., 1994) and Atlantic cods (Bentzen, et al., 1996). Several microsatellite studies have also demonstrated the ability to determine parentage, analyze pedigree, and identify species (Castro et al., 2004; McDonald et al., 2004). Microsatellites have been shown to be very useful for verifying pedigrees in prawn lines (Moore et al., 1999), and the pedigree tracing of hatchery populations (Herbinger et al., 1995; O'Reilly et al., 1998). Pedigree analysis and parentage assays are mainly done to deduce mating behavior and stock management. This proves to be very helpful since most newborn prawns cannot be tagged. Pedigree information is also necessary for the estimation of breeding values and the management of inbreeding. Inbreeding can be problematic in any closed breeding population because it can have notable deleterious effects on the animals (Kincaid, 1983, 1976). If the breeder has knowledge of the genetic relationships between individuals selected for use as broodstock, then inadvertent mating of close relatives can be avoided.

The major drawback of microsatellites is that they need to be isolated de novo from most species being examined for the first time. This is because microsatellites are usually found in non-coding regions where the nucleotide substitution rate is higher than in coding regions. Consequently, the strategy of designing universal primers matching conserved sequences, which was very effective for mitochondrial DNA (Kocher et al., 1989), is more problematic for microsatellites. However, the presence of highly conserved flanking regions has been

reported for some microsatellite loci in cetaceans (Schlotterer et al., 1991), turtles (FitzSimmons et al., 1995) and fish (Rico et al., 1996), allowing cross-amplification from species that diverged as long as 470 million years ago.

On the other hand, stuttering caused by slippage can occur in PCR especially when amplifying di-repeats. Many theories have tried to explain the phenomenon. One of the theories is mis-annealing after the strands segregate in replication due to repetitive nature of the sequence. Therefore, it will cause a few repeats added or removed from the segment , when it happen a few times where the re-annealing takes place (Alba, M. *et. al*, 1999). This phenomenon will cause problems during the population studies on species because it might hide certain alleles and homozygous excess. Some researchers claimed that stuttering could be reduced by working with larger repeats unit. This will also help to eliminate the need to use high-resolution gels and automated sequencer machines.

Null alleles which are not amplify alleles will also be a problem in microsatellite. It will underestimate heterozygosity and caused homozygosity excess. Many explanations for this, the most common one are mutations in the primer annealing site especially in the 3' end where the amplification starts (Kwok, et al., 1990). In addition, when PCR is involved in amplification, shorter and larger alleles are selected, , a phenomena that is known as "large allele dropout". Some of these alleles can be salvaged by adjusting contrast and DNA concentrations (Dakin & Avise, 2004).

2.2.2 Expressed Sequence Tags (ESTs)

Expressed sequence tag (EST), a type I marker is a small piece of DNA sequences with approximately 300-500 bp which are reversed transcribed from a cellular mRNA population. EST is generated by single-pass sequencing of complementary DNA clones obtained by reverse transcription of messenger RNA (Putney et al., 1983; Ferez et al., 2005). These markers are best known for physical mapping of genomes (Adams et al., 1991), gene expressions, and identification (Liu & Cordes, 2004) as it contains information that is acquired in expressed genes. This gives an insight as to what goes on inside each tissue, and the expression variation according to the development stage and changes caused by diseases. Expressed sequence tags (EST) – microsatellite (SSR) markers are developed from expressed genes, and utilized as transferable molecular marker since it is highly conserved (Perez et al., 2005). EST- SSR marker works well together since the EST provides gene identification whereas SSR expresses high level of polymorphism (Serapion et al., 2004). The advantages of the EST-SSR marker include the distributed abundant in genome, high level of polymorphism, locus-specific co-dominant inheritance, transferability across related species, and repeatability and clarity of scoring.

2.2.3 Mitochondrial DNA (mtDNA) Marker

Mitochondrial DNA is a single type of circular double stranded DNA including the light (L) strand and heavy (H) strand. The light strand is rich in cytosine while heavy (H) strand is rich in guanine making the size range between 17 – 17 kbps (Grey, 1989). During zygote formation, only nuclear DNA was inherited from the sperm. Therefore, mitochondrial DNA is known as maternally inherited, and both males and females receive their mitochondrial DNA

from the egg (Strachan & Read, 1999). There are several advantages of the mitochondrial DNA over nuclear DNA. According to Drake *et al.,* 1998, nuclear DNA undergoes relatively slow mutation compares to mitochondrial DNA. Thus, it requires longer nucleotide sequence than mitochondrial DNA in order to provide barcode capable of differentiating the different species. Out of the 37 genes that are available in Mitochondrial DNA, 6 different markers are developed which include protein coding region such as Cytochrome B, Cytochrome C Oxidase subunits (COI, COII, COIII) , ND1-6 (subunits 1-6 of the NADH reductase), non-protein coding region involving D-Loop control region, 12s rDNA and 16s rDNA (Sergio, 2000). Among all 6 markers available, protein coding region is considered as conserved markers while non-coding region such as D-Loop have much higher variability thus used in population studies (Brown *et al.*, 1996).

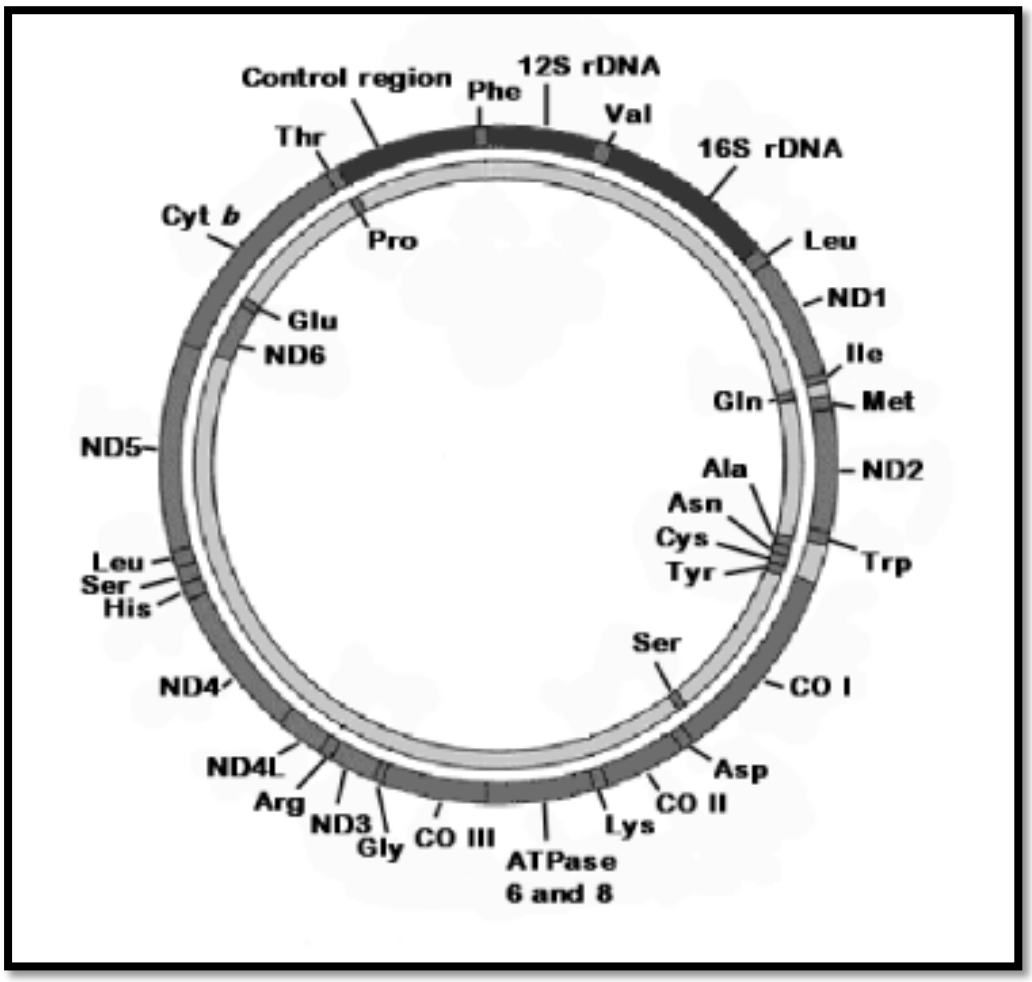


Figure 2.5: Schematic representation of the circular molecule of the "conserved" vertebrate mitochondrial genome. Genes outside and inside the circle are transcribed in H and L strands respectively. Protein-coding region are represented as follows; Cyt b – Cytochrome b; CO I, CO II and CO III – subunits I, II and III of the cytochrome oxidase; ND1 – 6 – subunits 1 to 6 of the NADH reductase. tRNA are represented by their three letter amino acid abbreviations . (Sergio Liz Pereira, 2000)

2.2.3.1 Cytochrome C Oxidase subunit I (COI)

Cytochrome C Oxidase subunit I gene, known as single short sequence of mitochondria DNA, which is able to code a large transmembrane protein found in the mitochondrion which is highly conserved among species. Cytochrome C Oxidase protein works as the terminal electron acceptor in the respiratory chain for reduction of oxygen to water (Waugh, 2007). A segment near the 5-Terminus of COI has been selected as a barcode region for some animal groups (Hebert *et al.*, 2003). This marker has become a popular marker in the identification of

stocks and analysis of mixed aquatic species. It also provides information on hybridization and introgression between aquatic species, and critical information for use in conservation and rehabilitation programs.

2.3 Importance of genetic studies on *Macrobrachium rosenbergii*

In Malaysia, wild stocks have been used to cultivate commercial cultures in many areas. Therefore, repeated harvesting from the wild stocks will eventually lead to extinction and loss of genetic diversity. Effective conservation may not be feasible unless data is available on genetic variation throughout its distribution range (Bhassu & Hassan, 2005). It's important to have awareness of the genetic makeup of wild stocks for conservation and breeding purposes, and there is a great necessity for one to assess these stocks for genetic information. Unfortunately, there isn't much available population genetics information on this prawn especially for the western form which is indigenous to Malaysia. Therefore, to overcome this problem, different genetic markers have been developed to identify genetic variations at the DNA level in *M. rosenbergii*. Charoentawee *et al.*, 2006 designed primer pairs for 20 microsatellite loci in the western form of *M. rosenbergii* which provide powerful tools for the conservation and management of wild stocks. In addition, eight single locus microsatellite markers were developed to characterize the Malaysian giant freshwater prawn, *M. rosenbergii*, by using the 5´ anchored polymerase chain reaction technique. These eight microsatellite loci were polymorphic with the levels of heterozygosity ranging from 0.6 to 0.8 (Bhassu *et al.*, 2008). This study will also develop microsatellite Type I and Type II loci to characterize western form of *M. rosenbergii.*

2.4 Stock Enhancement

Many scientists believe that stagnant production from capture fisheries means that aquaculture will play a major role in supplying the increased demand, albeit with a substantial increase in the use of fisheries resources (Tacon, 2003; Muir, 2005). Still, the potential for improved management of capture fisheries to increase yields, and thus complement aquaculture in filling the supply gap, is controversial (FAO, 2004). The possibility of increased production arises because several coastal fisheries no longer provide their potential benefits for two reasons: spawning biomass has been reduced below optimal levels, and the habitats that support fisheries production have been degraded (FAO, 2004). Reversing this mismanagement should improve production from these fisheries, but ecosystem level considerations suggest that management for conservation goals, such as restoring populations of top predators and the historical structure of food webs, will result in lower, rather than higher, overall yields (Jackson et al., 2001; Pauly et al., 2002; Garcia and Grainger, 2005). Increasing the productivity of capture fisheries for the long term will involve hard decisions about reducing fishing effort, removing excess fishing capacity, and building the enabling institutional arrangements to create property rights and/or other incentives to allow better management (Pauly et al., 2002; Hilborn et al., 2003; Garcia and Grainger, 2005). The hardships involve in such decisions, in terms of loss of livelihoods for fishers, may be offset eventually by the capacity of a restored and well-managed stocks to produce higher and more sustainable economic benefits for those remaining in the industry. Unfortunately, the reforms outlined above will be difficult to implement, and remain to be the key challenges for fisheries managers worldwide. The ‘restocking’ and ‘stock enhancement’, that could either reduce the time needed to rebuild certain capture fisheries to more productive levels, or increase the productivity of some ‘healthy’ fisheries

2.5 Ecological Studies

2.5.1 Benthic study

Benthic study is the study of benthos organism in the interested river. The benthos refers collectively to all aquatic organisms which live on, in or near the bottom (substratum) of water bodies. This includes organisms inhabiting both running and standing waters, and also applies to organisms from both saltwater and freshwater habitats. The benthos may subdivide based on size. Large benthic animals (those readily visible without the use of a microscope) are collectively referred to as macrozoobenthos or macroinvertebrates. Representatives include clams, snails, worms, amphipods, crayfish, and the larvae of many aquatic insects (e.g., dragonflies, mayflies, stoneflies, caddisflies, chironomid midges, and black flies.). Microscopes are essential to discern members of the microbenthos, e.g. nematodes, ostracods. Benthos organism's communities are well suited for use as biomonitoring tools, because the various benthic organisms have differing sensitivities to environmental stressors (Smith, 1971). By measuring the diversity of the benthic community, we can gain some insight into the level of human impacts on the aquatic system. Benthic invertebrates live longer than most planktonic organisms, and thus will indicate the effects of environmental conditions over time. They are relatively sedentary, therefore easy to sample and can serve as indicators of specific areas (Smith, 1971). In addition to serving as indicators of ecosystem condition, many benthic invertebrates are also important components of aquatic species diets and provide an important link in the food chain.

2.5.2 Water Quality

Water quality is the physical, chemical, and biological characteristics of water (Mitchell & William, 1994). It is most frequently used by reference to a set of standards against which compliance can be assessed. The most common standards used to assess water quality relate to drinking water, safety of human contact, and for health of ecosystems. For this project, several parameters will be assessed such as temperature, pH, dissolved oxygen, salinity, alkalinity, suspended solid, and ammonia concentration. Temperature and pH affects the solubility and, in turn, the toxicity of many other parameters. Generally, the solubility of solids increases with increasing temperature, while gases tend to be more soluble in cold water. Temperature is a factor in determining allowable limits for other parameters such as ammonia. Ammonia (NH_3^+) is a colorless gas with a strong pungent odor. It is easily liquefied and solidified, and is very soluble in water (Mitchell & William, 1996). It has been reported toxic to fresh water organisms at concentrations ranging from 0.53 to 22.8 mg/L (Smith, 2003). Toxic levels are both pH and temperature dependent. Toxicity increases as pH and temperature decrease. Plants are more tolerant of ammonia than animals, and invertebrates are more tolerant than fish (Smith, 2003). Alkalinity refers to the capability of water to neutralize acid. It essentially absorbs the excess H+ ions and protects the water body from fluctuations in pH. Alkalinity is important for fish and aquatic life because it protects or buffers against rapid pH changes (Mitchell & William, 1996). Living organisms, especially aquatic life, function best in a pH range of 6.0 to 9.0 (Smith, 2003). Alkalinity is a measure of how much acid can be added to a liquid without causing a large change in pH. Higher alkalinity levels in surface waters will buffer acid rain and other acid wastes and prevent pH changes that are harmful to aquatic life. Dissolved oxygen analysis measures the amount of gaseous oxygen (O_2) dissolved in an aqueous solution.

Oxygen gets into water by diffusion from the surrounding air, by aeration (rapid movement), and as a waste product of photosynthesis.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Introduction

This chapter will thoroughly explain how the studies were conducted. There are two different studies which are the ecological study and molecular study. Mainly for the ecological aspect, the study will focus on various data such as benthic study and water quality. The second part of this project would be molecular aspect study. The molecular technology in this case is the application of microsatellites and mitochondria as genetic markers to detect genetic variability between the organisms studied. Figure 3.1 below is a simplified version of flowchart showing the methodology of this study.

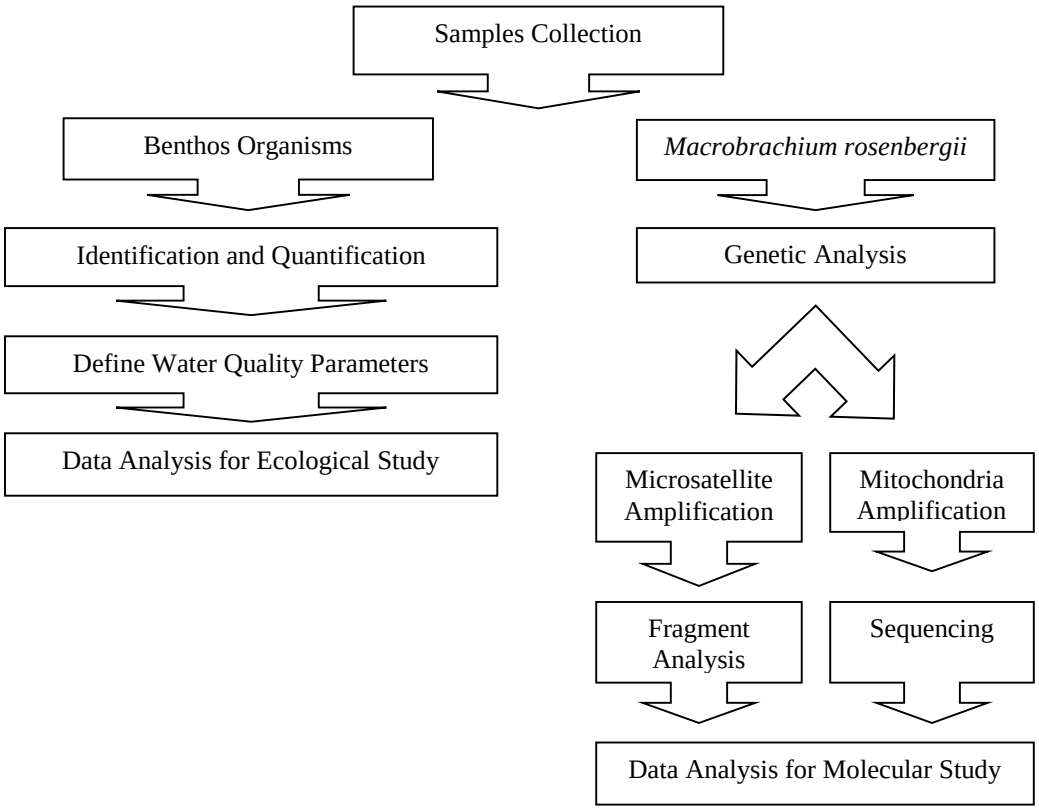


Figure 3.1: Methodology

3.2 Sampling Methods

Based on traditional practice by the fisherman, there are four ways that have been used in collecting prawn sample for this study, which are by fishing nets or specifically cast nets, fish traps, spear fishing and fishing rods. A cast net is a circular net with small weights distributed around its edge (Figure 3.2A). When the fishermen at Sungai Timun use this kind of technique, first they will use a very long pole to mark a specific location and then the bait is thrown into the water near the pole. The bait balls can be made of anything a prawn will eat. The most common bait that was used by the fisherman at Sungai Timun is the coconut flesh. After several minutes, the cast net is thrown as close to the bait as possible and the prawn are caught in the net. Fish trap is one of the most important methods that were used in this study. As fish traps usually can catch a lot of samples at one time and it's very easy to use. The fish traps or "kambang" in Malay words are made from small mesh green nylon strengthen with iron metal to make it square shape (Figure 3.2B). It has two funnels and usually the bait is tied up at the center of the traps. Then the traps were put along the river by tying it up with a stone to weigh it down the river. The traps were usually left overnight before being collected back. Spear fishing uses sharpened sticks has been a popular technique throughout the world for centuries. Unfortunately, this technique requires very calm shallow water and only can be done in the night with full moon light. At Sungai Timun, only the experienced fishermen can practice this technique as the fishermen must account for optical refraction at the water's surface, which makes the prawn appear to be further away. By experience, the fisher learns to aim lower to hit the target. Fishing rod, is a length of fishing line which is attached to a long, flexible rod or pole; one end terminates in a hook for catching the fish (Figure 3.2C). The hook is where usually the baits were lured. The fishermen of Sungai Timun usually use small prawns as bait. These small prawns usually can be caught near the bank of the river. Between the four methods,

the first two early methods that has been mentioned which are cast nets and fish traps catches small sizes of prawns while the other two methods catches bigger sizes of prawns.

Figure 3.2: Sampling Methods. 3.2A; The fisherman is using cast nets. 3.2B; The fisherman is using “kambang” or fish trap. 3.2C; The fisherman is using a fishing rod.
3.2D; Eckman Grab (Source: Private Picture)

3.3 Samples Collection

In this study, three kinds of sample were sampled that comprise of water from the river, benthos organisms and the prawns. All the samples were collected along the river, however five sampling stations were established on transects running along the river (Table 3.1). The stations were allocated by detecting the coordinate by using GPS system (Figure 3.4B). Water depths recorded ranged from 1.8 m to the 7.8 meter. Sampling was conducted for two years which was in 2009 and 2010 (Table 3.2).

Table 3.1: Point with the coordinate

Station	***Longitude***	***Latitude***
1	N02 26°249’	E102 03°783’
2	N02 25°955’	E102 03°062’
3	N02 25°253’	E102 03°034’
4	N02 25°454’	E102 02°191’
5	N02 25°972’	E102 01°768’

Benthos organisms were sampled with an Eckman Grab with an area of 202.5 cm^2 (Figure 3.2D). Each sampling station will be divided into three substations across the river. This was done so that benthos organisms were collected in three replicates. The samples were placed in a washing bucket containing 30 meshes per inch sieve screen. Hudson (1970) found that only organisms with a head diameter larger than 0.516 mm will be retained by the 30 mesh per inch sieve screen. Remaining organisms will be selected out by using a scalpel and transferred to 70% alcohol solution before being taken back to the laboratory for separation.

All freshwater prawns that needed for this study were obtained from two sites which are Kampung Acheh, Perak and Sungai Timun, Negeri Sembilan (Figure 3.3). There were 5 populations that had been used from both of these sites which were the Parental population, the F1, F2, F3 population and the Wild Type population. The Parental population was obtained from Kampung Acheh where the hatchery process operations for PL dissemination and restocking programs are used (http://perikanan.perak.gov.my/ppug/fpageppug.htm). While the Wild Type populations *M.rosenbergii* is the population that originally inhabited from Sungai Timun before the restocking program. The restocking program was done by releasing the juveniles originated from Kampung Acheh into the Sungai Timun periodically. Table 3.3 shows the dates of releasing the juveniles of the prawn to the Sungai Timun. 35 samples of Wild Type population which was consisted of 20 male and 15 female were collected in 2004 and 202 samples of Parental population which was consisted of 192 female and 10 male were collected earlier before the sampling for F1 population begun. After the releasing of juveniles to the rivers, an assumption is made that the prawn juveniles would survive in the rivers and once again samples were collected, these entire sample would include as the F1 population.. In this study, F1 samples is the prawn captured on 2008, F2 samples is the prawn captured in 2009 while F3 is the prawns captured in 2010. They were stored in ice at 4°C before being brought back to Genomics and Molecular Breeding Lab, University of Malaya where they were kept at -80°C until DNA extraction was done.

Table 3.2: Sampling months for F1 population

Month	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
2008		•			•		•			•	•	•
2009		•	•	•	•	•	•	•	•	•	•	•
2010	•	•	•					•	•	•		

Table 3.3: Stocking dates

Date	*Number of Stocks*
04 April 2007	40,000
04 April 2007	40,000
15 April 2007	22,000
22 June 2007	55,800
20 July 2007	27,000
10 June 2008	80,000
10 Oct 2008	80,000
14 May 2009	40,000

Figure 3.3: Location of prawn sampling;
A) Pusat Penetasan Udang Galah, Kg.Acheh, Perak B) Sungai Timun, Negeri Sembilan

3.4 Ecological Study

All benthos organism that were collected and separated into taxonomic groups and the numbers in each group were recorded with the use of microscope (100X enlargement) using Taxonomic keys. Representative specimens were selected and preserved in 70% alcohols for later identifications and the remainder were preserved in 10% formalin for weighing.

Other than benthos organism, the ecological study required collection of water quality parameters. In each station, water quality parameters were collected using Multi Parameter Water Quality Meter (Figure 3.4A). The sensor of this analyzer or called probe was submerged into the river for about 5 minutes to get the mean reading for each parameter. This analytical instrument, capable of measuring up to 13 critical water quality parameters (6 measured and 7 calculated), including pH, EC/TDS, ORP, DO, atmospheric pressure, and temperature. However for this study, only five parameters taken which were pH, salinity, ammonia, nitrate and phosphate.

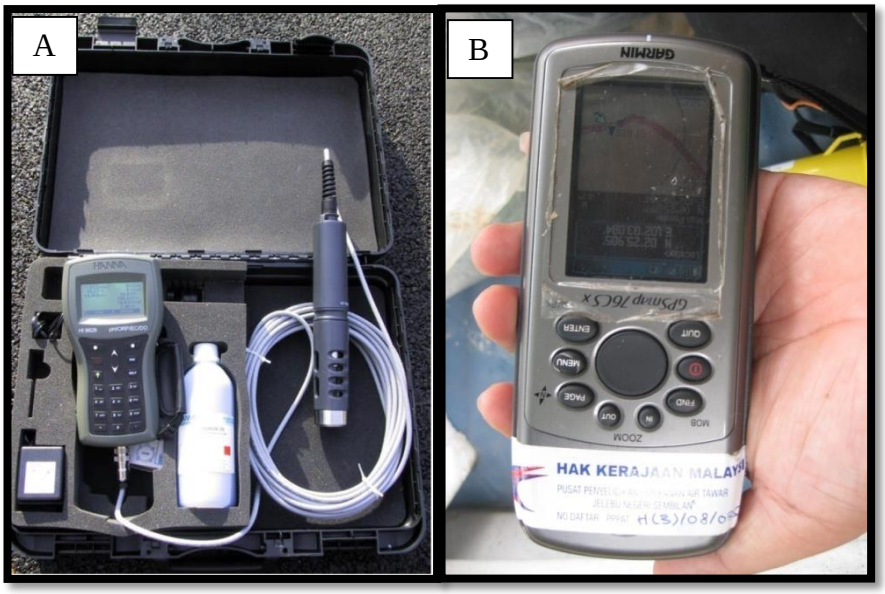


Figure 3.4: (A) Multi Parameter Water Quality Meter. (B) GPS System

3.5 Molecular Study

The molecular study which has been explained in chapter 2 is the application of the molecular markers in detecting the genetic variation between the samples from all the population collected. Two types of markers were used in this project which is microsatellite and mitochondria marker. This section of this chapter will thoroughly explain the processes of the molecular study that comprise of DNA analysis, developing all the primers needed and lastly the data analysis. The first step were the DNA analysis, where all the prawn samples needed were extracted. While in the second step, how all the primers including genomics SSR (Type II) and EST SSR (Type I) were developed and used through the study were described and the use of a maternal genetic marker which is the mitochondria marker.

3.5.1 DNA Analysis

3.5.1.1 DNA Extraction

All prawn sample from each population were extracted using GF1 DNA EXTRACTION KIT (Vivantis). Firstly, 20-30 milligrams of tissue sample was cut by using the scapula and using a balance machine to weight the sample to get the appropriate size. Then the sample was smashed (for more efficient lysis) and filled into 1.5 ml microcentrifuge tube. A 250 µl of buffer TL (tissue lysis) and 20 µl Proteinase K was added into the tube which then the solutions were mixed thoroughly by pulsed vortexing to obtain homogenous solution. Next, 12 µl of lysis enhancer was added and mix immediately. Then, the solution was incubated at 65ºC for 1-3 hours in a shaking water bath or mix occasionally during incubation to ensure thorough digestion of the sample. For better result, the tissue can be incubated overnight (not more than 18 hours) at 37ºC. Next, 20 µl of RNAse A was added into the solution before the

solutions will be incubated at 37ºC for 10 minutes. Homogenization would be the next step, before about ~600 µl of buffer TB was added and mix thoroughly by pulsed-vortexing until a homogenous solution were obtained before the solution was incubated for 10 minutes at 65ºC. Another 200 µl of absolute ethanol was added and once again mix thoroughly and immediately by pulsed-vortexing until a homogenous solution were obtained to prevent any uneven precipitation of nucleic acid due to high local ethanol concentrations. The resulting mixture was transferred into a column assembled in a clean collection tube (provided in DNA extraction kit) by pipetting out approximately 600 µl with the micropipette. Then the tubes was centrifuged at a speed of 5000 ×g for 1 minute before flow through was discarded. This step was repeated until all the solution was transferred to the column. Subsequently, 750 µl of Wash buffer was added and later centrifuged at 5000 ×g for 1 minute before the flow through was discarded. This step was repeated twice. The second last step was the column drying. It has been done by centrifuged the column at 10000 ×g for 1 minute to remove all traces of ethanol. Afterwards, the column was transferred to the new 1.5 ml microcentrifuge tubes and 80 µl of preheated Elution buffer was added directly to the column membrane and the column stand at room temperature for 30 minutes (> 1 hour for more concentrated DNA). Lastly, the column was centrifuged at 5000 ×g for 1 minute to elute DNA. The DNA was stored at 4ºC.

3.5.1.2 DNA Quality

The quality of the DNA after the extraction process was tested visually by running the DNA samples on a 1.0% agarose gel in 1X TBE buffer and using Lamda Hind ΙΙΙ ladder to determine the approximate length (in bp). The gels were ran at 78V for about 45 minutes using Power Pack (BioRad®). The gel was then stained with Ethidium Bromide for 15 minutes exactly before being visualized under UV light in an Alpha Imager™ 2200 Transilluminator.

3.5.2 Genomics SSR

3.5.2.1 5' Anchored PCR Amplification

Two degenerate primers were used to amplify the *Macrobrachium rosenbergii* genome: VJ2 (5'- NNN KKV RVR V$(CTC)_5$ -3') (Kumar *et al.,* 2003) and T79112 (5'- (K)2(YH)3Y$(GTT)_5$ -3') (See Leng Min,2008) and, where K = G/T, V = G/C/A, R = G/A, N = A/C/G/T, H = A/C/T and Y = T/C (IUB code). The $(CTC)_5$ and $(GTT)_5$ component of the VJ2 and T79112 primer respectively was designed to anneal to (CTC)n and (GTT)n, where $n \geq 5$, and the 10 nucleotides in both of the primers form the 'anchor'.

PCR was carried out in a total volume of 10 µl containing ~20 ng genomic DNA, 1X PCR buffer , 0.25 mM each of dATP, dGTP, dCTP, and dTTP (Promega), 0.5 µM of each primer, 3 to 5 mM MgCl2 (Promega), 1.5 U Taq polymerase (Promega) and deionized water. Amplification was performed in a MultiGene Thermal Cycler (Applied Biosystems) with an initial 3 minutes predenaturation at 96°C, followed by 40 cycles of denaturation at 94°C for 10 seconds, an appropriate annealing temperature for 10 seconds and an extension at 72°C for 30 seconds. A final extension step at 72°C for 7 minutes was included.

Table 3.4: List of 5' anchored primers used for obtaining PCR products for cloning.

No.	***Primer***	***Sequence (5'-3')***	***Annealing temperature (°C)***
1.	VJ2	NNN KKV RVR V(CTC)5	50
2.	T79112	(K)2(YH)3Y(GTT)5	47.5

3.5.2.2 Cloning

Day One

On the first day, several preparations were made before the cloning process started. One of it was LB Agar as medium. LB Agar was prepared by adding 35 gram of LB Agar powder to 1 liter of double distilled water. The solution then was autoclaved at 121°C for 20 minutes to ensure all the powder dissolved in the solution. 60 ml of the medium were poured to each sterilized plate. However before pouring the medium 250 µl ampicillin was needed to add to the medium. Ampicillin was used to detect the recombinant plasmid. Ampicillin solution were prepared by adding 0.05 gram Ampicillin powder to 1000 µl deionized water to make it 50mg/ml. The plates were then sealed and put in the 37°C oven. Other than this, X-gal solution was also prepared by adding 0.05 X-gal powder to 1000 µl before covering with aluminum foil and kept in -20°C.

Day Two

On the second day, all the PCR product were checked on 1% agarose gel to make sure there are amplification of the product by the primers. The PCR was done by using colorless buffer instead of green (Promega). In the evening, yT&A cloning vectors (Yeastern Biotech Co., Ltd.) were centrifuge for 30 seconds while ligation buffer A & B were vortexed. Next all the components were mix together according to the protocol under the laminar flow (Table 3.5). Ligation mixture was vortexed and was incubated overnight at 4°C.

Table 3.5: Ligation Mixture Protocol

Mixture	***Volume (µl)***
Ligation buffer A	1
Ligation buffer B	1
yT&A cloning vector	2
PCR product	5
yT_4	1

Day Three

On the third day, 20 µl of X-gal was added by separating it into each plate by using L-Rod. The plates then were incubated 37°C for at least an hour. While waiting for the plate in the incubator, competent cell were prepared by using ECOS^{TM}(Yeastern Biotech Co., Ltd.) according to manufacturer instructions with minor modification. ECOS^{TM}were thawed for 25 seconds under the tap water before 10µl of ligation mixture were transferred into the ECOS^{TM}tubes. The mixture was flicked a few times to make sure the solution was mix. After that, it was incubated immediately in 42°C water bath for 35 seconds. Then 100 µl of the mixture were spread evenly onto the main plate using glass beads. The plates were sealed and incubate overnight at 37°C.

Day Four

The sub cloning process was done on the fourth day of cloning. The blue white colonies in each plate were observed. All the white colonies were transferred to the new plate using wooden sticks. The plates were then incubated at 37°C for 2 days.

Day Six

On the sixth day of cloning, LB broth were prepared by mixing 2 gram of LB broth with 100 ml of deionized water and autoclaved it. 5 ml of bearable temperature LB broth and 5µl ampicillin were added into each falcon tube. Each colony was transferred into each tube and all the tubes were put in shaker at 37°C for 12 to 16 hours.

Day Seven

The growth of the competent cells were observed in each tube. Twelve recombinant clones from each primer were randomly selected for plasmid extraction. Plasmid extraction was done by using HiYield Plasmid Mini Kit instruction protocol (Yeastern Biotech Co., Ltd.). 1.5 ml of bacterial culture was centrifuge for 1 minute at 14000 rpm in the centrifuge tube. The supernatant was discarded. Then 200µl of PD1 buffer was added to the tube before vortexing the mixture to resuspend the cell pellet. Next 200 µl of PD2 buffer was added and mixed by inverting the tube for minimum 10 times. The lysate was left in room temperature for 2 minutes until cleared. The next step was the neutralization when 300 µl of PD3 was added and mixed immediately by inverting the tube for 10 times before being centrifuged at 14000 rpm. Meanwhile PD Columns were placed in a collection tubes that been provided. The clear lysate was applied to the PD Column before centrifuged at 8000 rpm for 30 seconds. Then the flow through was discarded and the PD Columns were place back in the collection tubes. In the wash step, 400 µl of W1 was added in the PD column and centrifuged at 8000 rpm for 30 seconds. The flow through was discarded and the PD Columns were put back in the collection tubes. Similarly, this step was repeated but using 600 µl Wash Buffer. Next, the PD Columns were centrifuged at 14000 rpm for 2 minutes to dry the column matrix. The dried PD Columns were

transferred on a clean microcentrifuge tube before 50 µl of Elution Buffer was added into the centre of the column matrix. After an hour left in room temperature, the columns were centrifuge for 2 minutes at 14000 rpm. Finally, the flow through which is the purified DNA was then checked on 1% agarose gel.

3.5.2.3 DNA Sequence

The best presented bands on the 1% agarose gel of the purified DNA were selected. Consequently, all the purified DNA samples selected were submitted to the 1st BASE for sequencing purpose. The DNA sequences were then submitted to the National Center for Biotechnology Information (NCBI) GenBank database by using Sequin Application Version 11.75 in FASTA format. Sequin is a stand-alone software tool developed by the NCBI for submitting and updating sequences to the GenBank, EMBL, and DDBJ databases.

3.5.2.4 Designing Primers Flanking Microsatellite Regions

The primers flanking the microsatellite repeat regions were designed using the PRIMER3 program (Rozen and Skaletsky, 1997) with product lengths ranging from 150 to 300 basepairs as the controlled parameter. The primers sequences were sent to the 1st BASE for synthesizing.

3.5.2.5 Microsatellites Amplification

Optimization of PCR parameters

Some of the chemical volume depends on the expected size of microsatellites such as the volume of Magnesium Chloride ($MgCl^2$), GoTaq® Green Flexi Buffer 5×, GoTaq® flexi DNA Polymerase and the volume of primer (forward and reverse). This is because, each primer have its own requirement so that it can function properly. The optimization process was done to get the optimum results from the different reaction parameters, such as

- Different Magnesium Chloride ($MgCl^2$) concentrations
- Different annealing temperatures
- Different quantities of Taq polymerase

The other parameters such as DNA templates, types of Taq Polymerase (GoTaq® flexi DNA Polymerase), types of thermal cycler (Multigene®) and types of Metaphore gel (Agarose®) were kept constant.

Polymorphisms Testing

All pairs of primers where then tested for their polymorphisms with 30 wild *Macrobrachium rosenbergii* individuals from Timun River, Negeri Sembilan. The basic method of PCR is divided by two which is the first part when all the content were mixed to produce the reaction volume before running in the PCR machine. The PCR was performed in Multigene® Thermocycler. The reaction volume consisting of :

a) 2.5 mM Magnesium Chloride ($MgCl^2$) (Promega®)

b) 1 × GoTaq® Green Flexi Buffer (Promega®)

c) 500.0 µM of each dNTPs (Promega®)

d) 1.5 units GoTaq® flexi DNA Polymerase (Promega®)

e) 50.0 pmole primer (1^{st} Base)

f) 20.0 µg of DNA sample

g) Deionized distilled water to make up the reaction volume to 10µl

The second part of the PCR is the PCR protocol. The primer of microsatellites for, all the temperatures are the same except for the annealing temperatures which depends on the primer. This temperature is the most important as if the temperature is not right, the primer cannot anneal to the DNA strands, the DNA Polymerase cannot function properly and the replication process would not take place. The temperature for the annealing process is known by optimization through PCR. Below is the protocol (Table 3.6) that will be followed in this study.

Table 3.6: Protocols for PCR amplification

Process	***Temperature (°C)***	***Time***
Predenaturation	94.0	3 minutes
Denaturation	94.0	30 seconds
Annealing	Depends on the Primer	30 seconds
Extension	72.0	30 seconds
Final Extension	72.0	7 minutes
Storage	4.0	∞

Gel Electrophoresis

The electrophoresis of the PCR products, 4 % Metaphor Gel (Agarose®) was prepared and let it dried in room temperatures for about 2 hours so that it becomes solid. Once dried, it was then placed in a tank of buffer, TBE 1×. Dyed PCR samples (as colourless buffer was used) were pipetted into the well of the gel, 5µl each and lastly the 20bp ladder (Vivantis) and their designated dyes was mix together with a micropipette before being the last to be pipette into a well. 78 V and 150 ampere Electric source (BioRad®) was attached to the tank and electrophoresis was carried out until the Bromophenol blue dye reached the final level accordingly which took around 2 hours and 30 minutes.

Staining

When the gel electrophoresis was done, the gel was removed from the plate and immersed in liquid Ethidium Bromide for 15 minutes to be stained. Lastly, it was observed under UV light in an Alpha Imager^{TM} 2200 Transilluminator. Photographs was taken of the bands that were detected.

3.5.2.6 Labelling primer

The polymorphic primer pairs identified were sent to 1st BASE for labelling. The forward sequence of all the polymorphic primers were labelled with 5' FAM fluorescent dye so that it can be determined with GeneScan^{TM} 500 LIZ size standard and ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) to confirm the polymorphism.

3.5.2.7 Fragment Analysis

The entire PCR product will be wrapped in aluminum foil to prevent the product from exposing to light because the florescent primer is light sensitive and will affect the Fragment analysis. In fragment Analysis, 2.0 µl of PCR product was added with 10.0µl formamide or "Hidi". Then 0.2µl of internal standard size (GS-250) that act like a ladder was mixed to the solution. The solution will be heated at 95 ºC for 3-5 minutes by using a heat block or PCR machine. After the heating process, the solution will be placed immediately on ice for at least 5 minutes before loading the solutions into MicroAmp® Optical 96-Well Reaction Plate. Lastly, the plate was placed into the Genetic Analyzer machine to produce graph scale result.

3.5.3 EST SSR

3.5.3.1 Development of EST SSR

In this study, we are detecting microsatellites in exonic region of *Macrobrachium rosenbergii*, by using contigs developed from short read sequences acquired by RNA sequence. Next-Generation Solexa Single-End Sequencing was carried out on total hepatopancreas RNA into RNA short reads. The short read sequences of hepatopancreas were assembled into contigs using an assembler programme, SOAPdenovo. Assemblies were performed with k-mers of 15, 17, 19, 21, 23, and 25 using default parameters. The k-mer that produces most contigs with the highest N50 values when chosen. The contigs were compared against genomic data of *Macrobrachium rosenbergii*, acquired via sequencing by RAMs (random amplified microsatellites). This is to confirm that the data acquired. Microsatellite repeat motifs/SSRs were identified from contigs. SSRs lying too close to 5' and 3' ends were excluded from study.

The original contigs that contain the microsatellite repeat motifs were BLASTed used BLASTx. BLASTx was used for gene annotation purpose. The database used was non-redundant protein database, and the criteria set was that E-value $<10^{-5}$ with low complexity filters enabled. The genes annotated was separated into 3 groups, unknown gene product, known gene product, hypothetical proteins. Classification of known genes was based on Gene Ontology, using BLAST2GO. WEGO was used for data illustration purpose. This is done to find out what proteins did the contigs with SSR code for.

3.5.3.2 Designing Primers

Primer was designed using PRIMER3 program (Rozen and Skaletsky, 1997) with the criteria's set for the EST microsatellite are as follows:-

a) 6 repeats or above for dinucleotide microsatellites.
b) 5 repeats or above for trinucleotide microsatellites.
c) 4 repeats or above for tetranucleotide microsatellites.
d) 3 repeats or above for pentanucleotide and above.
e) Position of the microsatellite must be at least 17-20 nucleotide positions from both ends of the EST sequence.

All the primers sequences were sent to the 1st BASE for synthesizing.

3.5.3.3 Validation of EST SSR

These primers were then used to validate at genomics level in 10 wild *Macrobrachium rosenbergii* individuals from Timun River, Negeri Sembilan. This process was done to make sure that the EST SSR primers were able to amplify microsatellite region in the *Macrobrachium rosenbergii.* The PCR amplifications were performed in a 10-µl reaction mixture containing 20 ng genomic DNA, 1X PCR buffer (Promega), 0.25 mM of each dNTP (Promega), 0.5 µM of each forward and reverse primer, 2.5-5.0 mM MgCl2 (Promega), 1.5 U Taq polymerase (Promega) and deionized water. PCR profiles involved: initial denaturation at 95°C for 3 minutes, followed by 35 cycles at 94°C for 30 seconds, annealing temperature were between 45°C to 60°C for 30 seconds, and extension at 72°C for 30 seconds with a final

extension at72°C for 7 minutes in a MultiGene Thermal Cycler (Applied Biosystems). PCR products were run on a 1% agarose gel at 60 V for an hour.

3.5.3.4 Polymorphism Testing

The working primer pairs identified were then tested for polymorphism. (Refer to the Polymorphism Testing, Gel Electrophoresis and Staining subtopic in 3.5.2.5 for this procedure).

3.5.3.5 Labelling primer

Refer to 3.5.2.6 subtopic for this procedure.

3.5.3.6 Fragment Analysis

Refer to 3.5.2.7 subtopic for this procedure

3.5.4 Genetic Diversity Study of Macrobrachium rosenbergii

3.5.4.1 Marker Selection

In this study, there were two categories of microsatellites markers that were used to detect genetic diversity for all the samples. The Genomics SSR is the Type II marker while EST SSR is the Type I marker. In this chapter, several methodologies have been discussed in developing both of these markers. However, several markers that was published in the paper were also used in this study. Thus, to make sure that all the primers that has been used have high discriminating and polymorphic power in differentiating all the samples, PIC value has been applied. Polymorphic information content (PIC) refers to the value of a marker for detecting polymorphism in a population (PIC, Botstein et al., 1980). PIC depends on the number of detectable alleles and the distribution of their frequencies, which equals 1 minus the sum of the square of all allele frequencies (Z.J. Liua, J.F. Cordes, 2004). For example, the PIC of a microsatellite marker with four alleles of frequency 0.25 each should be **$1- [(0.25)^4 + (0.25)^4 + (0.25)^4 + (0.25)^4] = 0.984375$**. Comparison of PIC values between each primer can be served as parameters to differentiate the power of each primer.

3.5.4.2 Microsatellite Amplification

Refer to 3.5.2.5 subtopic for this procedure. This procedure was applied to samples from all populations which Wild Type, Parental and F1 populations.

3.5.4.3 Fragment Analysis

Refer to 3.5.2.7 subtopic for this procedure.

3.5.5 Mitochondrial DNA Marker

"Universal" DNA primers for polymerase chain reaction (PCR) amplification of a 710-bp fragment of the mitochondrial cytochrome c oxidase subunit I gene *(COI)* were used in this study because it appears to be among the most conservative protein-coding genes in the mitochondrial genome of animals (Brown, 1985), which was preferable to keep track of the lineages of the river population. Below are the primer sequences that were used to amplify the COI region in all of the samples (O. Folmer *et al.*, 1994).

MITO MRF

5'-GGTCAACAAATCATAAAGATATTGG-3'

MITO MRR

5'-TAAACTTCAGGGTGACCAAAAAATCA-3

171 samples has been amplified with the mitochondrial marker and it comprises 3 months of F1 populations, parental population and also wild population (Table 3.7). The reasons for only three months of F1 sample has been amplified because the duration of the grow-out period of juvenile prawns generally took around 150 to 180 days in the river before the prawn is sexually matured to breed (L.R.D'Abramo *et al*.,2003). As the stocking programme was started on April 2007, it will take around 6 months before the parental populations can breed with the wild population in the river and produce F1 generation. The prawns will take another 6 months to grow up as only adult prawn were sampled in this study.

Table 3.7: Population type with total sequences

Samples	***Total Sequences***
F1 (July 2008)	19
F2 (July 2009)	42
F3 (August 2010)	47
Parental	43
Wild Type	30

3.5.5.1 Mitochondrial Amplification

A series of experiments has been done in order to obtain optimal condition for the primer pairs. The experiment include manipulation of parameters such as annealing temperature, concentration of DNA, concentration of *Taq* polymerase, and concentration of Magnesium chloride ($MgCl^2$). The condition for optimization of the primers contained a total volume of 20 µl reaction mixture. The PCR was performed in Multigene® Thermocycler. The reaction volume consisting of:

a) 2.4 µl of Magnesium Chloride ($MgCl^2$)(2.5 mM) (Promega®)
b) 6µl of 1 × GoTaq® Green Flexi Colourless Buffer (Promega®)
c) 0.5µl of each dATP, dGTP, dCTP and dTTP (10mM) (Promega®)
d) 0.6µl Promega GoTaq® Polymerase (5u/µl) (Promega®)
e) 0.5µl of each primer (10µM) (1st Base)
f) 4.0µl of DNA sample
g) Deionized distilled water to make up the reaction volume to 20µl

Below is the PCR protocol that will be followed in this study for 40 cycles.

Table 3.8: Protocols for PCR amplification

Process	***Temperature (ºC)***	***Time***
Predenaturation	96.0	3 minutes
Denaturation	96.0	30 seconds
Annealing	45.0 – 55.0	30 seconds
Extension	72.0	30 seconds
Final Extension	72.0	7 minutes
Storage	4.0	∞

3.5.5.2 Gel Electrophoresis

Electrophoresis is then carried out on 1 % Agarose gel in order to confirm the presence of the PCR products. Dyed PCR samples (as colourless buffer were used) were pipetted into the well of the gel, 5µl each and lastly the 100bp ladder (Vivantis) and their designated dyes was mixed with a micropipette before being last to be pipette into a well. 60 V and 150 ampere electric source (BioRad®) was attached to the tank and electrophoresis was carried out until the yellow colour dye reached the final level accordingly.

3.5.5.3 Staining

Once electrophoresis was completed, the gel was removed from the plate and immersed in liquid Ethidium Bromide for 15 minutes to be stained. Finally, it was observed under UV light in an Alpha Imager^{TM} 2200 Transilluminator. Photographs are taken of the bands that

were detected. Then the gel containing the expected DNA band will excise and will be purified to confirm the occurrence of the DNA.

3.5.5.4 Gel Purification

Gel purification was carried out using NucleoSpin® Extract II according to the protocol provided by the company (MACHEREY-NAGEL).The gels that contained the selected bands with the correct bands size were excised from the agarose gel with a clean scalpel before transferred into 1.5 ml microcentrifuge tube. The process was done in the dark room under UV illuminator. Next, the sliced gel was then crushed with tips before 200µl of Buffer NT was added to the gel. Now, the crushed gel was incubated for 10 minutes in 50°C until all the gels were completely dissolved. After incubation, a NucleoSpin® column was placed into a new 2ml collection tube provided and the samples were loaded into it. Furthermore, the solubilized gel was centrifuged at 11,000×g for 1 minute and the flow through was discarded. 600µl of Buffer NT3 was added into the tube and centrifuged at 11,000×g for 1 minute before the flow through was discarded. This step was repeated with another 200µl of Buffer NT3. Moreover, the columns were centrifuged at 11,000×g for 2 minutes for dying process. The column wer transferred into a new 1.5 microcentrifuge tube before preheating Elution Buffer (preheated at 70°C) was then added to the centre of the column. The columns were then incubated overnight in the room temperature. Finally the columns were centrifuged for a minute at 11,000×g. The elution contained the purified DNA fragment was stored at -20 °C for sequencing application.

3.5.5.5 DNA Sequencing

The purified DNA was checked first on 1% agarose gel to ensure the success of gel purification. The succeeded product of purified DNA was subjected to sequencing reactions in both directions using the BigDye® Terminator v3.1 Cycle Sequencing Kit following the manufacturer's instructions. There were two steps in DNA sequencing; first step was PCR the purified product and the second step was ethanol/EDTA precipitation. The first step, PCR mixture was prepared by using the purified DNA. The 20µl PCR mixture consisting of:

Table 3.9: PCR Mixture for DNA Sequencing

Reagent	***Quantity***
BigDye®	1µl
Buffer	3µl
Template (Purified DNA)	5µl
Primer (3.2 pmol)	1µl
Deionized water	10µl

Below is the protocol that will be followed in this study.

Table 3.10: Protocol for DNA Sequencing

Process	***Temperature (ºC)***	***Time (minutes)***
Predenaturation	96.0	1 minute
Denaturation	96.0	10 seconds
Annealing	50.0	5 seconds
Extension	60.0	4 minutes
Storage	4.0	∞

There are two methods in carrying out ethanol/EDTA precipitation which is by using tube or 96-well reaction plate. However, the plate method is simpler than the tube method. First, the PCR products were transfer to MicroAmp® Optical 96-Well Reaction Plate. Then the plate was briefly spun before 5µl of 125mM EDTA and 60µl of 100% ethanol were added. The plate was then sealed with Masterclear real-time PCR Film (Eppendorf). The sealed plate was incubated at room temperature for 15 minutes. It was centrifuge at 2000×g for 45 minutes and all the mixture was discarded by inverting the plate on the tissue. Another 60µl of 70% of ice cold ethanol were added and it was centrifuged at 1650 for 15 minutes at 4°C. Similarly to the final step, all the ethanol was discarded by inverting the plate on the tissue. As a result, the pellet formed was dried at 90°C for 1 minute followed by 50°C for 5 minutes in the Eppendorf Thermocycler. Finally, 10µl of Hi-Di formamide was added before it was denatured at 95°C for 5 minutes in the Thermocycler. After denaturing, the plate was placed on ice for at least 5 minutes and then assembled into ABI plate sandwich. The sandwich is assembled in this order, the black baseplate, the clear reaction plate, the grey rubber septa, and the white plate retainer. The ABI plate sandwich was placed on the autosampler carefully. The plate is now ready to be processed by the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems).

3.6 Ecological Data Analysis

In ecological studies, the analysis of the benthic organisms the identification was made at the level of class and order but in some cases, the organisms were identified up to genus or species, using several Taxonomic Key that been published (Hartman, O.,1941,1944,1947,1954) (Robert L. et al., 2003). The picture of each species was taken and documented to make as future reference. The data were then transfer to Microsoft Excel 2003 including water sample parameters result obtained from Multi Parameter Water Quality Meter.

3.7 Molecular Data Analysis

3.7.1 Microsatellite data analysis

In this data analysis step, the microsatellite result from the graph of fragment analyzer were visually analyzed and scored directly by using GeneMapper® Software. The genotypes were determined by recording the single high peak for Homozygous and the double high peaks for the Heterozygous between the expected size neglecting other low peak which correspond to mismatches or unstable primer-template matches (Barral *et al*., 1993) or contaminates that can also cause excess background noise. The scored data were kept in Microsoft Excel 2003. The data were analyzed with four different kind of software which is GENEPOP, POPGENE, MICROCHECKER, and NTSYS to get different kind of results.

Table 3.11: Software those were used to analysis data

Software	***Result Produced***
GENEPOP version 4.0	Hardy Weinberg Equilibrium
	Allele Frequencies
	Heterozygosity
	Genetic differentiation (F_{ST})
	Linkage Disequilibrium
POPGENE version 1.31	Inbreeding Coefficient (F_{IS})
	Genetic Distance
MICROCHECKER version 2.2.3	Null alleles
NTSYS version 1.6	Dendograms

3.7.1.1 Allele Frequencies

Allele frequency or gene frequency is the relative frequency of a particular allele at a genetic locus in a population. It is usually expressed as a proportion or a percentage.The formula is:

$$f(Ai) = [\, n(Ai\ Ai) + n\,(Ai\ Aj)] / N$$

Where:

f (Ai) = frequency of allele Ai

n (Ai Ai) = number of individuals homozygous for allele

n (Ai Aj) = number of individuals heterozygous for Ai and Aj , $i \neq j$

N = total number of individuals in sample

3.7.1.2 Heterorozygosity

Heterozygosity is a widely used measurement for quantifying genetic variation of polymorphic loci. The heterozygosity value ranges from zero to 1. The observed heterozygosity (H_o) for a locus means the proportion of individuals in a population which are heterozygous at the locus. It is estimated using the formula:

$$\mathbf{Ho = n\ (AiAj)/\ N}$$

Where:

n (AiAj) = number of individuals with genotype AiAj , $i \neq j$;

N = total number of individuals in sample.

AiAj = alleles at the locus.

Some forces such as inbreeding, population persistence, and genotypic fitness may cause low observed heterozygosity.

The expected heterozygosity (H_e) is defined as the estimated fraction of all individuals who would be heterozygous based on the known allele frequency. It is determined for a locus as follows (Nei, 1978).

$$\mathbf{He = N\ (1- \Sigma\ Pi\ ^2)/N-1}$$

Where Pi is the frequency of the *i*th allele and N is the sample size. Deviation observed from the expected can be used as an indicator of important population dynamics.

3.7.1.3 Hardy-Weinberg Equilibrium

Hardy and Weinberg developed a concept based on Mendel's principles of inheritance, which is known today as the Hardy-Weinberg Principle. It essentially says that "in a large, randomly breeding (diploid) population, allele frequency will not change from one generation to the next, assuming no mutation, gene migration, genetic drift or selection and genotypic frequencies are related to gene frequencies by the formula:

$$\mathbf{P = p^2\ H = 2pq\ Q = q^2}$$

When a population meets all of the Hardy-Weinberg conditions it is said to be in Hardy-Weinberg equilibrium (HWE). Conformity to HWE can be determined by a statistical method called χ^2 (chi-squared) test, which is used to compare the observed (O) versus the expected (E) numbers.

$$\mathbf{\chi^2 = \sum (O- E)^2 / E}$$

The degree of freedom for the test is estimated as the number of expected genotypes minus the number of observed alleles. If by the chi-square test result it was concluded that the population was not in Hardy-Weinberg equilibrium, it was indicated that the population may be affected by one of the following forces: migration, mutation, genetic drift, selection and others which resulted in changes to the gene and genotypic frequencies.

3.7.1.4 Inbreeding Coefficient (F_{is})

F_{is} values help to detect departures from Hardy-Weinberg by measuring the amount of heterozygote deficiency or excess observed in the sample.

$$\mathbf{F_{IS} = 1 - HI / HS}$$

Where HS is an average expected heterozygosity within populations and HI is an average observed heterozygosity within populations. (Yýlmaz & Brahim, 2002)

3.7.1.5 Genetic differentiation (F_{st})

F_{st} values help to understand the degree of population differentiation within species.

$$\mathbf{F_{ST} = 1 - HS / HT}$$

Where HS is an average expected heterozygosity within Populations and HT is an average expected heterozygosity in total population. (Yýlmaz & Brahim, 2002)

3.7.1.6 Genetic distance

The genetic relationship between any two populations is a function of differences between them in allele's frequencies and this relationship is expressed in terms of a genetic distance. It is also a measure of the dissimilarity of the genetic material between different species or populations or individuals of the same species population. On the other hand, it can be used to estimate the time that has passed since the populations existed as a single cohesive unit.

The estimated genetic distances may vary among loci, therefore the most accurate measurement would be obtained by averaging many loci (Nei, 1978). Several distance measurements are available in the literature, but Nei's standard genetic distance, D, (Nei, 1978) had been widely used in studies of many populations, species and genera. In this study, the Nei's standard genetic distance was used by the following formula:

$$\mathbf{D_N = \ln \mathit{I}}$$

Where $I = J_{XY} / (J_X J_Y)^{0.5}$,and J_X, J_Y, and J_{XY} are arithmetic means across all loci.

3.7.1.7 Dendogram

Dendrogram is a visual representations of the genetic relationship based on the genetic distance by following the equation (Lynch, 1990)

$$\mathbf{D2ij = -ln [S'ij/ (Si\ Sj)]}$$

A pairwise matrix of distance was generated by using the GENEPOP software. Dendogram were constructed based on the unweighted pair group method with arithmetic averaging (UPGMA) employing the NTSYS software version 1.60.

3.7.2 Mitochondrial data analysis

In this data analysis step, all the mitochondrial result which is the sequences was visually analyzed by using GeneMapper® Software. Prior to data analysis, several examination need to be done to exclude the probability of Nuclear Sequence of Mitochondrial Origin (NuMts) as the primer used to amplify COI region in this study were universal primer for invertebrate species (Lopez *et al.*, 1994). A careful examination of sequencing gels and chromatogram is one of the early methods to identify NuMts other than produced high quality of DNA from DNA extraction. Two bands on the agarose gel or two peaks of complimentary strands at corresponding positions should be resisted (Sorenson *et al.*, 1998). Therefore, all sequence reads were checked against chromatograph data using Chromas Software (Technelysium Pty Ltd) to ensure the high quality of the sequence and to remove ambiguous bases. These successful amplified COI sequences were then blast in NCBI BLAST program to confirm the amplified products were came from the COI full sequence. After successfully blasting all the partial sequence, they were aligned using ClustaW (Thompson et al, 1997) with default parameters and trimmed at both ends to exclude any unsure sequences using MEGA 4 software version 4.1 (Tamura *et al*., 2007) . The sequence has been checked with any signs of unusual amino acid substitutions, stop codons and length mutations in protein-coding regions (Sorenson *et al.*, 1998). After cleaning all the sequences, the software DnaSP v5 (Librado & Rozas, 2009) was used to summarise haplotype information. In this case of understanding the evolutionary of *M.rosenbergii*, a haplotype network which connects all the haplotypes by mutational step was performed using software TCS version 1.21 (Clement *et al*., 2000). The network was transform into figure by using Microsoft PowerPoint. In order to measure the population distance, the F_{st} value which determines the population differentiation was then calculated using software Arlequin version 3.11 (Excoffier *et al*., 2005). This population

comparison was performed by computing pairwise differences with a number of permutation of 100 and significant level at $p<0.05$. Other than that, several results were determined by using this software which is molecular diversity and neutrality test.

CHAPTER FOUR

RESULTS & DISCUSSION

4.1 Introduction

All data for microsatellite analysis were produced from scored data of fragment analysis graph while for the mitochondrial analysis were produced from the sequences of COI gene. However, only three months of sampling samples were amplified by both markers. This is because the duration of the grow-out period of juvenile prawn generally took around 150 to 180 days in the river before the prawn is sexually matured to breed (L.R.D'Abramo *et al*.,2003). The stocking programme started in April 2007, it takes six months before the parental populations can grow to adult stage for them to able to breed with the wild populations in the river. The F1 prawns then will take another 6 months to grow up and this cycles were repeated for the next generation which the F2 and F3 generation. As only adult stage prawn were allowed to be sampled in this study, we derived a hypothesis, prawn sampled between the grow up period of each generation will be the same with generation before as both of these generation still not mate with each other yet.

The ecological analysis focused on benthic and water quality data. The benthic data was only collected for 6 months period which is from Mac 2009 to August 2009 from each station along the river as lack of resources to continuously collect the data for the whole year. In another hand, water quality data was collected throughout two years which from April 2008 to April 2010 in station along the river.

Table 4.1: Population type with total samples

Samples	***Total (Microsatellites Analysis)***	***Total (Mitochondrial Analysis)***
F1 (July 2008)	32	19
F2 (July 2009)	52	42
F3 (August 2010)	54	47
Parental	55	43
Wild Type	35	30
TOTAL	**228**	**171**

4.2 Molecular Study

4.2.1 DNA Analysis

The DNA was extracted from the freshwater prawn muscle tissue by using Vivantis GF1 DNA extraction kit. Even though the kit was time consuming, it is quite effective method as the kit extraction method yielded high molecular weight and high purity of DNA. Several samples were quantified using Spectrophotometer instrument and the result can be seen in table 4.2. Next, the quality of the DNA was visualized on a gel based on their intensity and thickness of the bands. The DNA samples were electrophoresed using 1% agarose gel in 1 TBE buffer and using Lamda Hind ladder to determine the approximate lengths (in bp). The gels were run at 60V for 50 minutes before been stained with Ethidium Bromide for 15 minutes. Figure 4.1 shows the UV image of 4 samples DNA extracted using this method by using Lamda Hind III ladder.

Table 4.2: Example of quantification results using Spectrophotometer.

Sample	***Concentration***	***260/280***
Wild Type 1	0.159	2.040
Wild Type 2	0.052	1.800
F1 1	0.103	2.020
F1 2	0.060	1.960
F2 1	0.092	0.000
F2 2	0.020	1.970
F3 1	0.113	1.980
F3 2	0.649	2.020
Parental 1	0.781	2.110
Parental 2	1.213	1.3062

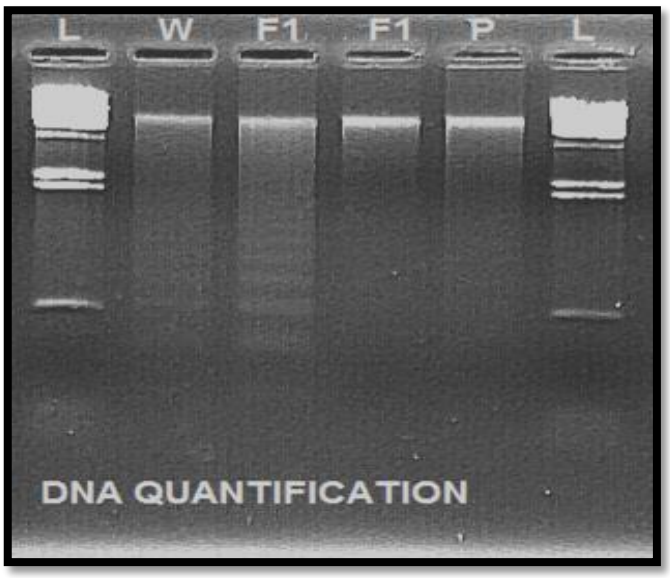


Figure 4.1: Examples of UV image of DNA Quality check

4.2.2 Isolation and Characterization of Novel Microsatellite (Type II Markers).

Random amplified microsatellite primers (RAMs) were optimized to produce sharp and clear bands on the gel. The amplified products were then cloned into yT&A cloning vectors (Yeastern Biotech Co., Ltd.). The blue-white screen technique was used to screen the successful ligations between the cloning vector and the DNA of interest. If the ligation was successful, the vector is then transformed into competent cell (bacteria) and become white colony in the presence of X-gal while become blue if the ligation was failed. The entire white colonies were then transferred to the new plate to sub clone before twelve recombinant clones from each primer were randomly selected for plasmid extraction. However, 9 of 12 clones were chose from VJ2 primers to be sequence while the other 6 clones were chose to sequence from T79122 primer. All the clone sequences were then submitted to the NCBI GenBank before use it to design novel microsatellite primers (Table 4.3).

There are 36 primers that been isolate from 9 recombinant clones from T79112 primer and 26 primers from 6 clones of VJ2 primers two. All primers are optimized to test their level of polymorphisms (Table 4.4). Out of 62 primers that were test, 14 of them showed reliable polymorphism (Table 4.5). Among 14 markers, one were dinucleotide microsatellites (VJ2_10a), three were trinucleotide repeats T79112_9f (T79112_7e, T79112_9c and T79112_9f) and the other ten sets of primers were tetranucleotide repeats were present. The characteristics of the 14 polymorphic loci are summarized in Table 4.5. The average allele number of these markers was 3.5/locus, with a range of 2 to 6 per locus. The observed heterozygosity ranged from 0.0952 to 0.9524. Two loci show higher than the expected heterozygosity which is T79112_8d and T79112_9f but not shows significant difference between both of the value. The other twelve loci show lower than expected heterozygosity. Such results could be explained by several hypotheses, including methodology bias, null alleles (Jarne and Lagode, 1996) and founder effect during the introduction (Mei et al., 2003). All these parameters and tests were computed using the Genepop version 3.4 software (Raymond and Rousset, 1995). These newly developed single locus microsatellites for western subspecies of *Macrobrachium rosenbergii* will indeed assist the documentation for stock enhancement programs that will be used in this study to document genotypic data, which could facilitate the knowledge on the genetic diversity of both the wild and parental population.

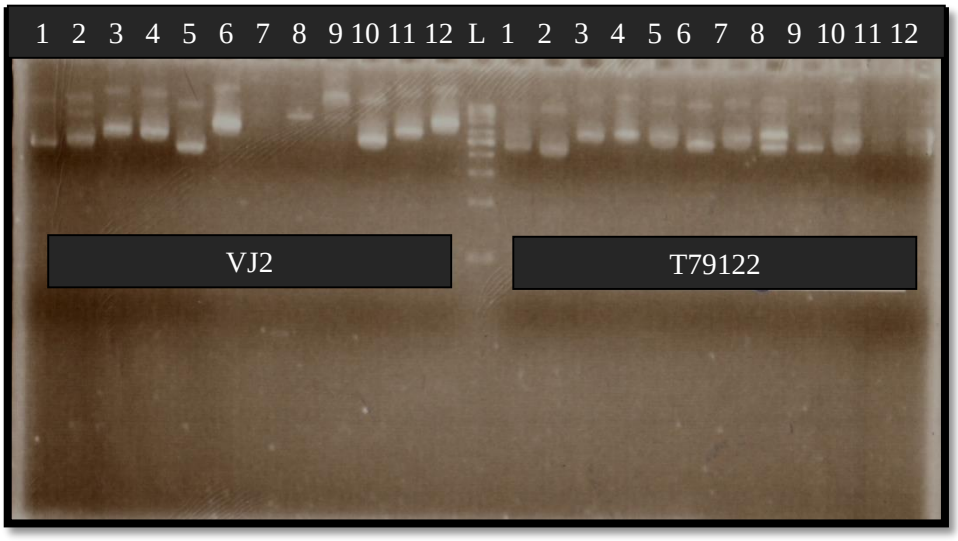


Figure 4.2 is the UV image for Plasmid Extraction of Vj2 and T79122 clones. Different clone was used for each lane with 1 kb DNA Ladder (Lonza, USA). Only clone number 3, 4, 6, 10, 11 and 12 were chose to sequence from Vj2 primer. Only clone number 1, 3, 4, 5, 6, 7, 8, 9, and 10 were chose to sequence from T79122 primer.

Table 4.3: All the clones and its primers with GenBank Accession Number

Clone Name	***GenBank Accession Number***	***Primer Design***
SBhMT79112_1	GQ853090.1	T79112_1a
		T79112_1b
		T79112_1d
SBhMT79112_2	GQ853091.1	T79112_3a
		T79112_3b
SBhMT79112_2	GQ853092.1	T79112_4a
		T79112_4b
		T79112_4c
		T79112_4d
		T79112_4e
SBhMT79112_4	GQ853093.1	T79112_5a
		T79112_5b
		T79112_5c
		T79112_5d
		T79112_5e
		T79112_5f
SBhMT79112_5	GQ853094.1	T79112_6a
		T79112_6b
SBhMT79112_6	GQ853095.1	T79112_7a
		T79112_7b
		T79112_7c
		T79112_7d
		T79112_7e

SBhMT79112_7	GQ853096.1	T79112_8a
		T79112_8b
		T79112_8c
		T79112_8d
		T79112_8e
		T79112_8f
SBhMT79112_8	GQ853097.1	T79112_9a
		T79112_9b
		T79112_9c
		T79112_9d
		T79112_9f
		T79112_9h
SBhMT79112_9	GQ853098.1	T79112_10a
		T79112_10b
SBhMVJ2_1	GQ853099.1	VJ2_3a
		VJ2_3b
		VJ2_3c
SBhMVJ2_2	GQ853100.1	VJ2_4a
		VJ2_4b
		VJ2_4c
		VJ2_4d
SBhMVJ2_3	GQ853102.1	VJ2_6a
SBhMVJ2_4	GQ853102.1	VJ2_10a
		VJ2_10b
SBhMVJ2_5	GQ853103.1	VJ2_11a
		VJ2_11b
		VJ2_11c
		VJ2_11d
SBhMVJ2_6	GQ853104.1	VJ2_12a
		VJ2_12b
		VJ2_12c
		VJ2_12d
		VJ2_12e
		VJ2_12f
		VJ2_12g
		VJ2_12h
		VJ2_12i
		VJ2_12j
		VJ2_12k
		VJ2_12l

Table 4.4: Optimized parameters for all the developed microsatellite primers

Primer Name	*Expected Size (bp)*	*Repeat Type*	*T^a(°C)*	*$MgCl_2$*
T79112_1a	183	AAG	60.0	1.2µl
T79112_1b	183	AAG	60.0	1.2µl
T79112_1d	285	CCA	57.0	1.2µl
T79112_3a	239	CATT	53.8	1.2µl
T79112_3b	227	GAT	52.0	1.2µl
T79112_4a	295	TAAAT	54.7	1.2µl
T79112_4b	197	CTT	54.3	1.2µl
T79112_4c	182	AGGG	53.8	1.2µl
T79112_4d	265	TTTTG	53.8	1.2µl
T79112_4e	178	AGAC	53.8	1.2µl
T79112_5a	161	CTTTT	49.9	1.2µl
T79112_5b	195	CT	52.3	1.2µl
T79112_5c	272	TTTA	53.8	1.2µl
T79112_5e	272	ATTTT	55.5	1.2µl
T79112_5f	244	TGA	52.8	1.2µl
T79112_6a	242	AAAT	48.5	1.2µl
T79112_6b	242	TGA	52.8	1.2µl
T79112_7a	161	CTTTT	48.0	1.2µl
T79112_7b	161	TCGA	52.8	1.2µl
T79112_7c	195	CT	48.0	1.5µl
T79112_7d	272	ATTT	48.5	1.5µl
T79112_7e	272	ATT	55.5	1.2µl
T79112_8a	263	GGAA	55.5	1.2µl
T79112_8b	255	AAAT	53.8	1.2µl
T79112_8c	245	AATT	55.7	1.2µl
T79112_8d	225	AAAG	52.8	1.5µl
T79112_8e	207	AAAT	52.3	1.2µl
T79112_8f	207	ATA	55.7	1.5µl
T79112_9a	198	TGT	52.8	1.2µl
T79112_9b	236	CAA	52.3	1.2µl
T79112_9c	157	TAA	52.8	1.2µl
T79112_9d	157	ACA	52.8	1.2µl
T79112_9f	179	AAC	52.8	1.2µl
T79112_9h	236	AAT	48.5	1.2µl
T79112_10a	299	GCTT	49.4	1.2µl
T79112_10b	239	GAAT	55.5	1.2µl
VJ2_3a	293	AGGC	56.0	1.2µl
VJ2_3b	286	TTGG	52.8	1.2µl
VJ2_3c	162	CAAT	55.7	1.2µl
VJ2_4a	188	TTAG	55.7	1.2µl
VJ2_4b	192	TGTT	54.2	1.2µl
VJ2_4c	176	TCTG	53.8	1.2µl
VJ2_4d	257	ACACAT	55.5	1.2µl
VJ2_6a	298	AAAG	52.3	1.2µl
VJ2_10a	246	CA	55.5	1.2µl

VJ2_10b	246	GAAA	55.7	1.2µl
VJ2_11a	278	AGGC	55.7	1.2µl
VJ2_11b	278	GGAA	55.7	1.2µl
VJ2_11c	208	GCCT	56.0	1.2µl
VJ2_11d	234	TAAA	48.9	1.2µl
VJ2_12a	192	AAAC	55.7	1.2µl
VJ2_12b	192	AACA	55.5	1.2µl
VJ2_12c	163	ACAA	56.0	1.2µl
VJ2_12d	163	ACA	55.5	1.2µl
VJ2_12e	163	ACAA	56.0	1.2µl
VJ2_12f	199	CAA	55.5	1.2µl
VJ2_12g	246	CAAA	55.7	1.2µl
VJ2_12h	246	ACAA	54.2	1.2µl
VJ2_12i	246	CAAA	53.8	1.2µl
VJ2_12j	156	ACAA	55.5	1.2µl
VJ2_12k	156	ACAA	55.7	1.2µl
VJ2_12l	156	CAAA	55.7	1.2µl

T^a = annealing temperature; $MgCl_2$ = Magnesium Chloride

Table 4.5: 14 novel polymorphic microsatellite markers from the *Macrobrachium rosenbergii*

Primer	***Primer sequences (5'-3')***	***Repeat Motive***	***Repeat Type***	***No. of alleles***	***Product size***	***T^a (°c)***	***H^o***	***H^e***	***PIC***
T79112_7d	**F:**TTCTTCCTTTGTCGACTTGT **R:**TGTCCAACCACTCTAACTCC	ATTT	Tetra	3	272	48.5	0.3333	0.2846	0.3291
T79112_7e	**F:**TTCTTCCTTTGTCGACTTGT **R:**TGTCCAACCACTCTAACTCC	ATT	Tri	4	272	55.5	0.7143	0.5006	0.5632
T79112_8d	**F:**GTATGCATGCGAAAATACTG **R:**GCGTCACTTATGATCCTGTT	AAAG	Tetra	3	225	52.8	0.2381	0.2973	0.4353
T79112_9c	**F:**TGACAGTGGTGATGATGATT **R:**TTTCACGCTGATAGTTGTTG	TAA	Tri	3	236	52.3	0.3810	0.3159	0.2981
T79112_9f	**F:**ACAATTGACAGTGGTGATGA **R:**ATATCCTCCTCGGTAGCTTT	AAC	Tri	5	179	52.8	0.7143	0.8060	0.4582
T79112_10a	**F:**CCACTTGTGAGAAGCCATAA **R:**TTCGTATGCCGTACAATATG	GCTT	Tetra	2	299	49.4	0.0952	0.0929	0.1901
VJ2_4c	**F:**TGTTTATCGGGGTCATAAGT **R:**GGCTTTACAAAACCATCAAC	TCTG	Tetra	3	192	53.8	0.9524	0.7398	0.3678
VJ2_6a	**F:**AGGTGGATTTCTGTAGCTGA **R:**CATTAGTCCCTCCAAAACTG	AAAG	Tetra	3	298	55.5	0.9048	0.7991	0.3599
VJ2_10a	**F:**CGACAGAATCAAGGCTTTAC **R:**GATTCCAAACCTCCTCTTCT	CA	Di	6	246	55.5	0.4353	0.3762	0.5111
VJ2_10b	**F:**CGACAGAATCAAGGCTTTAC **R:**GATTCCAAACCTCCTCTTCT	GAAA	Tetra	3	246	55.7	0.2534	0.1982	0.1821
VJ2_11d	**F:**TTTTCCTACATTGGCTCAGT **R:**TTTTTAGGCCCCAATGGT	TAAA	Tetra	2	234	48.9	0.7633	0.6733	0.2101
VJ2_12a	**F:**CTAAAGTCAAACTGCCATGA **R:**CTCCTATTGTTTTGGTCCTG	AAAC	Tetra	4	192	55.7	0.3211	0.2181	0.3419
VJ2_12i	**F:**GAGAAAACAAAGGAAGAGCAA **R:**CCTCTGTTGTTTATCCCGTA	CAAA	Tetra	4	246	53.8	0.2631	0.2113	0.3819
VJ2_12j	**F:**AAACGTCTACGGGATAAACA **R:**GTCAATTGGGATATTGATGG	ACAA	Tetra	4	156	55.5	0.2532	0.1621	0.1233

T^a = annealing temperature; H^e = expected heterozygosity; H^o = observed heterozygosity; PIC = Polymorphic Information Content

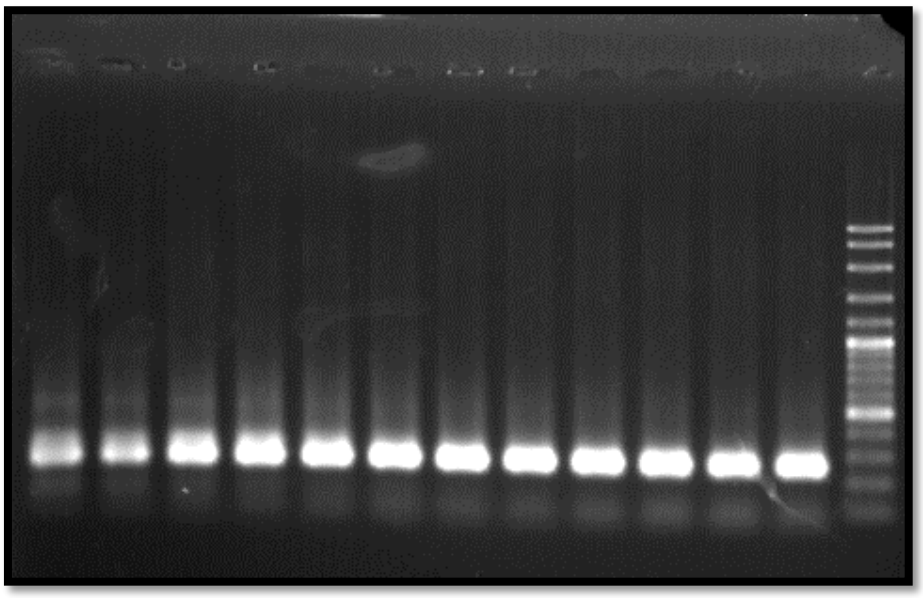


Figure 4.3 is the UV image for Primer Vj2_12i. For each lane, different temperature was used. The temperature for the first lane was 48.0°C, and increase for each lane 48.5°C, 48.9°C, 49.9°C, 51.1°C, 52.3°C, 52.8°C, 53.8°C, 54.2°C, 55.5°C, 55.7°C and 56.0°C respectively. Lane L marked as the lane for 20 bp DNA Ladder (Lonza, USA).

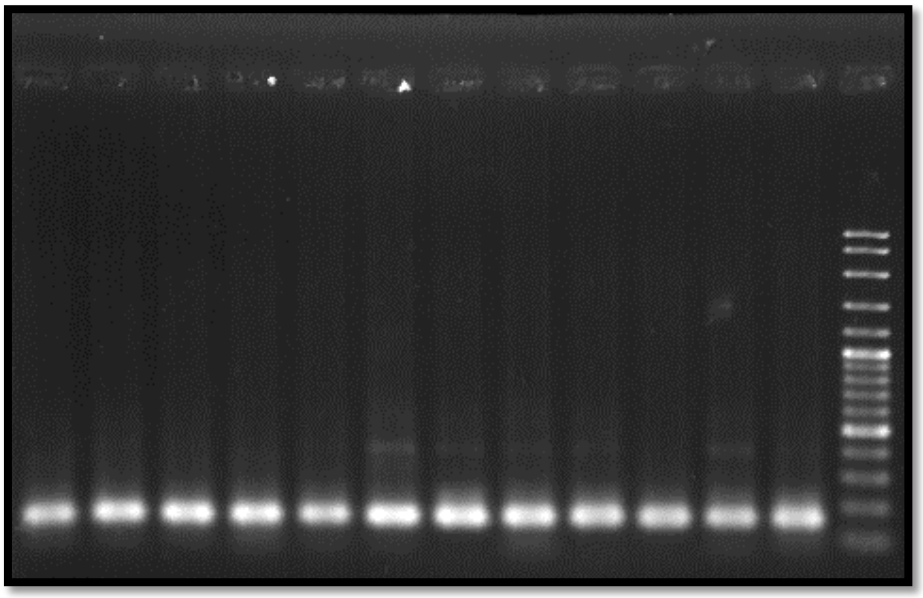


Figure 4.4 is the UV image for Primer T79122_12j. For each lane, different temperature was used. The temperature for the first lane was 48.0°C, and increase for each lane 48.5°C, 48.9°C, 49.9°C, 51.1°C, 52.3°C, 52.8°C, 53.8°C, 54.2°C, 55.5°C, 55.7°C and 56.0°C respectively. Lane L marked as the lane for 20 bp DNA Ladder (Lonza, USA).

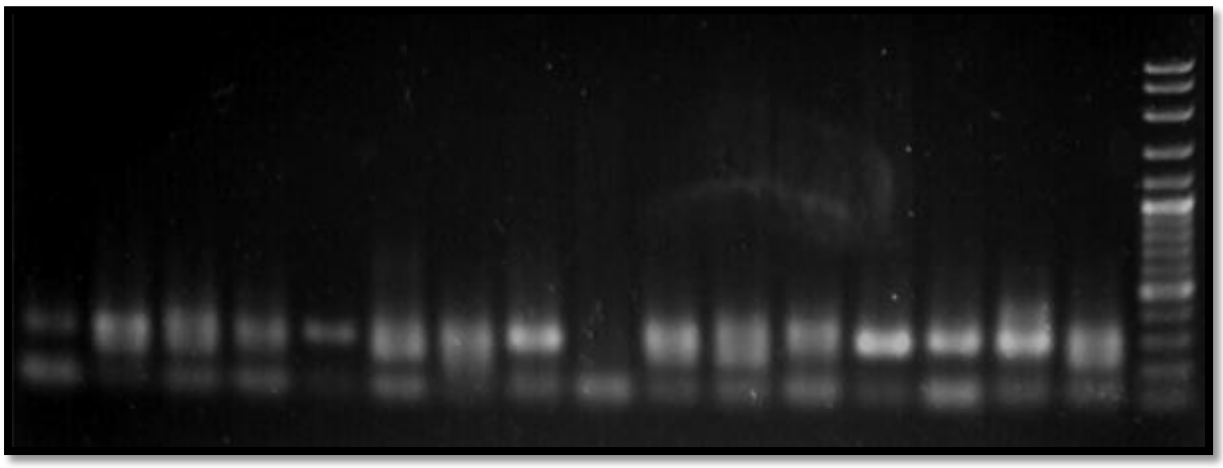


Figure 4.5 is the UV image of Polymorphisms Testing of T79122_7e primer on 1% agarose gel with 20 bp DNA Ladder (Lonza, USA).

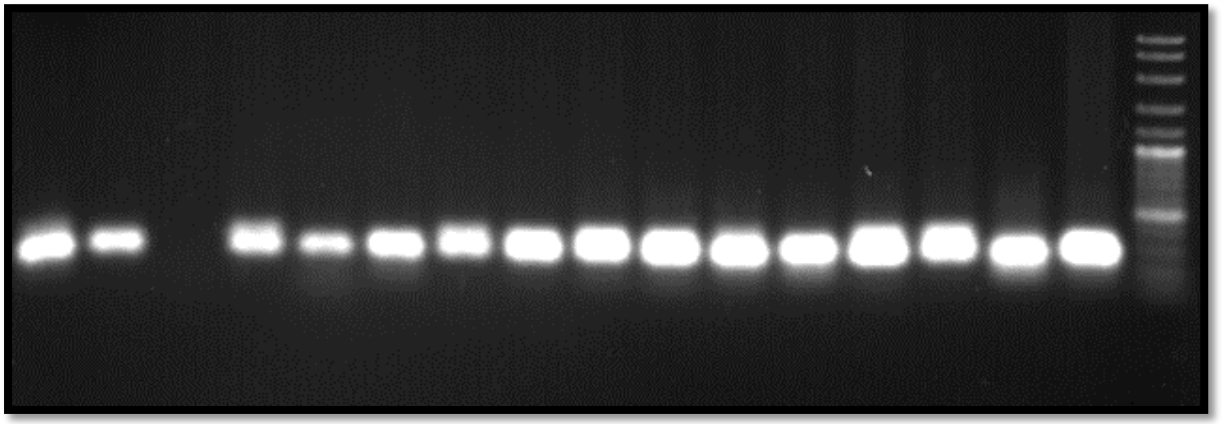


Figure 4.6is the UV image of Polymorphisms Testing of Vj2_10b primer on 1% agarose gel with 20 bp DNA Ladder (Lonza, USA).

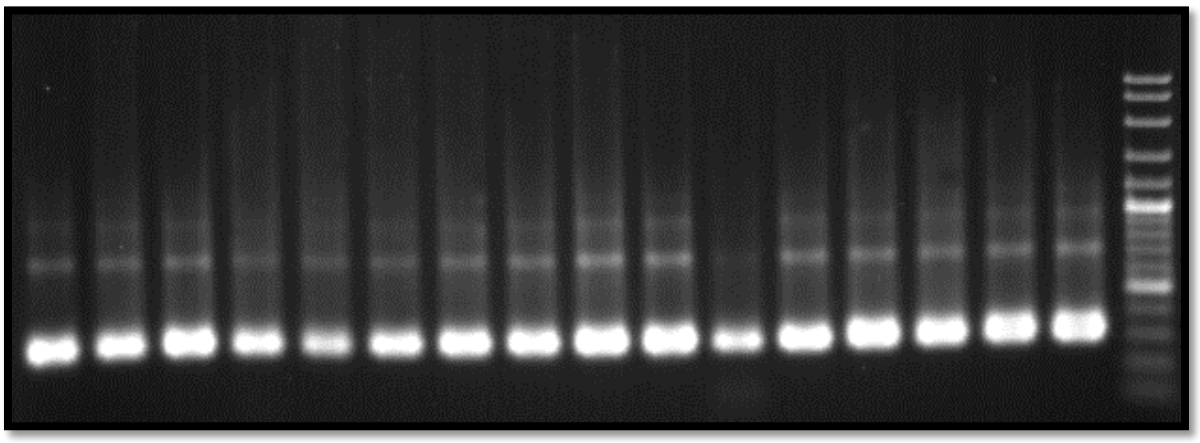


Figure 4.7 is the UV image of Polymorphisms Testing of Vj2_3c primer on 1% agarose gel with 20 bp DNA Ladder (Lonza, USA).

4.2.3 Development and Characterization of EST SSR (Type I Marker)

A total of 23595 EST sequences were assembled. 149 sequences with microsatellites were extracted. 15 dinucleotide, 75 trinucleotide, 7 tetranucleotide, 25 pentanucleotide and 27 hexanucleotide repeats were found. Trinucleotide microsatellites were most abundant, possibly due to suppression of other microsatellites in coding regions due the risk of frameshift mutations (Rajeev et al. 2005). For EST sequences with microsatellite, it was observed that 105 (70%) of the sequences were unannotatable, 24 (16%) were known gene products, where 20(13%) sequences were predicted. Overall for the entire transcriptome, 18366 (77%) were unannotatable, 2827(12%) sequences were known gene products and 2402(10%) sequences were predicted (Qi Bin, 2011). A comparison of EST with microsatellite and overall transcriptome EST showed similar distribution pattern for cellular component ontology (Figure 4.8). Most of EST with microsatellite was found in the region that controls metabolic function

Forty one sequences with microsatellites were selected from 149 sequences randomly to be tested at genomic level, on 32 unrelated individual prawns from Kampung Sungai Timun (Table 4.6). All primers worked but only 21 primers were found polymorphic at genomic level (Table 4.7). A total of 118 alleles were obtained across the 21 loci with the average about 5.62 alleles per loci. The overall allele frequency ranged from 0.0156 to 0.9531. The observed heterozygosity ranged from 0.0938 to 0.9375. Two loci show lower than the expected heterozygosity which is MR4 and MR29 but not shows significant difference between both of the value. The other 19 loci show higher than expected heterozygosity. Mistyping of heterozygous individuals might explain why the observed heterozygosity in this all the population is significantly difference with the expected heterozygosity.

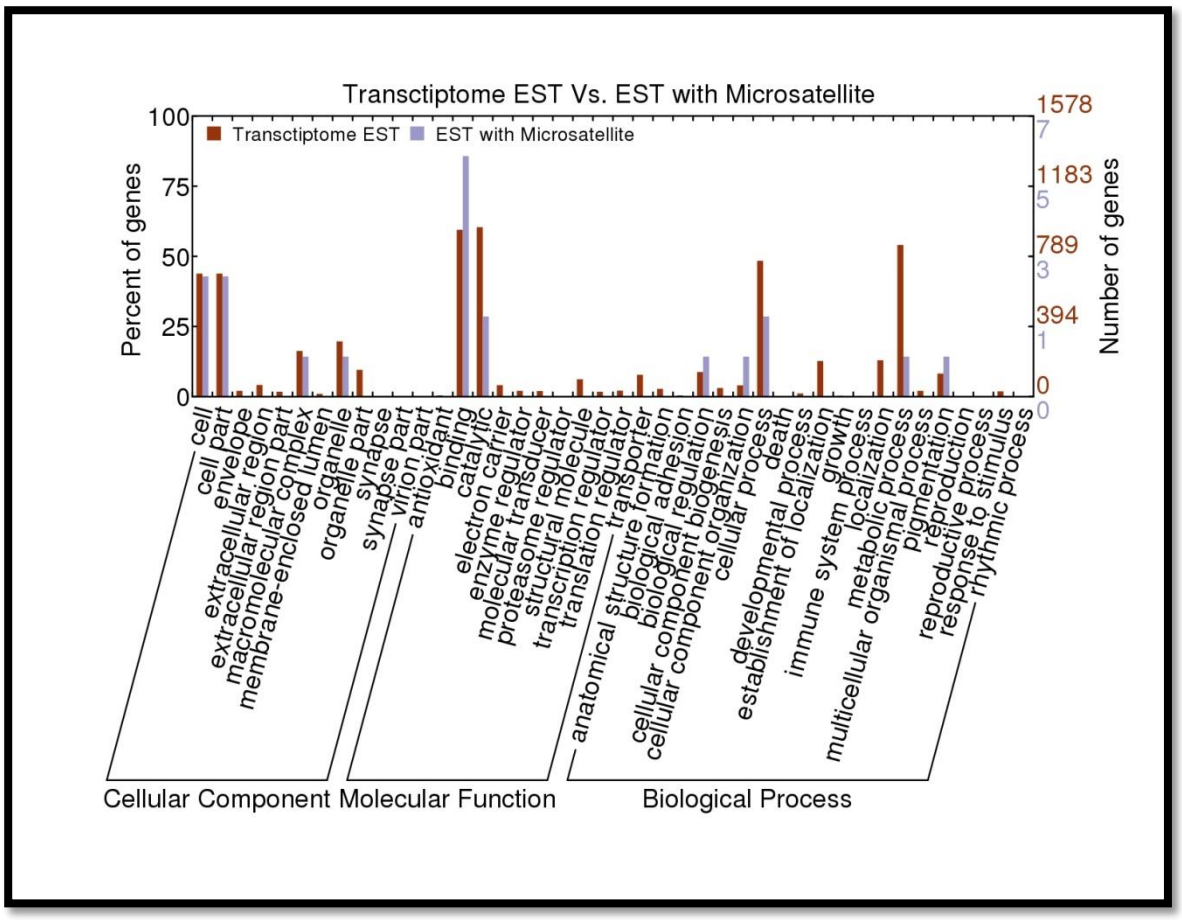


Figure 4.8: Transcriptome EST Vs EST with Microsatellite gene oncology graph (Qi Bin, 2011)

Table 4.6: Validation of all the EST SSR developed

Primer	*Sequence*	*Expected Size*	*Repeat Motive*	*Repeat Type*	*T^a (°c)*	*Single Band*	
						DNA	*CDNA*
MR 1	F AGTCTTTAGGTGAAACTGG R- CCTAACCCTGTAGCAAATA	641	ACT	Tri	51.9	Y	X
MR 2	F -AGGTTCAGGTTACTGAAGAG R – CAGTCGACAAGATAGTGAAT	706	TGTTTG	Hexa	55.4	Y	X
MR 3	F - GGCATTTCTCTCATAAGC R - GAAAGACACTCGTCGAAC	695	TCT	Tri	48.0	Y	X
MR 4	F - TCATGTCTCTTGAGTCTTTC R – GTTATCTGATCGTCACAGTT	311	AT	Di	53.9	Y	X
MR 5	F - ACTTGACTTCGATCTTGAC R – GAGAGCTAGATGACACAGAA	302	CTTGAT	Hexa	54.0	Y	Y
MR 6	F - CTATTGCCAGGCCAAAAA R – TACCACCAAACTCCACTAC	200	GGCA	Tetra	58.0	Y	Y
MR 7	F - AAGTCTTGTTGTTGGACAG R – ACTATAGGGGGTTCAGTAAA	303	CAAAAA	Hexa	54.0	Y	Y
MR 8	F - CCAGTGATTGAAATCGTC R – AAGAAGAGGCCACGACAG	284	CGC	Tri	55.3	Y	Y
MR 9	F - ACATGTCGTTTACAGTTACC R – AACAGCTACAACAACGTAAG	203	GAA	Tri	**NI**	X	X
MR 10	F - CTTGGTACAATGCTTATCAC R – AACATCACTCTGAAGTCTTG	320	GTG	Tri	51.9	Y	Y

MR 11	F - AAAGCAGAGAAGCGTAAG R – GTCACCTGCATTGTGTTT	201	GAC	Tri	49.1	Y	Y
MR 12	F - GAGTCTGTACATTCTGTGTT R – TGGCAGCCAATTCAGTATAG	200	GAG	Tri	**NI**	X	X
MR 13	F - ACATTCCACACTAAAGTACG R – CAAGACATGCACTCTTTACT	204	AGC	Tri	50.9	Y	Y
MR 14	F - GTAAGGCTGATATACCTGAG R – TGTAGAGGAATCTGTTGAAG	322	GGT	Tri	53.9	Y	Y
MR 15	F - TTAGACTGTTGACCACCATA F - CACCTCAAGGATACTCTCA	298	CTG	Tri	50.6	Y	Y
MR 16	F - GGTGGAATATTCTTGGTACT R – GAGCGAAGAGATAGACAAA	213	TGTCG	Penta	**NI**	X	X
MR 17	F - AGTCAAAACGAAGACTCC R – GTGTCTCTTCTGGGTAAATA	321	AGAGGC	Hexa	**NI**	X	X
MR 18	F - GTACGTGTTGCGTACTTG R – CAGAACGTATGTCCACAG	213	**NI**	**NI**	53.9	Y	X
MR 19	F - AGCGTTAAGTAGCATTACAC R – ATGACATTTGACAGAGGAC	260	AGGCG	Tri	56.9	Y	Y
MR 20	F - CTTGTGCAAGAGATTAT	248	AAG	Tri	55.3	Y	Y
MR 21	R - GGTGATGCCTTTGTTATAC F - GAGTGATGATGACGATGATA R - CTCTCCTGAACTTTCTTCTT	288	GAG	Tri	52.4	Y	Y

MR 22	F - CTGACTAGATAGCATTTTCC R - GTTACCAAAAGAACAGTCAG	706	TTGAT	Penta	55.3	Y	X
MR 23	F - AGCCTTCAAGATATTCTCAC R - GTTAGGAGAAGGTGGATAAA	224	AATTC	Penta	52.4	Y	Y
MR 24	F - GTAGTTAGCGGAACATGG R - TATCCCTCTTTCTCTCTCTC	303	GGGAG	Penta	55.3	X	X
MR 25	F - CAACAATGCTGGAACAAAG R - TACCATCGTCCTCCTGAT	200	GGA	Tri	55.3	X	X
MR 26	F - TATTTTCCTTTCTCCAGAGG R - CTTCTCCATCATCATCATC	202	AACC	Tetra	55.3	Y	Y
MR 27	F - TCAGTCTCTAGACTCCATGA R - GTGGATAACCTTTACAGACA	282	ATG	Tri	55.3	Y	Y
MR 28	F - ATGTAGATTCCTTGGTGAG R - GTATCCAGCTCAAGAACAC	279	TCTGT	Penta	55.3	Y	Y
MR 29	F - GGGAAAAGCGACACATATAA R - ACTAGTAGCTGCTCTTTGTG	213	GAA	Tri	55.3	Y	Y
MR 30	F - TCAGAGGGAGATGAGTATAA R - GAACTGTATGCTCAGTTCAT	296	GT	Di	55.3	X	X
MR 31	F - CAGCGTATACTTTCTTCACT R - CCATTTCCATTGGTGTAG	274	CGTGAC	Hexa	53.3	Y	Y
MR 32	F - TCATCGTCCTCTTAGTCAT R - AGTAGAGGAAGGTGTAGGAG	267	TCC	Tri	**NI**	X	X

MR 33	F - GATGTCTCTCAAGGACTTCT R - CCAATAGTAAGGCACTCTAA	324	GAC	Tri	**NI**	X	X
MR 34	F - ACGTTCACTTGACATCAG R - CTGGTACTTCTTCTTCCTCT	266	AGATGA	Hexa	**NI**	X	X
MR 35	F - CCTGTGATCGAGTGAGTA R - GGGTGTGTATGAAAAATGTG	275	CTG	Tri	**NI**	X	X
MR 36	F - GTTTCGTGAATCCTCAAG R - ATGATGATGGTGATCTACG	220	TTG	Tri	**NI**	X	X
MR 37	F - TCTTCTTCTAGCTTCTAGGTC R - CTTCTTTGTTTACGTTCCTC	303	TTC	Tri	52.7	Y	Y
MR 38	F - GTTCAAGAAGCTGAATGTAG R - CAGTCTGTCAGTTTCAAATC	283	TTTTA	Penta	59.6	X	Y
MR 39	F - GAGCATGACATTGTGAAGA R - GAGTAAAGTGCCCAGGAC	200	GTG	Tri	50.9	Y	Y
MR 40	F - TATGAAGTAACCTGCTCTGT R - GCAGCTCCTATAGAATTAAG	300	TAA	Tri	48.0	Y	Y
MR 41	F - ACTAACACTAGGTCAAAGCA R - AGAAGAAGAAGAAGATGTCG	722	TCT	Tri	**NI**	X	X
MR 42	F - CTAAACTCAAGCCAGAACA R - TGCAGTTGTTACGTGTACTA	7009	CCAGAA	Hexa	**NI**	X	X

T^a = annealing temperature; NI = No Information; Y = Produce single band; X = Not produce single band

Table 4.7: 21 novel polymorphic EST markers from the *Macrobrachium rosenbergii*

Locus	*Primer sequences (5'-3')*	*Repeat Motive*	*Repeat Type*	*No. of alleles*	*Product size*	*T^a (°c)*	*H^o*	*H^e*	*PIC*
MR1	**F:**AGTCTTTAGGTGAAACTGG **R:**CCTAACCCTGTAGCAAATA	AAT/ACT	Tri	3	655	51.9	0.5556	0.5003	0.402
MR4	**F:**TCATGTCTCTTGAGTCTTTC **R:**GTTATCTGATCGTCACAGTT	AT	Di	11	311	53.9	0.7333	0.7599	0.721
MR5	**F:**ACTTGACTTCGATCTTGAC **R:**GAGAGCTAGATGACACAGAA	CTTGAT	Hexa	5	412	54.0	0.5938	0.4802	0.4400
MR6	**F:**CTATTGCCAGGCCAAAAA **R:**TACCACCAAACTCCACTAC	GGCA	Tetra	6	200	58.0	0.9375	0.7192	0.665
MR7	**F:**AAGTCTTGTTGTTGGACAG **R:**ACTATAGGGGGTTCAGTAAA	CAAAAA	Hexa	5	298	54.0	0.5312	0.4345	0.3700
MR8	**F:**CCAGTGATTGAAATCGTC **R:**AAGAAGAGGCCACGACAG	CGC	Tri	9	275	55.3	0.9032	0.8509	0.8160
MR10	**F:**CTTGGTACAATGCTTATCAC **R:**AACATCACTCTGAAGTCTTG	GTG	Tri	9	237	51.9	0.5484	0.4907	0.4690
MR13	**F:**ACATTCCACACTAAAGTACG **R:**CAAGACATGCACTCTTTACT	AGC	Tri	2	204	50.9	0.8125	0.4901	0.3660
MR15	**F:**TTAGACTGTTGACCACCATA **R:**CACCTCAAGGATACTCTCA	CTG	Tri	5	522	53.9	0.4000	0.3893	0.3480
MR19	**F:**AGCGTTAAGTAGCATTACAC **R:**ATGACATTTGACAGAGGAC	AGGCG	Penta	6	255	56.9	0.8710	0.6943	0.6260
MR20	**F:**CTTGTGCAAGAGATTATCC **R:**GGTGATGCCTTTGTTATAC	AAG	Tri	4	247	55.3	0.9375	0.7316	0.6670
MR21	**F:**GAGTGATGATGACGATGATA **R:**CTCTCCTGAACTTTCTTCTT	GAG	Tri	6	282	52.4	0.6562	0.5055	0.4510
MR23	**F:**AGCCTTCAAGATATTCTCAC **R:**GTTAGGAGAAGGTGGATAAA	AATTCC	Hexa	4	222	52.4	0.8438	0.5670	0.4600
MR26	**F:**TATTTTCCTTTCTCCAGAGG **R:**CTTCTCCATCATCATCATC	AACC	Tetra	6	200	55.3	0.9375	0.6230	0.5430

MR27	**F:**TCAGTCTCTAGACTCCATGA **R:**GTGGATAACCTTTACAGACA	ATG	Tri	4	283	55.3	0.9375	0.6503	0.5860
MR28	**F:**ATGTAGATTCCTTGGTGAG **R:**GTATCCAGCTCAAGAACAC	TCTGT	Penta	3	411	55.3	0.1250	0.1205	0.1140
MR29	**F:**GGGAAAAGCGACACATATAA **R:**ACTAGTAGCTGCTCTTTGTG	GAA	Tri	5	209	53.3	0.7200	0.7249	0.6730
MR31	**F:**CAGCGTATACTTTCTTCACT **R:**CCATTTCCATTGGTGTAG	CGTGAC	Hexa	4	445	53.3	0.0938	0.0923	0.0890
MR37	**F:**TCTTCTTCTAGCTTCTAGGTC **R:**CTTCTTTGTTTACGTTCCTC	TTC	Tri	8	308	52.7	0.7500	0.5749	0.4990
MR38	**F:**GTTCAAGAAGCTGAATGTAG **R:**CAGTCTGTCAGTTTCAAATC	TTTTA	Penta	6	285	59.6	0.8387	0.5648	0.4830
MR40	**F:**TATGAAGTAACCTGCTCTGT **R:**GCAGCTCCTATAGAATTAAG	TAA	Tri	7	303	48.0	0.8065	0.7070	0.6380

T^a = annealing temperature; H^e = expected heterozygosity; H^o = observed heterozygosity; PIC = Polymorphic Information Content

4.2.4 Genetic Diversity Study of Macrobrachium rosenbergii

The development of markers process produced 35 polymorphic markers consist of 21 Type I marker (EST SSR) and 14 Type II marker (Genomics SSR). However, not all the primers were used to assist in verifying the genetic diversity as lack of resources such as time and money. So, 13 primers were chosen based on the PIC value of each primer. Co-dominant markers exhibiting PIC value higher than 0.5 are regarded as highly informative (Botstein *et al.*, 1980) Another 5 primers were selected from previous studies done by (Subha *et al.*, 2009)

All samples were then tested with all of the selected optimized primers. The banding patterns were first observed in 1% Agarose gel before later diagnosed in more sensitive equipment which is Fragment Analyzer. These banding patterns were locus specific and were used to differentiate between the individual and between the populations. All amplified microsatellites bands were scored with GeneMapper® (Applied Biosystems) and Peak Scanner Software v1.0 software (Applied Biosystems). The scored data were transferred to the table in appendix. Next, the scored data was convert into GENEPOP (Raymond & Rousset 1995) format by using CONVERT v1.31 (Glaubitz, 2004) as an input file for Microchecker software v2.2.3 (Van Oosterhout *et al.*, 2004). This software was used to check genotyping errors due to non-amplified alleles (null alleles), short allele dominance (large allele dropout) and the scoring of stutter peaks, and also detects typographic errors.

Microchecker has showed that there are null alleles present in several loci in all population based on the methods described by Chakraborty *et al.*, (1992) and Brookfield (1996)(Table 4.8). As much as 8 loci were showed null alleles in at least 4 out of 5 populations (Locus MR13, MR38, LM13, LM16, LM04, LM08, T917_7e and T917_7d). There are no

sign of large alleles drop out and scoring error due to stuttering. A null allele can be defined as any allele at a microsatellite locus that is only weakly amplified or not visible after amplification and separation (O'Connell and Wright, 1997) and is recognized, together with population subdivision, as a major factor in depression of observed heterozygosity, compared with that expected on the basis of Hardy-Weinberg equilibrium. The prime reason for null alleles is thought to be deletion and insertion within the priming site of microsatellite DNA leading to an allele that will not amplify (Callen *et al.*, 1993).

Table 4.8: Microchecker result

Locus	***Null Present***				
	Wild Type	**Parental**	**F1**	**F2**	**F3**
MR4	NO	NO	NO	NO	NO
MR6	YES	YES	NO	YES	NO
MR8	YES	NO	NO	NO	NO
MR13	YES	YES	NO	YES	YES
MR19	YES	NO	NO	YES	NO
MR20	NO	NO	NO	NO	NO
MR27	NO	NO	NO	NO	NO
MR38	YES	YES	NO	YES	YES
MR40	NO	NO	NO	NO	NO
LM13	YES	YES	YES	YES	YES
LM08	YES	YES	YES	YES	YES
MBR5	NO	NO	NO	NO	NO
LM16	YES	YES	YES	YES	YES
LM04	YES	YES	YES	YES	NO

T917_7e	YES	YES	YES	YES	YES
T917_7d	YES	YES	NO	YES	YES
T917_9c	NO	NO	NO	NO	NO
VJ2_10a	NO	NO	NO	NO	NO

4.2.4.1 Alleles frequency

The information on the genetic variability levels of giant freshwater prawn *Macrobrachium rosenbergii* can be depicted by the number of alleles per locus of microsatellite markers. All populations showing high genetic diversity as the numbers of alleles were found in each primer is within 6 to 26 alleles. This showed that higher mutation rate occur in these 18 loci. Generally, among all 18 primers, EST SSR markers showed the higher allele numbers than Genomics SSR markers. However, in the table showed that Genomics SSR markers have higher Effective number of alleles (Kimura and Crow, 1994). The higher number of alleles was found in MR6 locus with 26 alleles with the lowest belongs to LM16 and T917_9c loci with 6 alleles. Figure 4.9 to figure 4.15 showed the example of allele's frequency of 6 different primers in each of the population. The others alleles frequencies of primers were shown in the Appendix F. A total of 131, 172, 133, 195 and 185 different alleles were found in all the loci in the Wild type, parental, F1, F2 and F3 populations respectively, which explains the extent of allelic richness possessed by all five populations.

There are two statistics frequently used to measure genetic diversity which is expected heterozygosity and allelic richness. However, gene diversity is probably used more frequently because the allelic richness of a sample is affected by the size of sample. In order to overcome this problem, rarefaction has been used as statistical method to produce unbiased estimates of

allelic richness (Hurlbert 1971; Smith& Grassle 1977; Leberg 2002). Based on the allelic richness (Figure 4.16) that were obtain by using HP-Rare software showed that all loci except MR4 have the same pattern of which Parental population have the highest allelic richness mean (7.12) while the F1 have the lowest mean (5.54). However the allelic richness is increasing in the F2 (6.80) and F3 population (6.66). Wild type population has a little bit higher than F1 population which is 5.87.

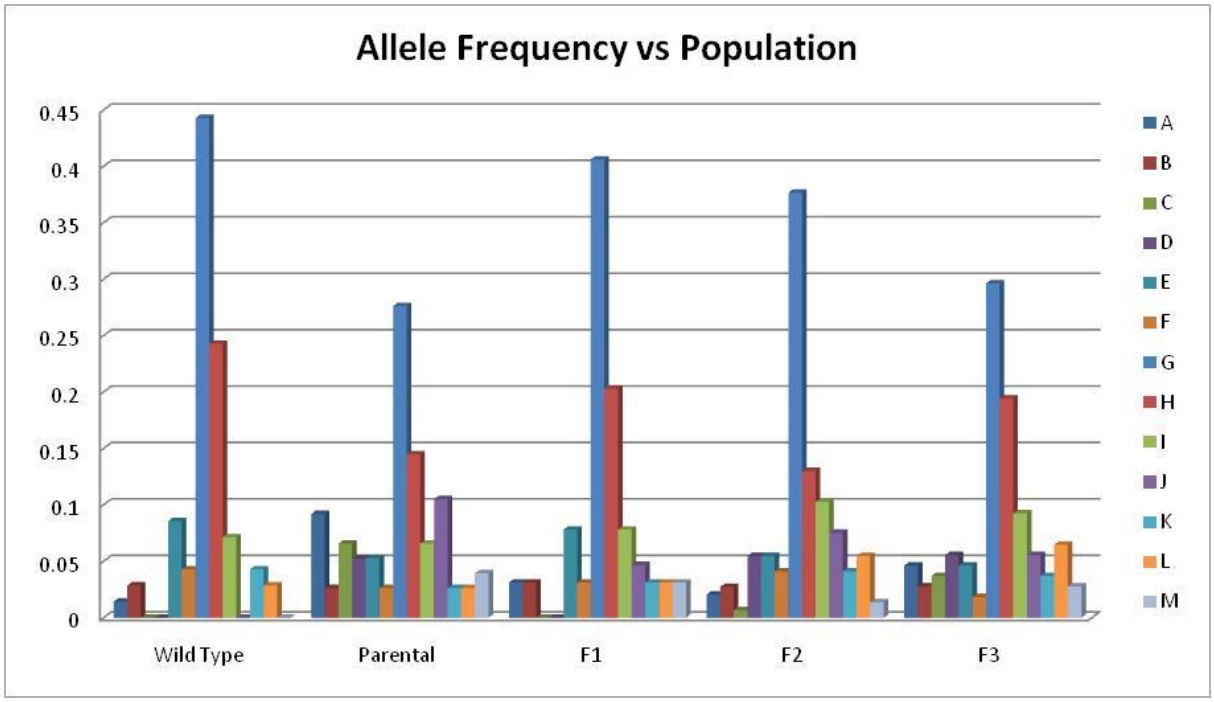


Figure 4.9: Percentage of allelic frequencies for all 5 populations produced by using Microsatellites primer (MR4) (Alphabets represented as different alleles)

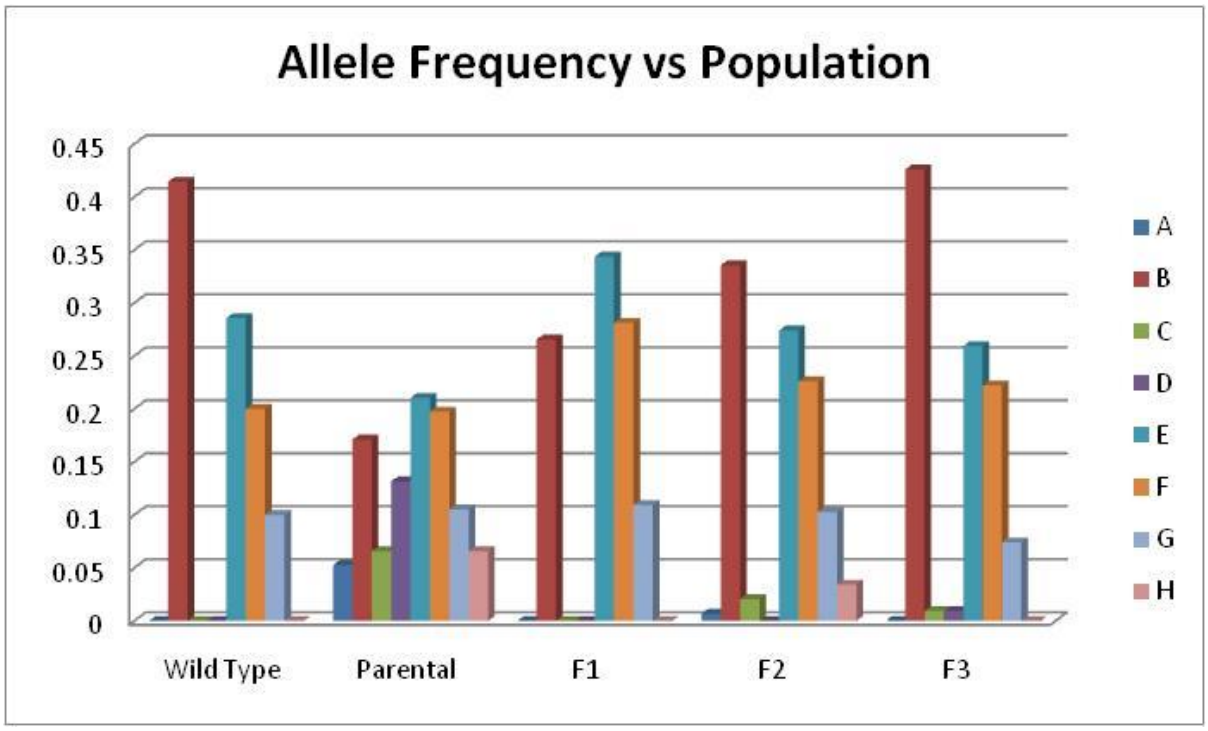


Figure 4.10: Percentage of allelic frequencies for all 5 populations produced by using Microsatellites primer (MR20) (Alphabets represented as different alleles)

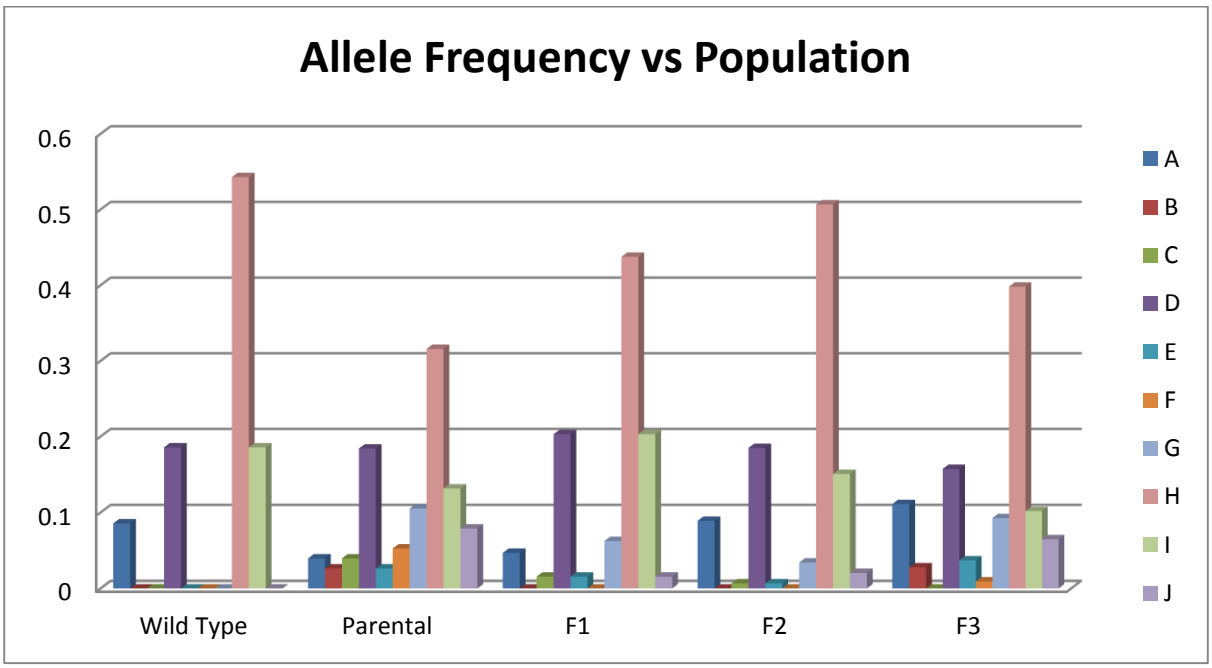


Figure 4.11: Percentage of allelic frequencies for all 5 populations produced by using Microsatellites primer (MR27) (Alphabets represented as different alleles)

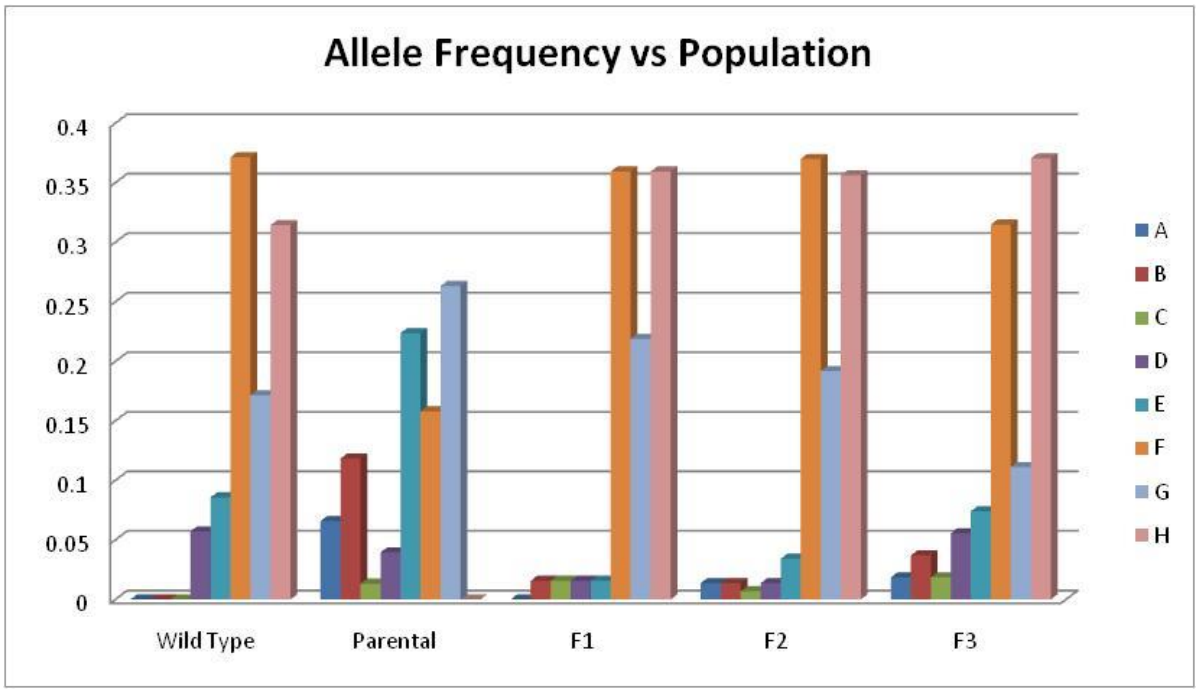


Figure 4.12: Percentage of allelic frequencies for all 5 populations produced by using Microsatellites primer (MR40) (Alphabets represented as different alleles)

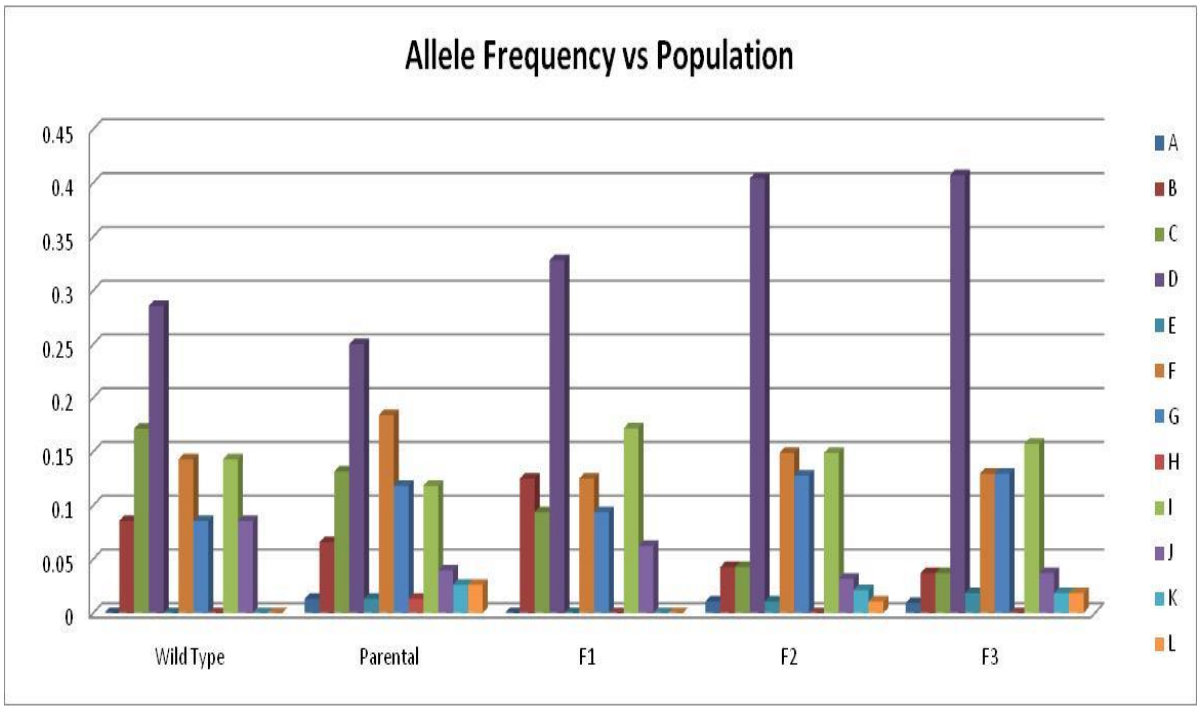


Figure 4.13: Percentage of allelic frequencies for all 5 populations produced by using Microsatellites primer (LM08) (Alphabets represented as different alleles)

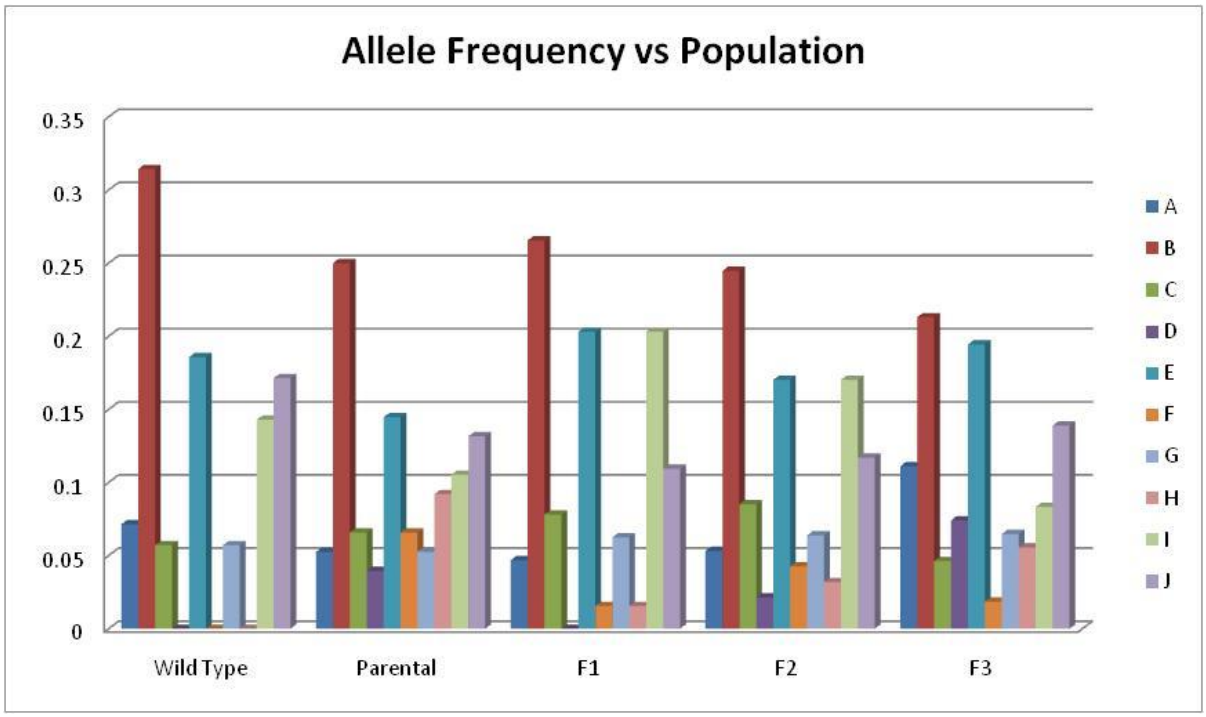


Figure 4.14: Percentage of allelic frequencies for all 5 populations produced by using Microsatellites primer (LM04) (Alphabets represented as different alleles)

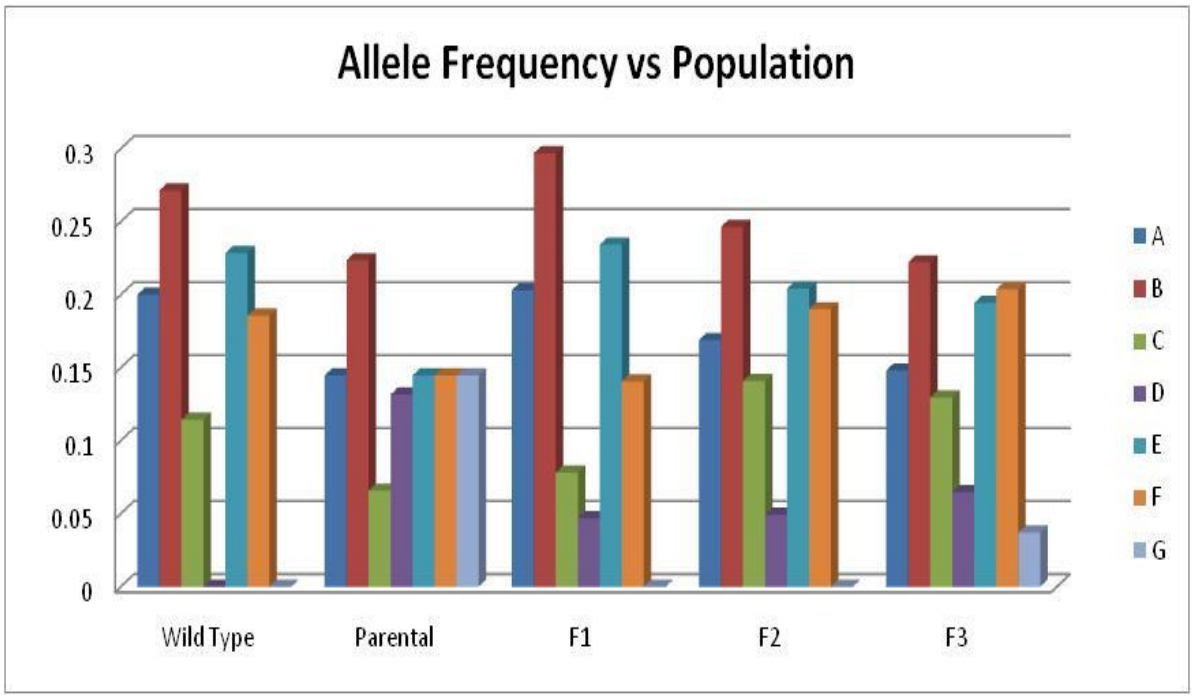


Figure 4.15: Percentage of allelic frequencies for all 5 populations produced by using Microsatellites primer (VJ2_10a) (Alphabets represented as different allele.

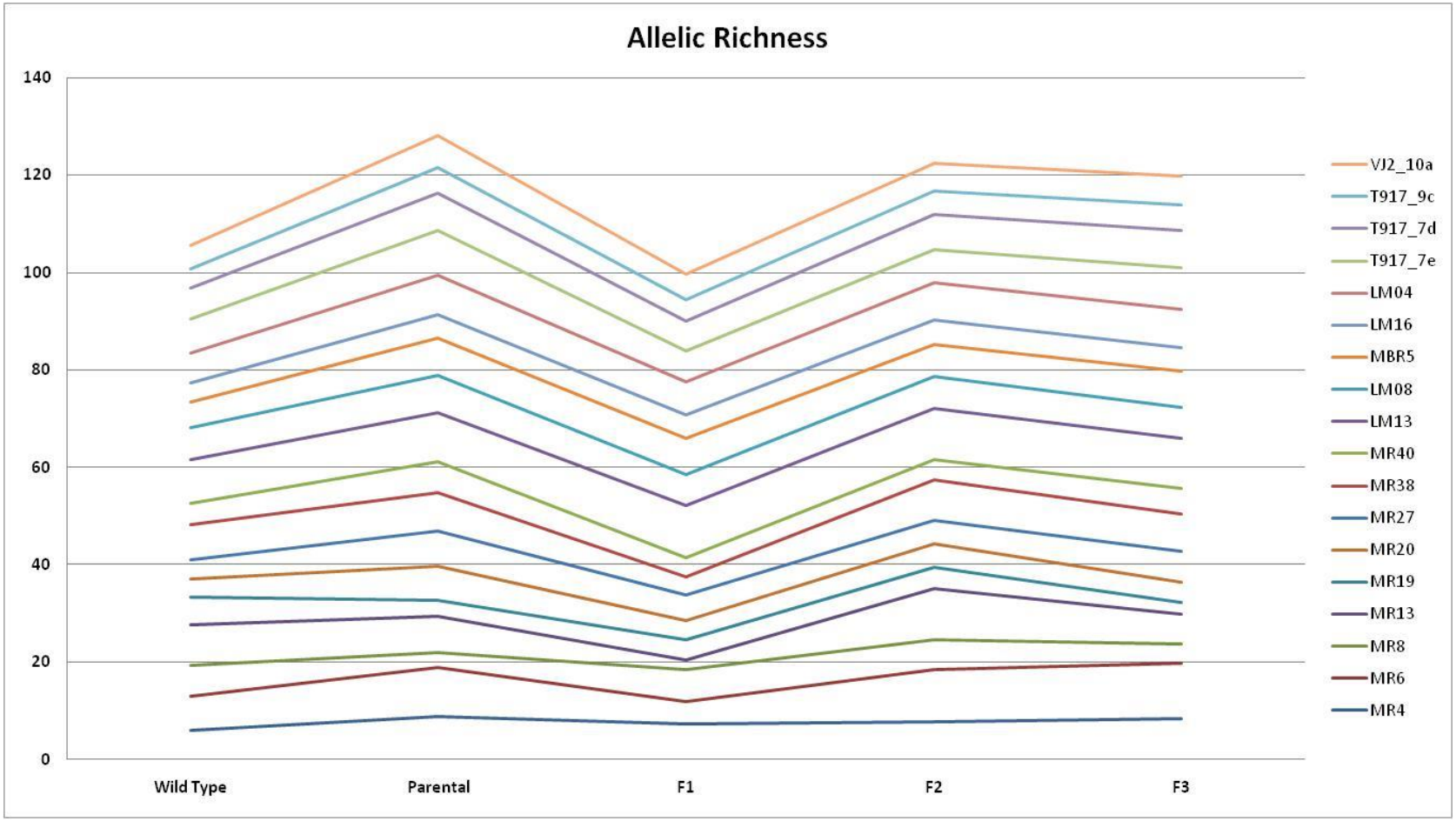


Figure 4.16: Allelic Richness of 18 loci for each population

4.2.4.2 Statistical results

There a lot of statistical result can be computed from the data. Table 4.9 showed the result for each primer and population. Different software were used to obtain several statically result based on their effectiveness of producing accurate value. POPGENE version 1.31 software was used to calculate number of alleles, number of effective alleles, observed heterozygosity and expected heterozygosity (Levene, 1949). The highest mean observed heterozygosity found in the F1 population with a value of 0.7188 while Wild Type Population had the lowest value of 0.5095. The effective alleles are the alleles that we can use to differentiate among individual and population (Kimura and Weiss, 1964). The tables showed MR6 and LM13 have the highest mean effective alleles with a value of 12 alleles while several primers such as MR8, MR19, LM16 and T917_9c have the lowest mean effective alleles with the value of 3 alleles. However out of 18 primers, 15 primers showed more than 40% of observed alleles were effective alleles with VJ2_10a the highest (79.70%). Primers MR8, MR19 and MR27 showed lower than 40% with the value of 19.40%, 23.55% and 37.64% respectively.

The Chi-square test were used to test for deviations from Hardy- Weinberg equilibrium in each locus in each population. The probabilities were showed in Table 4.9 while the calculations were showed in Appendix. Hardy Weinberg P- Value and Population differentiations are inferred using Wright's F-statistics (F_{IS} and F_{ST}) by POPGENE version 1.31software. Among the ninety tests (18 loci, 5 populations) were done, significant departures from Hardy Weinberg equilibrium (HWE) were observed after False Discovery Rate (FDR) correction in all population with F2 populations exhibiting HWE deviations at 17 loci (Except T917_9c). F1 and F3 population did not conform to HWE at 13 loci (Except MR4, MR8,

MBR5, T917_9c and VJ2_10a) and 15 loci (Except MR20, T917_9c and VJ2_10a) respectively of 18 loci tested while both Wild Type and Parental populations showed deviations from HWE in 14 out 18 loci.

Population specific Inbreeding Coefficient (F_{IS}), value were different in all population. Based on the figure 4.17 below, Wild type population have the highest positive value (0.1251) which mean that inbreeding occur in high frequency in this population that cause deficit of heterozygosity. Negative F_{IS} values indicate an excess of heterozygosity in the population shown in F1 population with the value of -0.1063. All populations showed evidence for recent reductions in population sizes indicated by the G/W Index values lower than the threshold bottleneck value of 0.68 (Garza and Williamson, 2001) which figure 4.18 showed wild type population has the most severe reduction in population sizes base on their value.

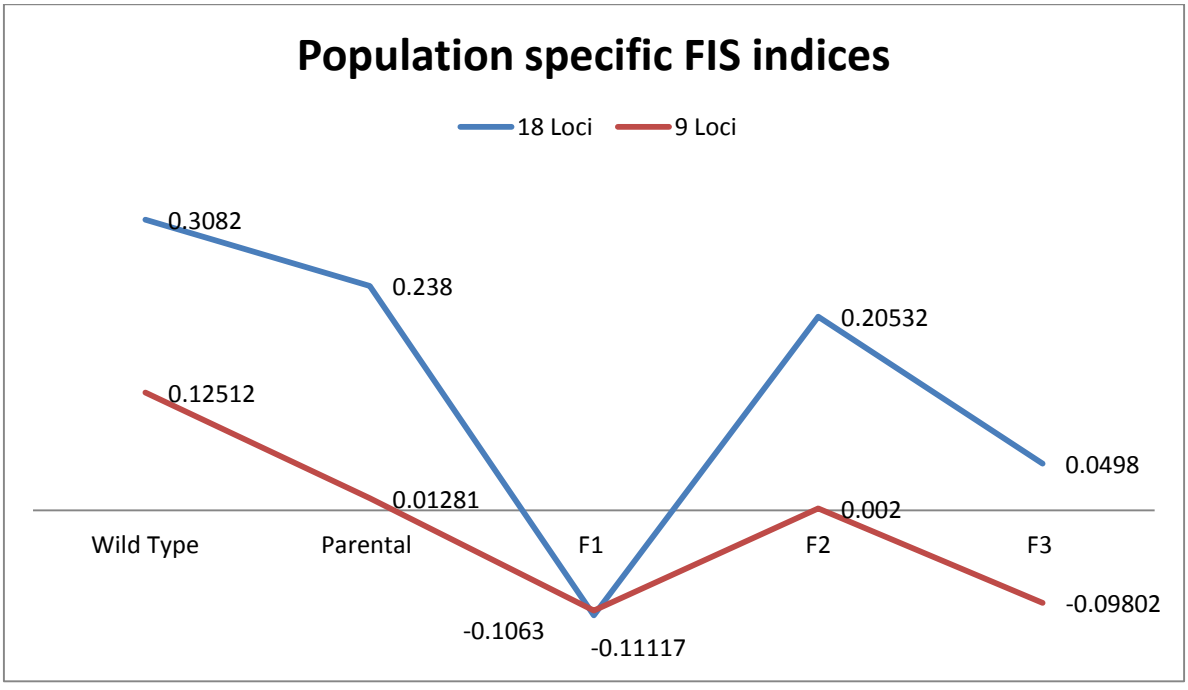


Figure 4.17: Population specific Inbreeding Coefficient (F_{IS}) of all five populations

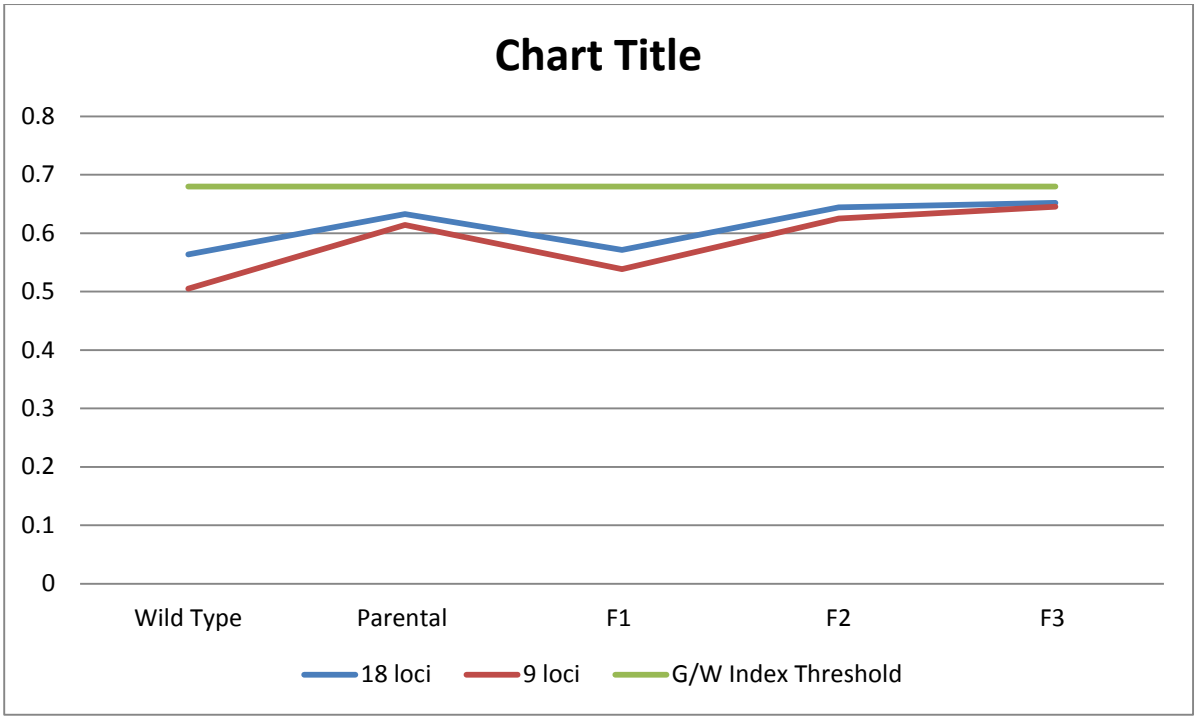


Figure 4.18: Garza and Williamson Index across all loci

Table 4.9: Level of heterozygosity of 5 populations of *Macrobrachium rosenbergii*.

Locus		***Wild type***	***Parental***	***F1***	***F2***	***F3***
MR4	N	35	38	32	73	54
	na	9	13	11	13	13
	ne	3.66	7.39	4.41	5.29	6.52
	H_O	0.7714	0.8421	0.7500	0.8630	0.8704
	H_E	0.7375	0.8761	0.7857	0.8167	0.8544
	P	0.93767	0.2582	0.48706	0.00547*	0.00007*
	Fis	-0.04675	0.03935	0.04615	-0.0571	-0.01881
MR6	N	35	38	32	73	54
	na	10	15	6	17	24
	ne	5.82	11.11	3.42	13.47	13.75
	H_O	0.6000	0.6053	0.9375	0.8493	0.9074
	H_E	0.8402	0.9221	0.7192	0.9322	0.9360
	P	0*	0*	0.01937*	0*	0*
	Fis	0.28884	0.34664	-0.30986	0.08944	0.03079
MR8	N	35	38	32	73	54
	na	9	4	9	14	7
	ne	3.15	1.69	5.90	2.83	2.74
	H_O	0.5429	0.3947	0.9375	0.5753	0.5370
	H_E	0.6923	0.4130	0.8438	0.6508	0.6411
	P	0*	0.0639	0.17972	0*	0.00025*
	Fis	0.21839	0.04475	-0.11311	0.11669	0.16354

MR13	N	35	38	32	73	54
	na	11	9	2	19	10
	ne	8.33	7.08	1.93	11.81	5.19
	H_O	0.0571	0.0526	0.8125	0.0000	0.3333
	H_E	0.8928	0.8702	0.4901	0.9217	0.8148
	P	0*	0*	0.00012*	0*	0*
	Fis	0.93686	0.94027	-0.67568	1	0.59318
MR19	N	35	38	32	73	54
	na	7	7	6	9	4
	ne	4.27	2.28	3.18	1.97	2.1
	H_O	0.0000	0.8947	0.8750	0.2329	0.9259
	H_E	0.7768	0.5695	0.6959	0.4949	0.5298
	P	0*	0*	0*	0*	0*
	Fis	1	-0.58339	-0.26255	0.53112	-0.76021
MR20	N	35	38	32	73	54
	na	4	8	4	7	6
	ne	3.30	6.56	3.57	3.98	3.29
	H_O	0.7143	0.7895	0.9375	0.7260	0.7963
	H_E	0.7068	0.8589	0.7316	0.7542	0.6963
	P	0.55778	0.02652*	0.00016*	0.0006*	0.37682
	Fis	-0.0107	0.08189	-0.2872	0.03758	-0.1344
MR27	N	35	38	32	73	54
	na	4	10	8	8	9

	ne	2.70	5.70	3.56	3.09	4.52
	H_O	0.9143	0.7895	0.9062	0.9041	0.8333
	H_E	0.6381	0.8354	0.7307	0.6812	0.7863
	P	0.00021*	0.00087*	0.02081*	0*	0*
	Fis	-0.44201	0.05572	-0.24515	-0.33016	-0.06047
MR38	N	35	38	32	73	54
	na	9	11	6	14	14
	ne	6.64	5.64	2.38	6.78	5.27
	H_O	0.0857	0.0000	0.8125	0.0274	0.5556
	H_E	0.8617	0.8337	0.5883	0.8583	0.8179
	P	0*	0*	0.00376*	0*	0*
	Fis	0.90183	1	-0.38966	0.96829	0.32283
MR40	N	35	38	32	73	54
	na	5	8	7	8	8
	ne	3.61	5.61	3.26	3.31	3.85
	H_O	0.8286	0.7368	0.7812	0.8356	0.8519
	H_E	0.7337	0.8326	0.7039	0.7026	0.7477
	P	0.00404*	0.00377*	0*	0*	0*
	Fis	-0.13138	0.11642	-0.11191	-0.19089	-0.14085
LM13	N	35	38	32	47	54
	na	12	15	16	16	16
	ne	9.28	10.98	12.19	12.20	10.98
	H_O	0.6857	0.7368	0.7188	0.7447	0.7037

	H_E	0.9052	0.9211	0.9325	0.9279	0.9174
	P	0.00009*	0.01195*	0*	0*	0*
	Fis	0.24514	0.20216	0.23209	0.1992	0.23466
LM08	N	35	38	32	47	54
	na	7	12	7	11	11
	ne	5.75	6.69	5.26	4.36	4.36
	H_O	0.0000	0.3684	0.0938	0.1277	0.1481
	H_E	0.8381	0.8618	0.8229	0.7788	0.7776
	P	0*	0*	0*	0*	0*
	Fis	1	0.57576	0.88768	0.83755	0.81093
MBR5	N	35	38	32	47	54
	na	7	10	9	10	10
	ne	4.27	7.29	6.42	6.17	6.67
	H_O	0.9143	0.8947	0.9375	0.8936	0.8889
	H_E	0.7768	0.8744	0.8576	0.8469	0.8579
	P	0*	0*	0.03905**	0*	0*
	Fis	-0.18004	-0.0236	-0.09476	-0.05574	-0.03646
LM16	N	35	38	32	47	54
	na	5	5	6	6	6
	ne	2.61	4.28	3.39	3.75	3.42
	H_O	0.0000	0.3158	0.1250	0.1489	0.0926
	H_E	0.6261	0.7768	0.7158	0.7412	0.7146
	P	0*	0*	0*	0*	0*

	Fis	1	0.59673	0.82766	0.8008	0.87148
LM04	N	35	38	32	47	54
	na	7	10	9	10	10
	ne	5.14	7.35	5.62	6.71	7.30
	H_O	0.5714	0.6316	0.5938	0.6596	0.7778
	H_E	0.8170	0.8754	0.8353	0.8602	0.8711
	P	0*	0.00115*	0*	0*	0.00485*
	Fis	0.30364	0.28126	0.29249	0.23518	0.10799
T917_7e	N	35	38	32	68	54
	na	8	12	8	10	13
	ne	7.21	9.86	5.77	6.25	8.68
	H_O	0.5429	0.6842	0.6250	0.7353	0.7593
	H_E	0.8737	0.9105	0.8398	0.8462	0.8930
	P	0*	0*	0*	0*	0*
	Fis	0.38211	0.25107	0.25882	0.1319	0.15101
T917_7d	N	35	38	32	66	54
	na	8	10	8	11	11
	ne	4.79	6.75	4.75	5.46	6.80
	H_O	0.6286	0.6842	0.6562	0.6667	0.7037
	H_E	0.8025	0.8632	0.8021	0.8229	0.8609
	P	0.00043*	0.00086*	0.00223*	0*	0*
	Fis	0.21921	0.20953	0.18421	0.19106	0.18395

T917_9c	N	35	38	32	71	54
	na	4	6	5	6	6
	ne	3.44	3.93	3.29	2.97	3.38
	H_O	0.6286	0.7368	0.7500	0.6620	0.7222
	H_E	0.7193	0.7554	0.7068	0.6679	0.7103
	P	0.04866**	0.21624	0.44294	0.12992	0.05089
	Fis	0.1277	0.02494	-0.0621	0.00889	-0.01697
VJ2_10a	N	35	38	32	71	54
	na	5	7	6	6	7
	ne	4.68	6.43	4.71	5.28	5.78
	H_O	0.6857	0.7368	0.6875	0.7324	0.8333
	H_E	0.7979	0.8558	0.8001	0.8163	0.8347
	P	0.09817	0.06805	0.21229	0.00243*	0.20527
	Fis	0.14241	0.14061	0.14268	0.10345	0.00167

Note: N -Number of samples, na – Number of alleles, ne – Number of effective alleles, H_O – Observed Heterozygosity, H_E – Expected Heterozygosity, P – Probability value estimates regarding deviation from Hardy Weinberg equilibrium, F_{IS} – Inbreeding Coefficient.
*Significant p-values ($p<0.05$) after FDR correction.
** Not significant p-values ($p<0.05$) after FDR correction.

4.2.4.3 Genetic diversity among populations

An analysis of molecular variance (AMOVA) was done using Arlequinn software (Schneider *et al.,* 2000) for measuring the variance of group, populations and individuals levels by applying the estimator of Weir and Cockerham (1984). AMOVA within and among the five populations were done twice. The first AMOVA was done by categorized 4 populations into two groups, wild type population as the parent group and offspring group while the second AMOVA was done by replacing the wild type population with the parental population of Kampung Acheh Sungai Perak as the parent group. The table 4.10 showed that Wild Type Population as the parent group has minimal differentiation with the value of 1.4 % variation. Similarly the 4.11 showed minimal differentiation between parent group and offspring group (0.03%). However, the variation within individuals accounted for 97.31% and 98.43% respectively.

Table 4.10: Analyses of molecular variance (AMOVA) of parent group (Wild Type population) and offspring (F1, F2, and F3 populations) group of *Macrobrachium rosenbergii* on 9 microsatellite loci

Source of Variation	***d.f.***	***Sum of squares***	***Variance components***	***Percentage of variation***	***Fixation indices***	***P-value***
Among groups	1	15.71	0.0422	1.40	0.0140	0.2463
Among populations within groups	2	26.35	0.1030	3.43	0.0348	0.000
Among individuals within populations	190	531.36	-0.06431	-2.14	-0.02248	0.9159
Within individuals	194	567.50	2.9256	97.31	0.0269	0.2874
Total	387	1140.92	3.0061			

Table 4.11: Analyses of molecular variance (AMOVA) of parent group (Parental population) and offspring (F1, F2, and F3 populations) group of *Macrobrachium rosenbergii* on 9 microsatellite loci

Source of Variation	***d.f.***	***Sum of squares***	***Variance components***	***Percentage of variation***	***Fixation indices***	***P-value***
Among groups	1	11.55	0.0010	0.03	0.0003	0.4853
Among populations within groups	2	26.35	0.1030	3.47	0.0347	0.0000
Among individuals within populations	193	532.56	-0.1203	-4.03	-0.0418	0.9902
Within individuals	197	591.00	3.0000	98.43	-0.0053	0.8446
Total	393	1161.46	2.9841			

Table 4.12 showed the population pairwise F_{ST} which estimated distance, according to the pairwise difference method that demonstrated significant differentiation that ranged from 0.02133 to 0.06844.

Table 4.12: Pairwise population distance represented by F_{ST} value calculated using 9 loci with "*" indicate significant F_{ST} ($p<0.05$).

Site	(1)	(2)	(3)	(4)	(5)
(1)	0.00000				
(2)	0.02133*	0.00000			
(3)	0.04169*	0.05136*	0.00000		
(4)	0.04674*	0.05030*	0.04485*	0.00000	
(5)	0.02238*	0.02651*	0.06844*	0.06223*	0.00000

*Note:*Site (1) = F3 population, (2) = F2 population, (3) = F1 population, (4) = Wild Type population, (5) = Parental population

4.2.4.4 Cluster Analysis and relationship trees

A schematic representation of the genetic structure of all the individuals under study is represented in figure 4.20 & 4.22. The STRUCTURE software uses allele frequencies from what is assumed to be potentially unlinked loci, and uses these figures to identify which set of genes are belong of which population. Based on the alleles frequencies the population is set (K), but K doesn't represented the number of population sampled but it can be anywhere from 1 to number of sampling locations. Thus, this software will theoretically indicate the total of populations that represent the data inserted. Thereafter the individuals too are assigned to their significant K populations (Pritchard, Stephens and Peter, 2000).

The log probabilities [Ln P(D)] associated with different numbers of genetic clusters (K), calculated from Bayesian clustering analysis of 232 individuals of 18 loci showed the highest value at K = 5[Ln P(D) = -15198.3], and the lowest value at K = 1[Ln P(D) = -16119.8] as shown in figure 4.19. According to degree of admixture (alpha value), K =5 gives the nearest value to zero indicates that the most appropriate number of genetic groups assigned for the given data set was five(Figure 4.20). Each population clustering separately and there was hybridization occur in each population.

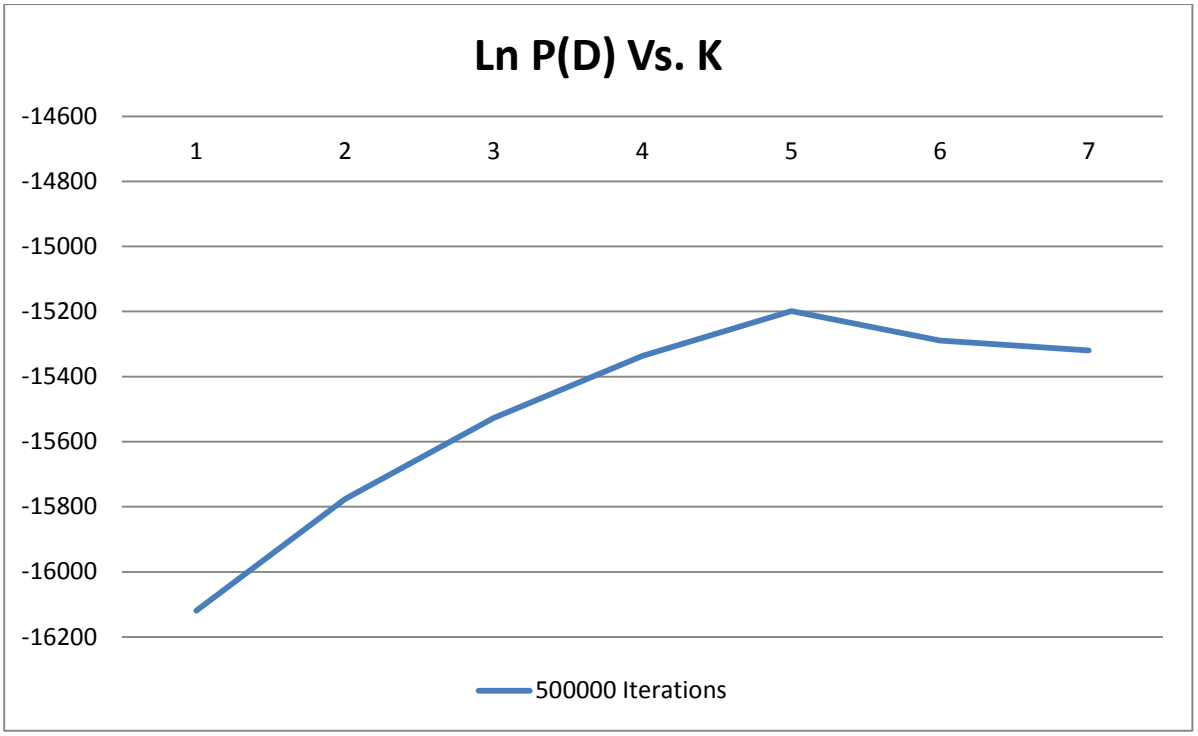


Figure 4.19: Results of Bayesian Cluster Analysis of log-likelihood ratios [Ln P(D)] for different K values. The peak of Ln P(D) identifies the best number of clusters in this study five.

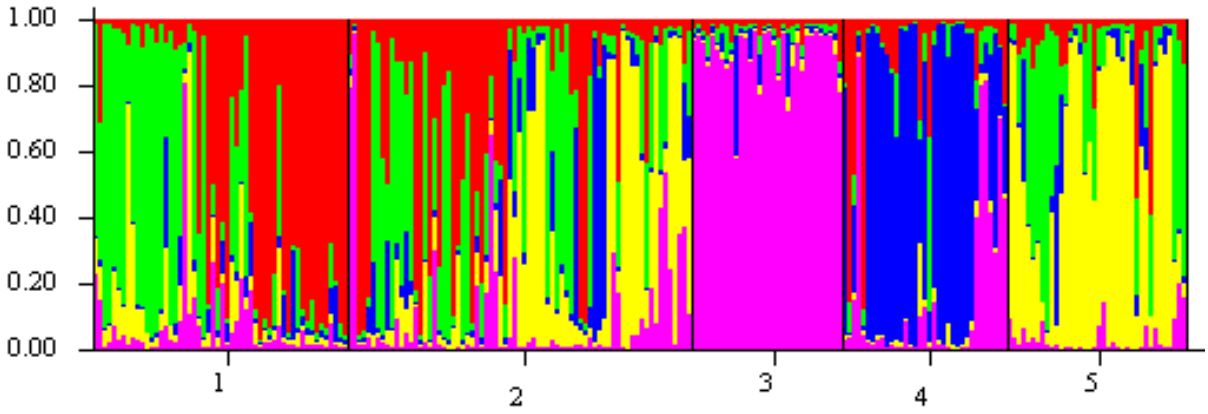


Figure 4.20: Results of the Cluster Analysis of generated by the STRUCTURE software. Different colour indicates different genetic clusters. Each column represents an individual whereby the height of the column segments shows the probability of assignment of the individual *M.rosenbergii* to the genetic clusters

After removing null alleles, the log probabilities [Ln P(D)] showed the highest value at K = 3[Ln P(D) = -6941.2], and the lowest value at K = 6[Ln P(D) = -7441.5] as shown in figure 4.21 . According to degree of admixture (alpha value), K =3 gives the nearest value to zero indicates that the most appropriate number of genetic groups assigned for the given data set was three (Figure 4.22). Clustering 1 showed clustering of F3, F2 and parental populations which an evidence of gene flow of migration occur between these population, The hybridizations that are seen in F2 and F3 populations is caused from Parental and Wild Type populations, however both of these populations most likely shared alleles from Parental rather than Wild Type population which explains the clustering of these three population. This has been supported with data from genetic identity and genetic distance. In the other hand, even wild type and F1 population cluster separately, they showed quite similar genetic makeup with each other than the other cluster. The minimal number of colour lined in F1 population showed a very high inbreeding occurred in this population. This has been supported by the data showed in $F_{IS.}$

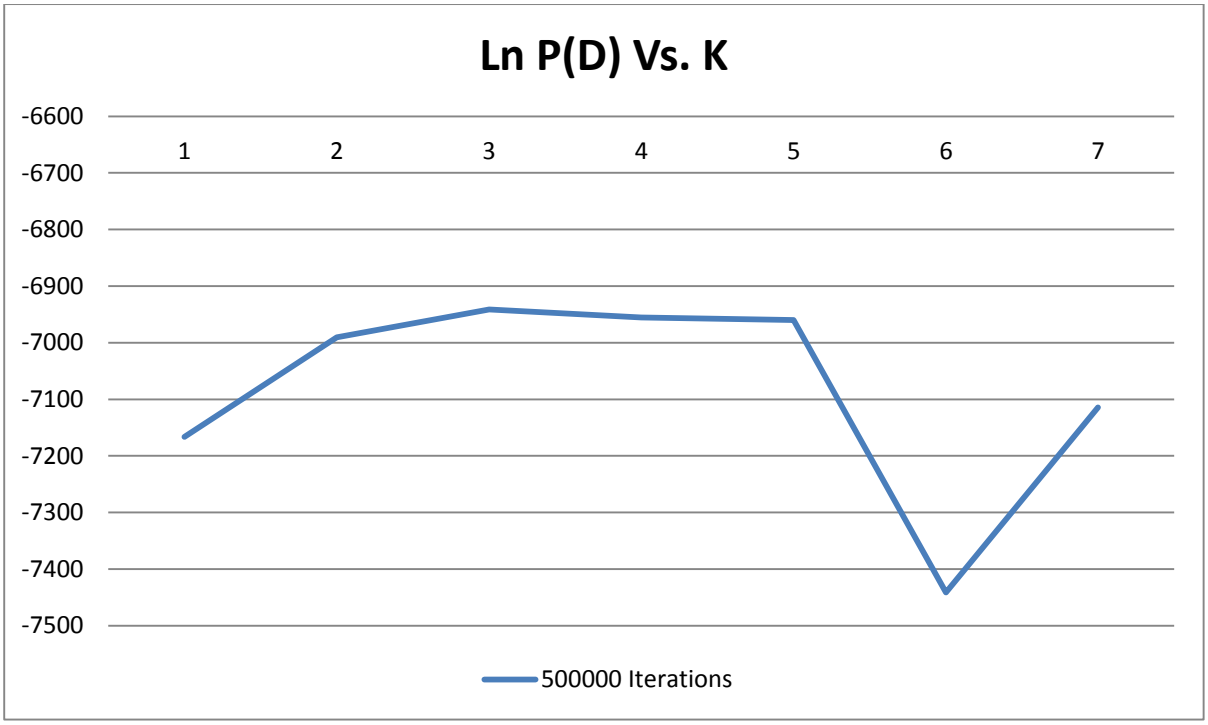


Figure 4.21: Results of Bayesian Cluster Analysis of log-likelihood ratios [Ln P(D)] for different K values. The peak of Ln P(D) identifies the best number of clusters in this study three.

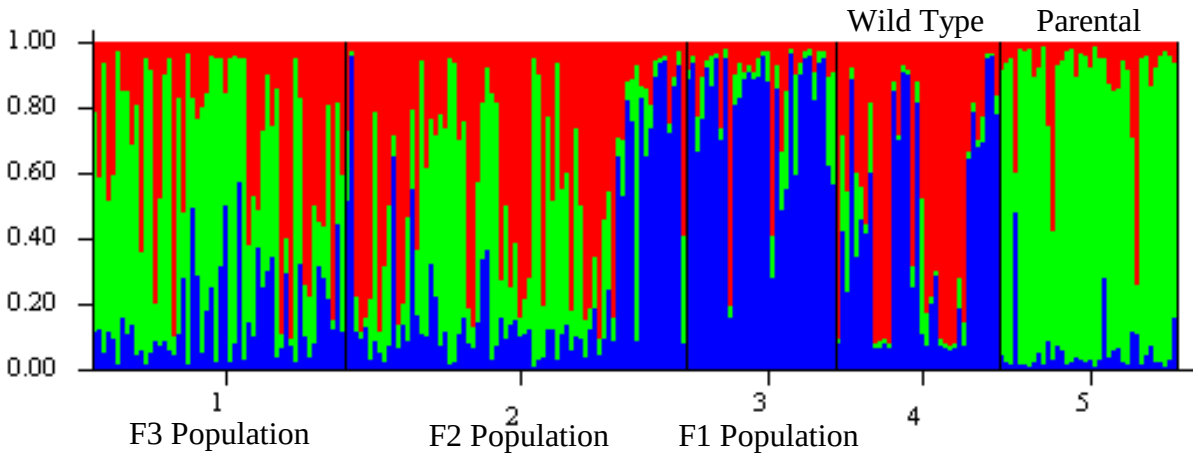


Figure 4 .22: Results of the Cluster Analysis of generated by the STRUCTURE software. Different colour indicates different genetic clusters. Each column represents an individual whereby the height of the column segments shows the probability of assignment of the individual *M.rosenbergii* to the genetic clusters

The genetic identity between F3 and F2 populations was the highest 0.9283 and the genetics distance between these two populations was the lowest 0.0744 (Table 4.13). This suggested both of the populations were the most related populations.

Clustering analysis of UPGMA (unweighted pair-group method using the arithmetic average) consensus tree of the populations was constructed based on Nei's (1978) unbiased genetic distance using the Phylip (version 3.67) with bootstrap value 1000 (Figure 4.23). UPGMA is clustering method that uses average pairwise distance among all individuals of the sample. The consensus tree divided these five populations 2 major clusters. F3, F2, Wild type and Parental Populations belonged to one group while F1 populations clustered by itself. The first major cluster subdivided into three groups. F3 and F2 populations belonged to a single group and the other two populations belonged to two different group.

Table 4.13: The values of genetic identity (above diagonal) and genetic distance (below diagonal between five populations of *Macrobrachium rosenbergii*

Pop ID	***F3***	***F2***	***F1***	***Wild Type***	***Parental***
F3	*****	0.9283	0.8273	0.8566	0.8667
F2	0.0744	*****	0.7846	0.8718	0.8977
F1	0.1896	0.2425	*****	0.7968	0.7476
Wild Type	0.1548	0.1371	0.2271	*****	0.7919
Parental	0.1430	0.1080	0.2909	0.2334	*****

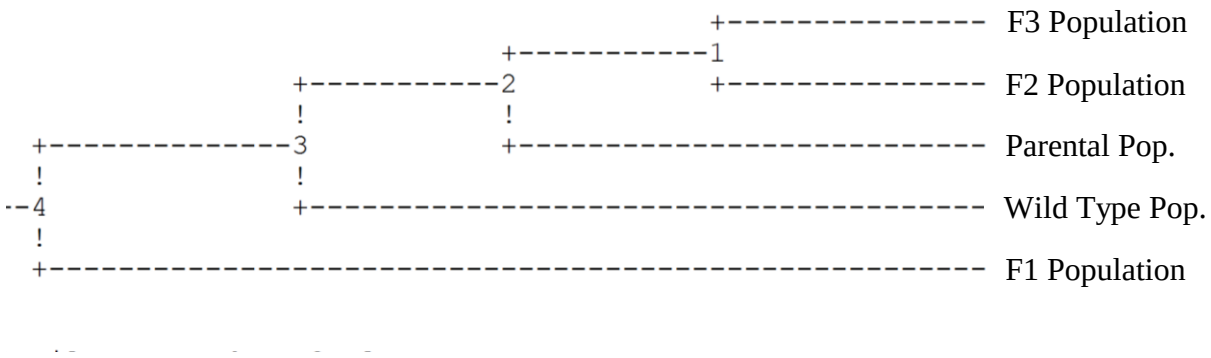


Figure 4.23: Consensus tree generated using 18 microsatellite loci

When 9 null alleles were removed, Population F3 and F2 populations gives the highest genetic identity which is 0.9416 and the lowest genetic distance is 0.0602 (Table 4.14). The F2 and F3 population seems to be the most related population from the result above. The consensus tree based on UPGMA, grouped the five populations into 2 major clusters (Figure 4.24). F3, F2, and Parental Populations belonged to one group while F1 and Wild Type populations clustered together as one group. The first major cluster subdivided into two groups. F3 and F2 populations belonged to a single group and the Parental population belonged to different group.

Table 4.14: The values of genetic identity (above diagonal) and genetic distance (below diagonal between five populations of *Macrobrachium rosenbergii*

Pop ID	***F3***	***F2***	***F1***	***Wild Type***	***Parental***
F3	*****	0.9416	0.8651	0.8560	0.9223
F2	0.0602	*****	0.8545	0.8666	0.9148
F1	0.1449	0.1572	*****	0.8492	0.7688
Wild Type	0.1554	0.1431	0.1634	*****	0.7847
Parental	0.0809	0.0891	0.2630	0.2424	*****

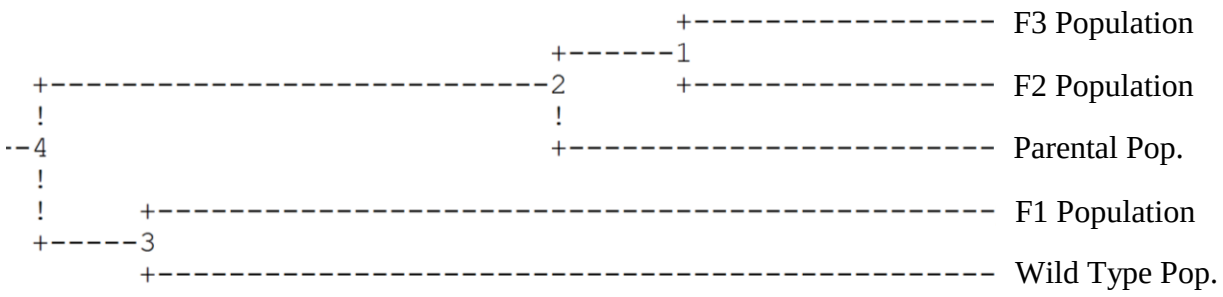


Figure 4.24: Consensus tree generated using 9 microsatellite loci

4.3 Mitochondria Marker Amplification

All samples included in genetic analysis shows 100% successful amplification by universal MITO MRF and MITO MRR primer set (0. Folmer *et al.*, 1994) with optimum annealing temperature of 50.6°C and optimum amount of 25mM $MgCl_2$of 1.2µl per reaction volume (Figure 4.25 and Figure 4.26). The expected amplified products (710 bp) were excised and purified from any chemical such as PCR mixture or EtBr that might be stuck on the template. Next, the purified products were checked on 1.0% agarose gel before processed it in Genetic Analyzer (Figure 4.27). The success of amplified COI sequence was shown in the form of Chromatogram diagram that produces sequence for each sample**.** The quality of each peak in Chromatogram for each sample was checked to eliminate ambiguous base especially the one that produces low quality peaks with the Chromas v2.33 software (Figure 4.28). There is no sign of double bands on the agarose gel both on the amplification product and purified product. Supported by chromatogram, these has proved all the sequence were not NuMts.

Successful amplified COI sequences were verified using NCBI BLAST program to confirm the amplified products were the COI full sequence of *M. rosenbergii*. All confirmed regions were aligned using ClustaW (Thompson et al, 1997) with default parameters and trimmed at both ends to exclude any unsure sequences using MEGA 4 software and resulted in a total length of 624 base pair amplicon for subsequent analysis. There is no sign of unusual amino acid substitutions and stop codons in the sequences found. 171 samples have been amplified with the mitochondrial marker and it comprises of 3 months of F1 populations, parental population and also wild population.

4.3.1 Mitochondrial Polymorphism

Among the 171 individuals in which the 624-bp sequences of Cytochrome Oxidase (COI) gene were analyzed, 591 sites (94.7%) were conserved with 33 sites (5.3%) were variables. The 33 variables sites consist of 13 Singleton variables sites and 20 Parsimony informative sites. The list of haplotypes assigned is shown in Table 4.15.

There are a total of 27 haplotypes identified from all the populations investigated and with the haplotype diversity value of h = 0.5719 being recorded (Table 4.16). Among the 27 generated haplotypes (hap), hap1 was the most common haplotype and was shared among all five populations with frequencies of these shared haplotypes were 46.7%, 45.5%, 73.7%, 85.7% and 68.1% in the wild type, brood stock, F1, F2 and F3 populations respectively. The genetic homogeneity is highest in F2 population than in others (haplotype diversity = 0.267). This result is due to the high number of haplotype 1 in this population. In Parental population, with the most heterogeneous population (haplotype diversity = 0.771), the most frequent haplotype is also 1, although there are other haplotypes in polymorphic frequencies as well. Twenty out of a 27 haplotypes were unique to the populations from one of the five populations. Specifically, six haplotypes (hap3,hap4,hap7,hap8,hap9 and hap10), uniquely representing wild type population, 7 haplotypes (hap14,hap 15,hap16,hap17,hap18,hap19, and hap20) were from the parental, each 2 haplotypes were from the F1 (hap21 and hap22) and F2 (hap23 and hap24) population and 3 haplotypes (hap25,hap26, and hap27) were from F3 population.Hap2 is shared between Wild type, Parental and F3 population while there also haplotype sharing between wild type and F1 and F2 population which is hap5. But one thing that interesting there no haplotype of parental population appears in F1 population but only appears in F2 population (Hap12) and F3 population (Hap13).

Nucleotide diversities were similar for the five populations. The nucleotide frequencies are 0.266 (A), 0.266 (T), 0.262 (C), and 0.205 (G). The transition/transversion rate ratios are k_1 = 59.111 (purines) and k_2 = 21.919 (pyrimidines). The overall transition/transversion bias is R = 17.868, where R = [A*G*k_1 + T*C*k_2]/ [(A+G)*(T+C)]. The majority of the nucleotide substitutions in the COI gene were transition substitutions (Table 4). There were 56transitions substitutions out of the 27 haplotypes while only 8 transversion substitutions were found in all the haplotypes.

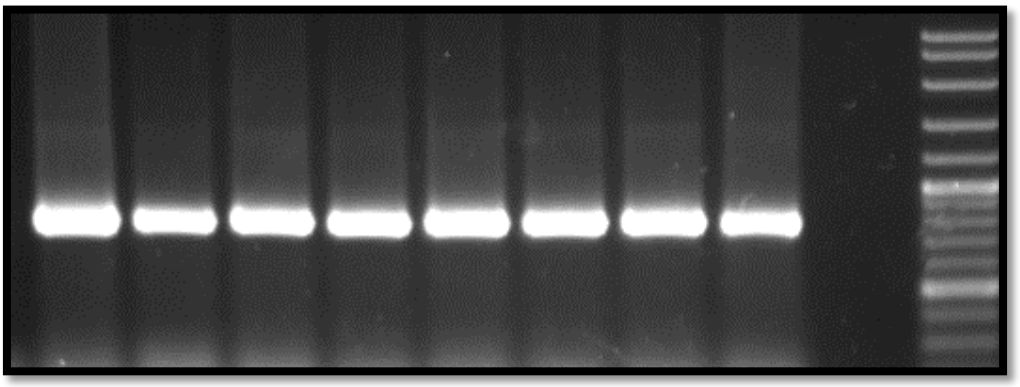


Figure 4.25: Example of COI Region amplification of Wild Type Population samples with 100 bp DNA Ladder (Lonza, USA).

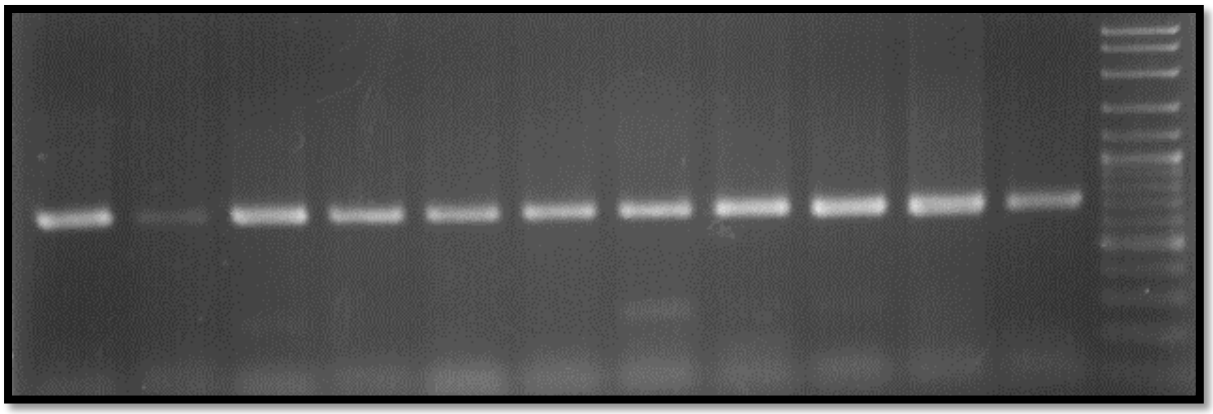


Figure 4.26: Example of COI Region amplification of Parental Population samples with 100 bp DNA Ladder (Lonza, USA).

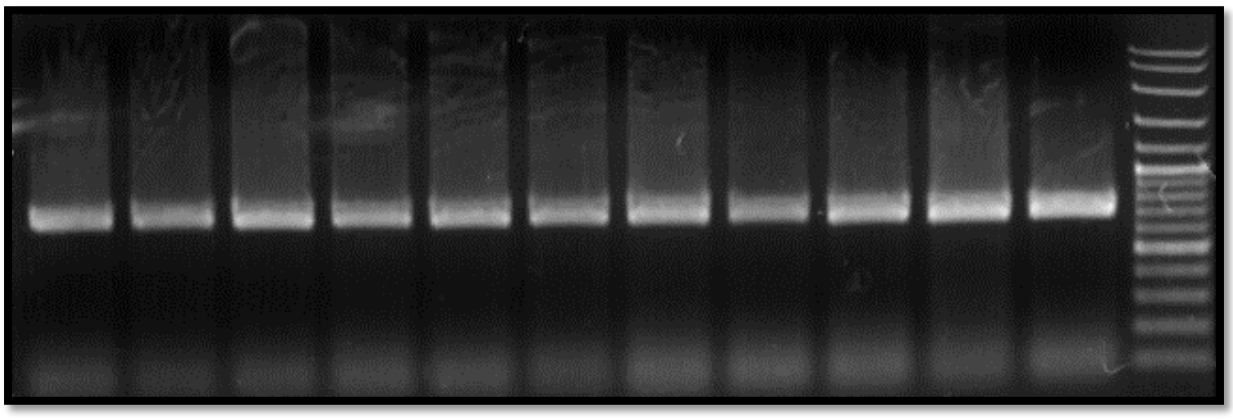


Figure 4.27: Examples of purified product of successful COI amplification of Parental Population samples with 100 bp DNA Ladder (Lonza, USA).

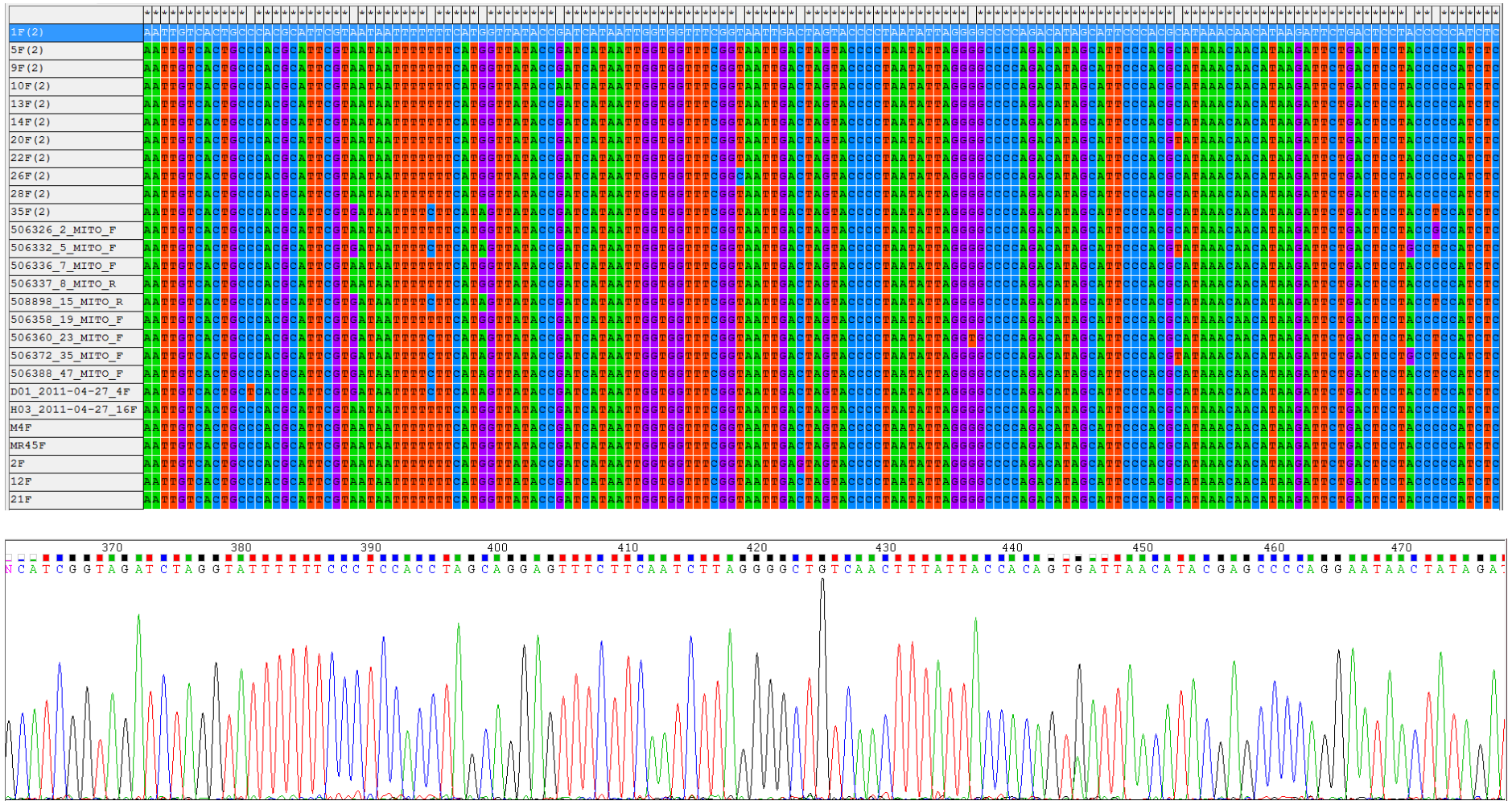


Figure 4.28: (Above) Results of a CLUSTALW alignment of 27 haplotype performed with MEGA 4.0 software, (Below) Examples of Chromatogram of amplified COI sequence using Chromas v2.33 software**.**

Table 4.15: Frequencies and haplotypic diversities of the different haplotypes in all populations

Haplotype	***Wild Type***		***Parental***		***F1***		***F2***		***F3***	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
1	14	46.7	15	45.5	14	73.7	36	87.8	32	68.1
2	6	20.0	1	3.0					4	8.5
3	1	3.3								
4	1	3.3								
5	2	6.7			1	5.7	1	2.4		
6	1	3.3					1	2.4	4	8.5
7	1	3.3								
8	1	3.3								
9	1	3.3								
10	1	3.3								
11	1	3.3	5	15.2						
12			1	3.0			1	2.4		
13			4	12.1					3	6.4
14			1	3.0						
15			1	3.0						
16			1	3.0						
17			1	3.0						
18			1	3.0						
19			1	3.0						
20			1	3.0						
21					1	5.7				
22					3	15.8				
23							2	4.9		
24							1	2.4		
25									2	4.3
26									1	2.1
27									1	2.1
Total	30	100	33	100	19	100	41	100	47	100
h	0.754		0.771		0.450		0.267		0.526	

N = number of individual, h = haplotypic diversity

Table 4.16: List of 27 unique haplotypes generated using software DnaSP v5 (Librado & Rozas, 2009).

HP #	**003**	**006**	**033**	**056**	**069**	**087**	**096**	**120**	**132**	**141**	**147**	**156**	**177**	**184**	**204**	**228**	**255**	**258**	**291**	**321**	**324**	**339**	**342**	**366**	**368**	**420**	**444**	**483**	**517**	**534**	**564**	**606**	**618**
1	C	C	G	C	G	A	A	C	A	T	G	G	T	C	G	C	A	C	A	T	T	G	C	G	T	G	G	G	A	A	A	A	G
2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	G	.
3	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.
4	.	.	.	.	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A
5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
6	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.
9	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	G	.
10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.
11	.	.	.	.	.	.	.	.	G	C	A	.	.	.	.	.	.	T	.	.	.	A	T	A	.	.	A	.	.	.	G	.	.
12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13	.	.	.	.	.	.	.	.	G	C	A	.	.	.	.	T	G	T	.	.	.	A	T	A	.	.	A	.	.	.	G	.	A
14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
16	.	.	.	.	.	.	.	.	G	C	A	.	.	.	.	.	.	T	.	.	.	A	T	A	.	.	A	A	.	.	G	.	.
17	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	.	.	.	.	G	.	.
18	.	.	.	.	.	.	.	.	G	C	A	.	.	.	**T**	.	.	T	.	.	.	A	T	A	.	.	A	A	.	.	G	.	.
19	.	.	.	.	.	.	.	.	G	C	A	.	.	.	.	T	G	T	.	.	.	A	T	A	.	A	A	.	.	.	G	.	A
20	.	.	.	.	.	.	G	.	G	C	A	.	.	.	.	.	.	T	.	.	.	A	T	A	.	.	A	.	.	.	G	.	.
21	.	.	.	.	.	.	.	T	G	C	A	.	.	.	.	.	.	.	.	C	C	A	T	A	**G**	.	A	.	**C**	.	G	.	.
22	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
23	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
25	.	.	.	.	.	.	.	.	.	.	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
26	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
27	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	G	.

4.3.2 Molecular Diversity

The molecular diversity indices for Theta S and Theta pi graphically illustrate the diversity of populations (Figure 4.29). These indices are unbiased estimators of population genetic structure, where Theta (θ), statistically summarizes the distribution of variation within and among populations when samples are assumed to represent characteristics of the larger group from which they are sampled, as if a larger number of samples were actually drawn from that population, random-effect sampling (Weir and Cockerham 1984; Weir and Hill 2002). Theta S defines the infinite site equilibrium relationship between polymorphic sites, sample size and θ, for a sample of non-recombining DNA (Tajima, 1989). Theta pi describes the infinite site equilibrium relationship between the mean number of pairwise differences and θ (Tajima, 1983).Note that these indices show moderate levels of diversity (Theta S = 3.849 – 4.544, Theta pi = 1.891- 6.004) for all populations studied except F2 population Theta S = 1.162, Theta pi = 0.284).

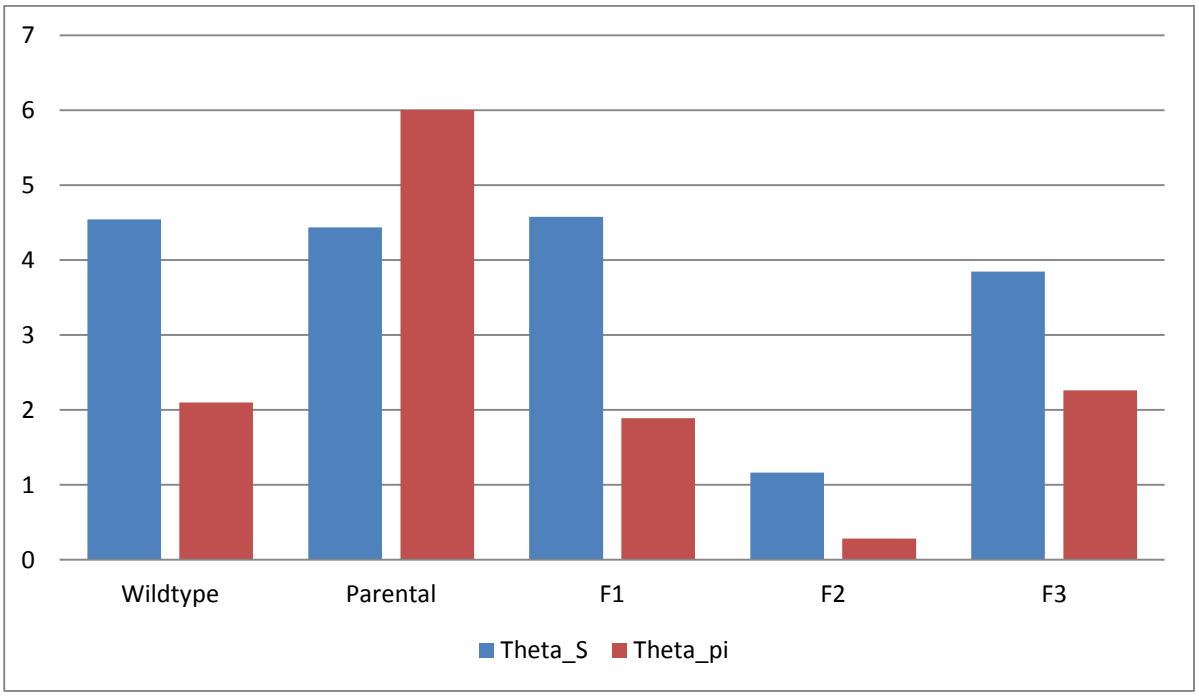


Figure 4.29: Molecular diversity indices, Theta S and Theta pi for all populations.

Tajima's test (Tajima, 1989) is based on the fact that under the neutral model estimates of the number of segregating/polymorphic sites and of the average number of nucleotide differences are correlated. If the value of D is too large (positive or negative), the neutral 'null' hypothesis is rejected. Tajima's D equals zero for neutral variation, is positive when an excess of rare polymorphism indicates positive selection and is negative in the excess of high-frequency variants, indicating balancing selection (Oleksyk, 2010).

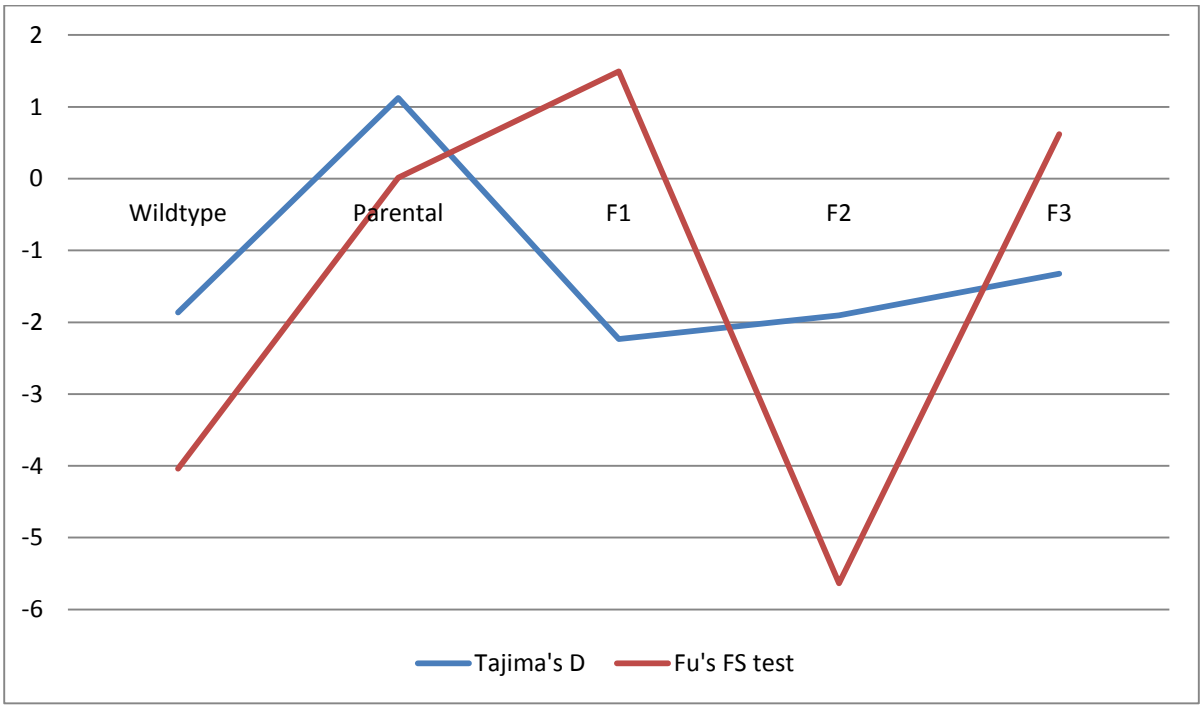


Figure 4.30: Neutrality Test for all populations.

Tajima’s test of departures from the neutral expectations detects a significant deviation from neutrality (Figure 4.30) from the wild type, F1, F2 and F3 populations (p-value <0.05). While in Fu and Li’s D* and F* statistics detected a significant deviation from neutrality for Parental, F2 and F3 population suggesting the existence of excessive rare nucleotide polymorphisms with respect to predictions of the neutral theory, with the possible effects of purifying, background selection, and/or population expansion. For the other two population the

wild type and F1 the D, D* and F* statistics were not significant, suggesting that the nucleotide substitutions of the COI gene are consistent with the neutral evolution theory.A positive value of Fu Test and Tajima is evidence for a deficiency of alleles, as would be expect from a recent population bottleneck or from overdominant selection that occurred in parent population. A positive value also can be seen in F1 and F3 population as in time, a lot of juveniles were been restock to this river.

The population's structure of *M.rosenbergii* were explained by the obtained mtDNA COI F_{st} values with the $p < 0.05$ as the cutoff point of significant population differentiation. Table 4.17 demonstrated significant differentiation that ranged from 0.0020 to 0.3381. The highest divergence value was between F2 population and Parental population and the lowest was between Wild type and F3 population.

4.17: Pairwise population distance represented by F_{st} value calculated using mtDNA COI sequence with bold and "*" indicate significant F_{st} ($p<0.05$).

Site	(1)	(2)	(3)	(4)	(5)
(1)	0.00000				
(2)	0.0538*	0.00000			
(3)	0.0105*	0.0138*	0.00000		
(4)	0.0020*	0.0907*	0.0345*	0.00000	
(5)	0.1977*	0.3381*	0.2352*	0.2221*	0.00000

Note: Site (1) = F3 population, (2) = F2 population, (3) = F1 population, (4) = Wild Type population, (5) = Parental population

4.3.3 The Parsimony network

The parsimony network (Figure 4.31) illustrated the relationship between the 27 *Macrobrachium rosenbergii* haplotypes. The colour differences show the different population while the size of the circle represented the frequency of haplotypes. The centre of the network is the most common haplotype which is hap1 that been shared between all populations. This network has been produce by using TSC 1.21 software and Microsoft Office PowerPoint 2007.

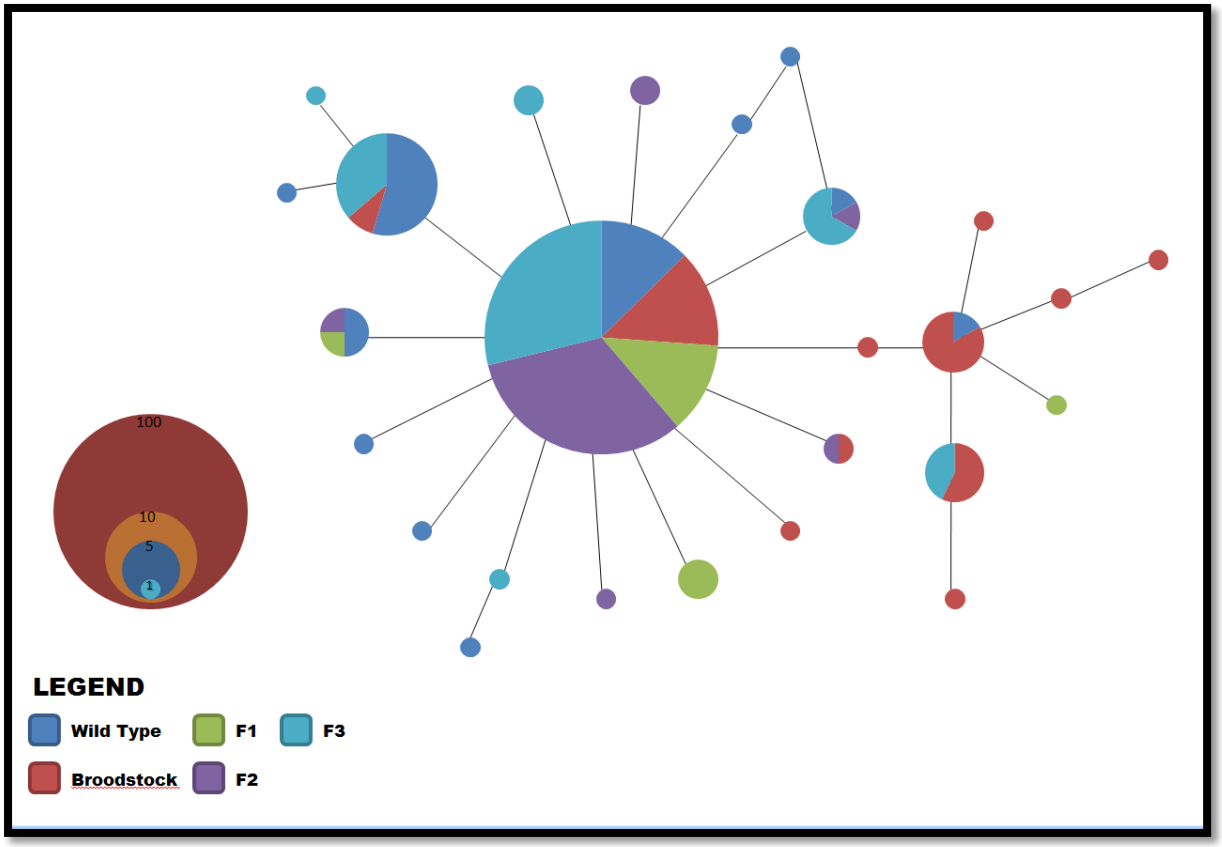


Figure 4.31: Haplotype network that connecting all generated haplotypes; the size of each haplotype is proportional to the number of each individual of corresponding haplotype; each line represents one mutational step.

4.4 Ecological Study

4.4.1 Benthic Study

Identification was made at the level of class and order, but in some cases the organisms were identified up to genus or species, using the Taxonomic Key (Hartman, O., 1941, 1944, 1945, 1947) (Robert L. et al., 2003). Several species have found abundantly on the bottom of Sungai Timun. Some of the examples with its picture were showed below:

Figure 4.32: Benthos Organisms; A) *Micrura lerdyi* (Aquatic worms), B) *Nereis sp.*, C) Scud (Order: Amphipoda) , D) Sowbug (Order:Isopoda), E) *Jassa sp.* F) *Marisa sp.*

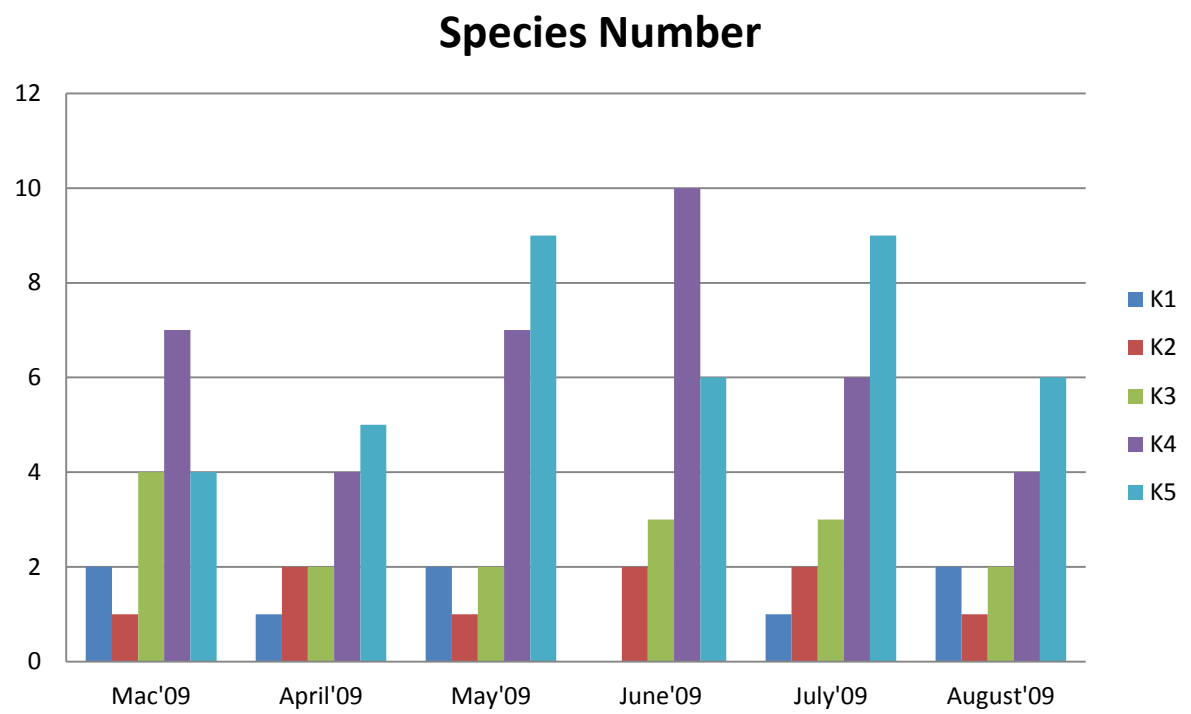


Figure 4.33: Number of Benthos species successfully captured using Eckman Grab in all stations

Benthic study of the Sungai Timun suggested that a maximum of 10 species were founded on the surface of river bed of Sungai Timun. This is the result of using Eckman Grab in all the stations with 3 substations in each station. Substation A and C were at the river bench while Substation B is at the centre of the river. Roughly, from data that been collected, higher number of benthos organisms were founded at Substations A and C than substation B. The figure above also showed that station 1, 2, and 3 have much more lower species number than station 4 and station 5 through all the sampling months.

4.4.2 Water Quality Analysis

Water quality data were used to determine the water quality status weather in clean, slightly polluted or polluted category and to classify the rivers in Class I, II, III, IV or V based on Water Quality Index (WQI) and Interim National Water Quality Standards for Malaysia (INWQS) every year. The water quality statistical analysis enables to identify the point source (PS) and non-point source (NPS) pollutions impairing the river.

Dissolved Oxygen (DO) is essential for aquatic life. A low DO (less than 2mg/l) would indicate poor water quality and thus would have difficulty in sustaining many sensitive aquatic lives. Figure 3.34 showed that DO for all the stations were ranging between 2.05 mg/l to 8.28 mg/l. All five stations showed higher value in November 2009 while much more lower value in August 2009. Based on Interim National Water Quality Standards for Malaysia (INWQS), this DO value classified the river in Class III.

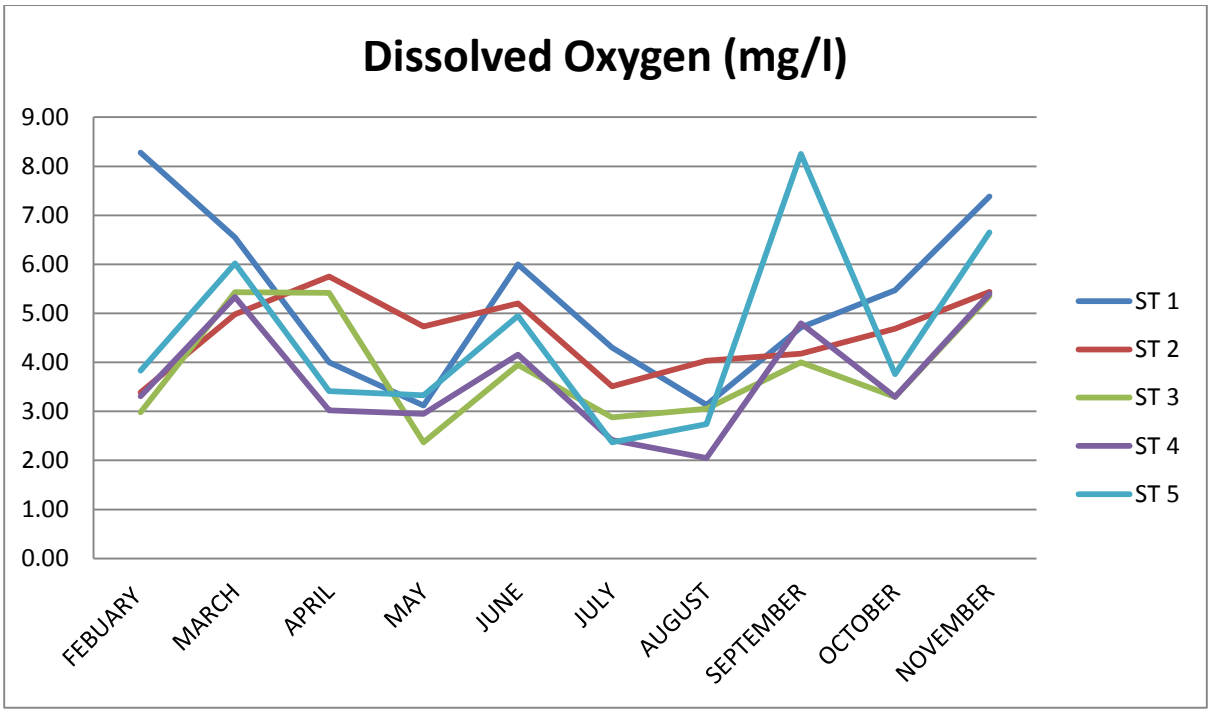


Figure 4.34: Dissolved Oxygen (DO) concentration for all five stations along the river from February 2009 until November 2009

The second parameter is pH which an indicator of the existence of biological life as most of them thrive in a quite narrow and critical pH range. In this case *Macrobrachium rosenbergii*, is best to avoid a pH below 6.0 or above 9.5. Figure 4.35 below showed fluctuation of pH between each month. March, June, and November 2009 showed pH value less than 6.0 while no stations in all the months showed pH value more than 9.5. Higher pH value can be correlate with the present of the algae on the water surface.

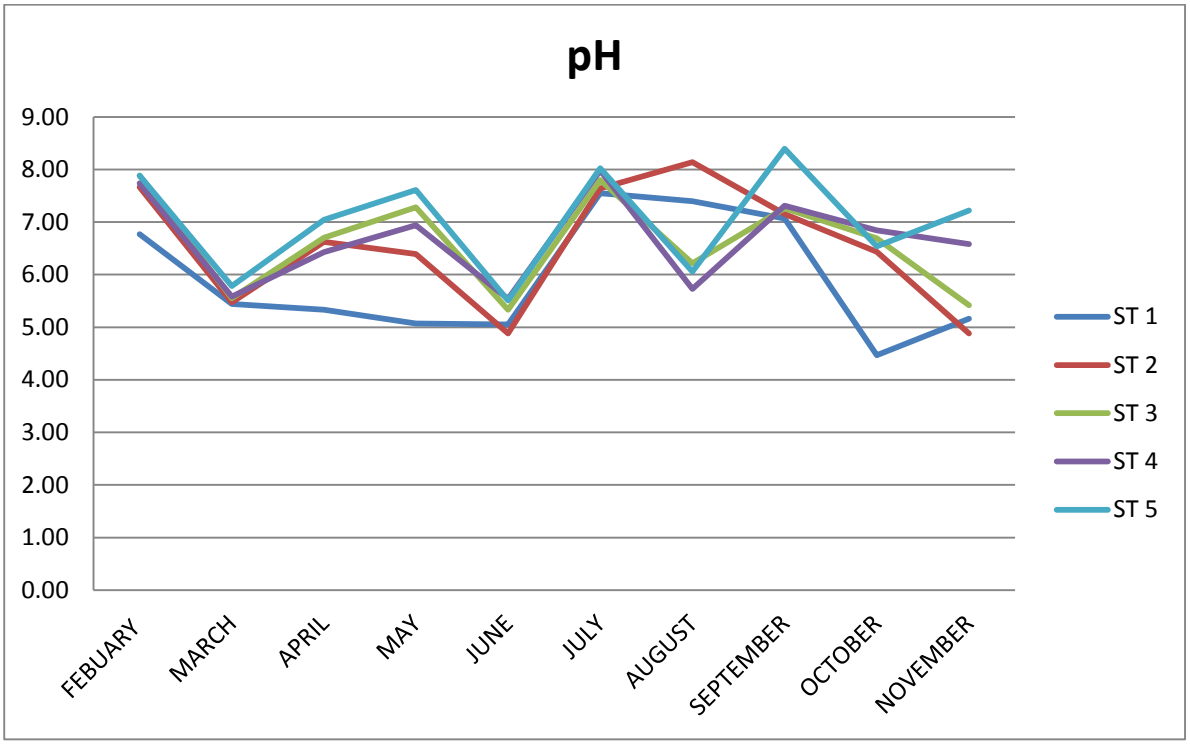


Figure 4.35: pH value for all five stations along the river from February 2009 until November 2009

Macrobrachium rosenbergii is one of the species that tolerate with different value of salinity based on their life stages. However, based on FAO report on 2002, the best salinity is between 10 ppt to 16 ppt where 12 ppt is usually the prawn hatched their eggs and the hatched larvae will migrate to brackish water with a salinity of 10 to 14 ppt. Figure 4.36 showed that the suitable salinity for the studied species can be found at station 3, 4 and 5 which this stations actually nearer to the sea that let the saline water mixed with freshwater.

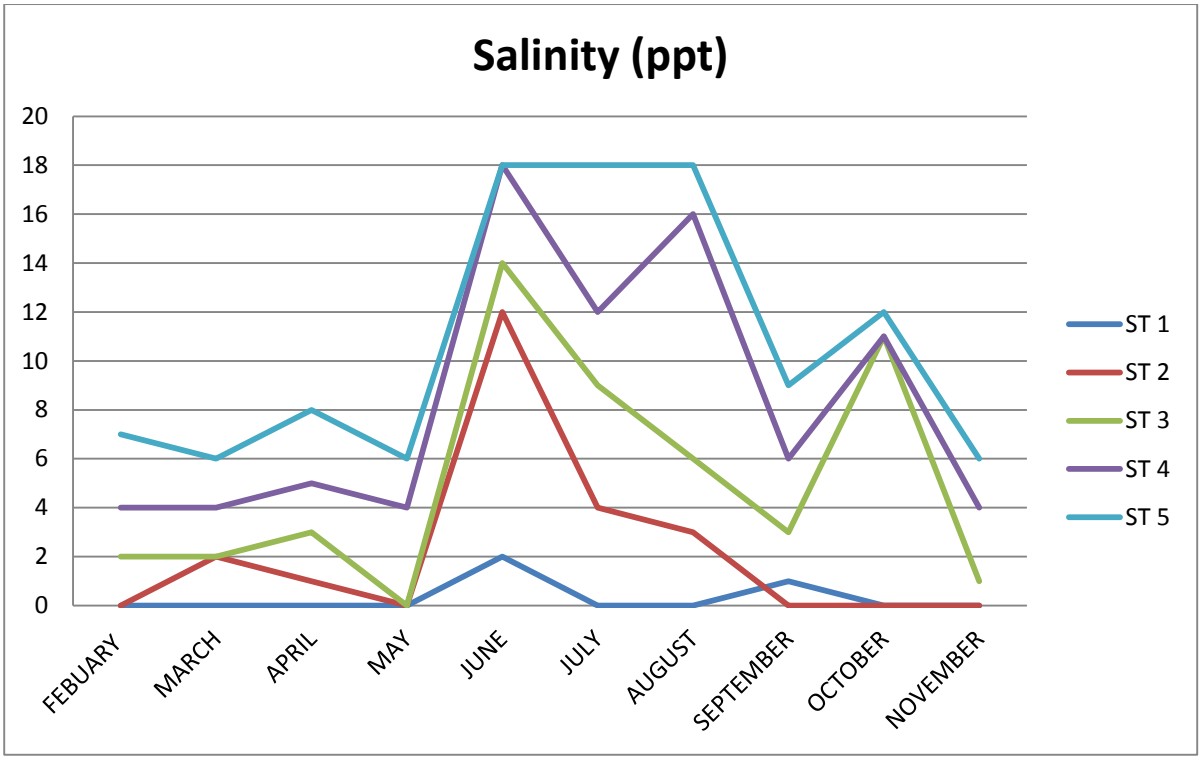


Figure 4.36: Salinity value for all five stations along the river from February 2009 until November 2009

Ammonia levels in excess of the recommended limits may harm aquatic life. Although the ammonia molecule is a nutrient required for life, excess ammonia may accumulate in the organism and cause alteration of metabolism or increases in body pH. It is an indicator of pollution from the excessive usage of ammonia rich fertilizers. Based on FAO 2002, the optimum concentration of ammonia is 0.3 ppm. Only at April 2009 showed low concentration of ammonia in 4 out 5 stations studied while all other stations showed higher concentration of ammonia (Figure 4.37). This suggest that the water has been moderately polluted based on Water Quality Index (WQI) and Interim National Water Quality Standards which in between Class III or Class IV

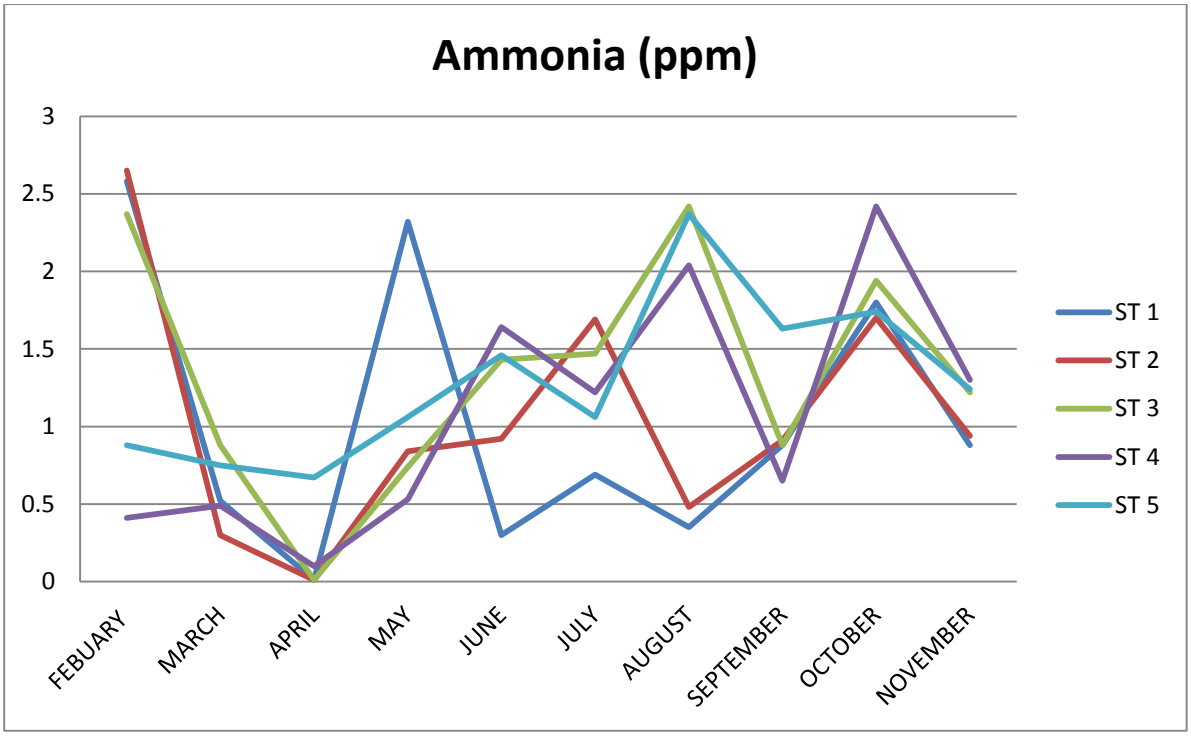


Figure 4.37: Ammonia concentration for all five stations along the river from February 2009 until November 2009

Higher concentrate of nitrate have been reported to associate with mortality in hatcheries (1.8 ppm), but no experiential or experimental definitive information about the toxicity of nitrite to prawns in pond situations is available. However, due to the usage of fertilizers many bodies of freshwater are currently experiencing influxes of nitrogen and phosphorus that will lead to phytoplankton (algae) and macrophyte (aquatic plant) production and caused decreasing of Dissolved Oxygen. Figure 4.38 showed <1.8 ppm of nitrate in all months at all studied stations that suggest the water still suitable for growing up prawn regardless the water bodies were polluted.

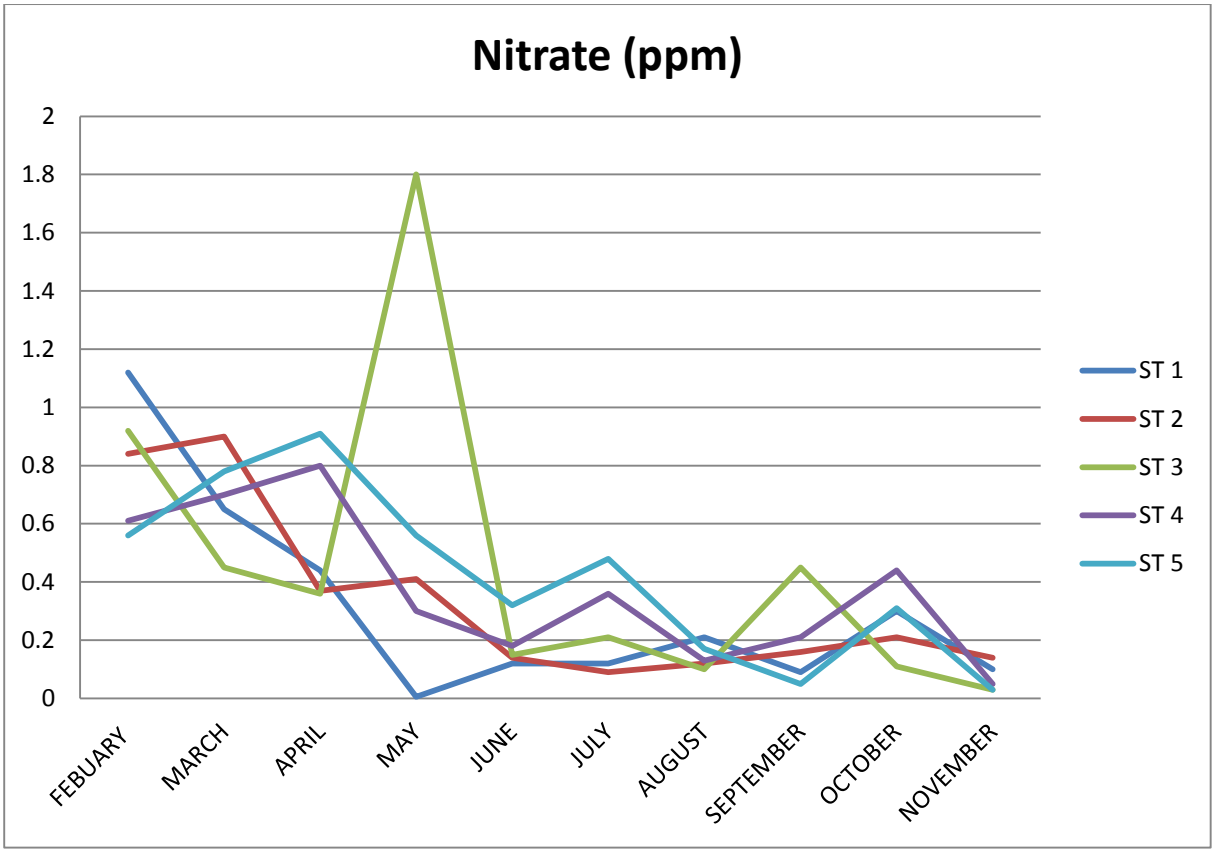


Figure 4.38: Nitrate concentration for all five stations along the river from February 2009 until November 2009

Total Suspended solids (TSS) are an indication of the amount of erosion that took place nearby or upstream. The series of sediment-induced changes that can occur in a water body may change the composition of an aquatic community. First, a large volume of suspended sediment will reduce light penetration, thereby suppressing photosynthetic activity of phytoplankton, algae, and macrophytes. This leads to fewer photosynthetic organisms available to serve as food sources for many invertebrates. As a result, overall invertebrate numbers may also decline, which may then lead to decreased prawn populations. Second the settling of suspended solids from turbid waters threatens benthic aquatic communities. Deposited particles may obscure sources of food, habitat, hiding places, and nesting sites. Most aquatic insects will simply drift with the current out of the affected area. Benthic invertebrates that prefer a low-silt substrate, such as mayflies, stoneflies, and caddis flies, may be replaced by

silt-loving communities of oligochaetae, pulmonate snails, and chironomid larvae that were not a food source to the prawn. Suspended sediment may also affect predator-prey relationships by inhibiting predators' visual abilities which is in this case the predator is the prawn.Based on Interim National Water Quality Standards for Malaysia (INWQS), water bodies that contain TSS concentration of <25 mg/l is classified in Class I which showed in June, July and August 2009. However, figure 4.39 showed the river quite heavily polluted with the concentration of TSS is more than 150 mg/l at the end and the beginning of the year.

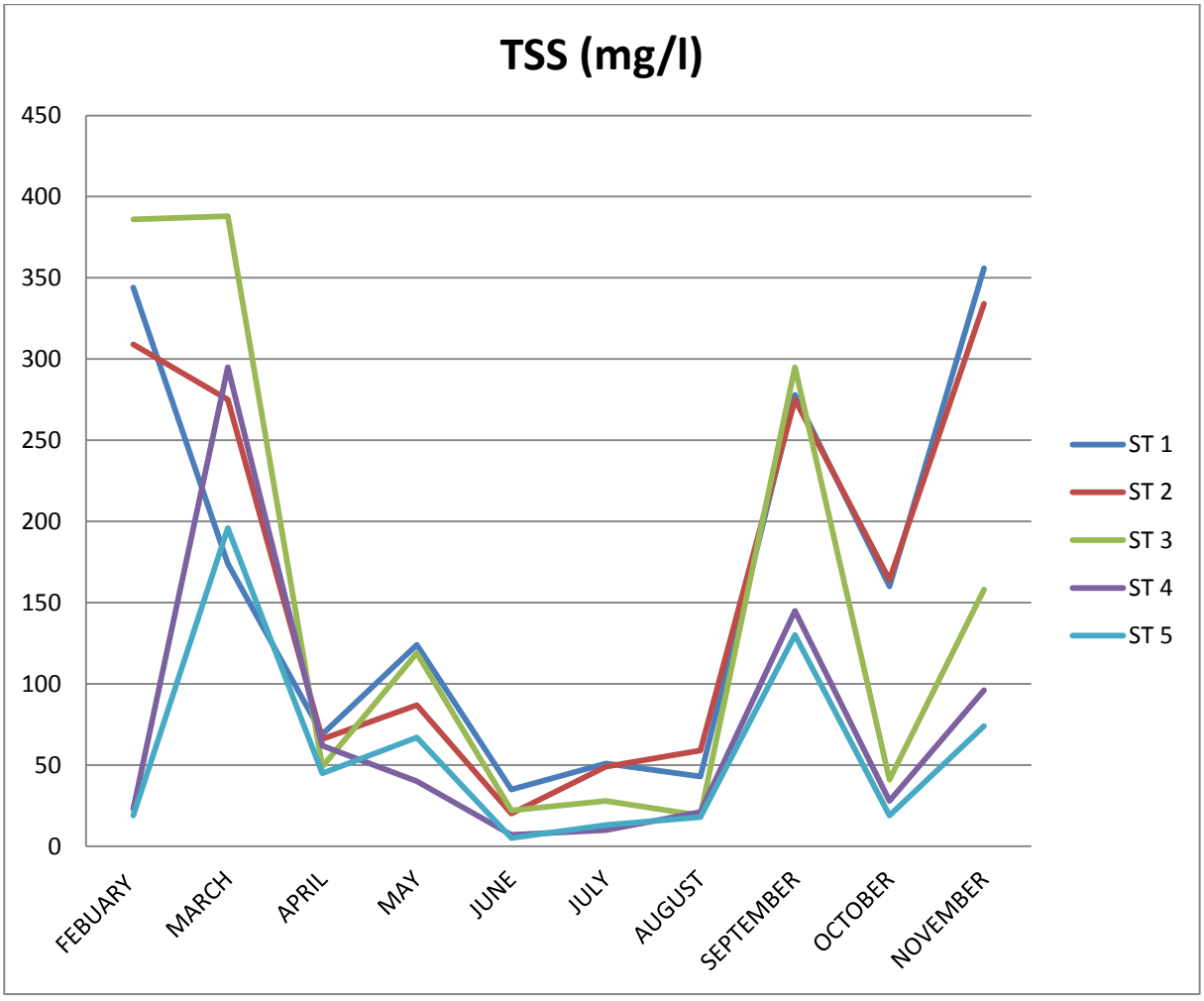


Figure 4.39: Total Suspended Solids for all five stations along the river from February 2009 until November 2009

CHAPTER FIVE

CONCLUSION

5.1 Marker selection

Molecular markers are classified into two categories: type I are markers associated with genes of known function, and type II markers are associated with anonymous genomic segments (O'Brien, 1991 ;). In this study, both of these categories of markers were tested with the studied organisms. EST markers or in the case of this EST SSR are type I markers because they represent transcripts of genes. While type II markers would be the genomics SSR and mitochondria DNA marker.

In this study, at least 35 novel primers were developed. There were14 type II markers and 21 type I markers. However, to make sure that all the primers being used have enough power to differentiate all the samples, PIC values were calculated. PIC refers to the value of a marker for detecting polymorphism in a population. By comparing the PIC value for each primer, 8 EST SSR were chosen from the 21 EST primers which had the highest PIC values. The same goes to the novel genomic SSR where 5 of them were choose from 14 of the novel markers. Another 5 more genomics markers were chosen from the published journal that showed good performance in detecting genetic variation in the sample of *Macrobrachium rosenbergii*.

Null alleles were occurred in five loci namely LM 4, LM 8, LM 12, LM 13 T79112_8d and VJ2_10a as a result of a checking using the of Microchecker software A null allele can be defined as any allele at a microsatellite locus that is only weakly amplified or not visible after

amplification and separation (O'Connell and Wright, 1997) and is recognized, together with population subdivision, as a major factor in depression of observed heterozygosity, compared with that expected on the basis of Hardy-Weinberg equilibrium. The prime reason for null alleles is thought to be the deletion and insertion within the priming site of microsatellite DNA leading to an allele that will not amplify (Callen *et al.*, 1993).

Table 5.1 All categories of Microsatellites marker that were used in this study

Locus	***Null Present***	***Oosterhout***	***Chakraborty***	***Brookfield 1***	***Brookfield 2***
MR4	No	0.2123	0.2321	0.0162	0.0162
MR6	No	0.1931	0.1745	0.1549	0.1549
MR8	No	0.3128	0.2623	0.1171	0.2128
MR19	No	0.1281	0.2162	0.2111	0.2111
MR20	No	0.1273	0.3421	0.0011	0.0011
MR27	No	0.2125	0.2623	0.0011	0
MR29	No	0.0182	0.0244	0.0222	0.0222
MR40	No	-0.1244	-0.0619	-0.0226	0
T79112_7e	No	0.0784	0.1078	0.0775	0.0775
T79112_8d	Yes	0.2134	0.3037	0.1676	0.1676
T79112_9c	No	0.2783	0.2871	0.0671	0
VJ2_10a	Yes	0.1872	0.2457	0.1536	0.1536
VJ2_12j	No	0.1259	0.2653	0.0919	0.919
LM4	Yes	0.1755	0.2264	0.119	0.119
LM8	Yes	0.2808	0.3895	0.2275	0.2275
LM12	Yes	0.0793	0.0925	0.0794	0.0794
LM13	Yes	0.3181	0.5093	0.282	0.282
MBR5	No	-0.0821	-0.0718	-0.0705	0

5.2 Microsatellite Analysis

Microsatellites typing of all three populations indicated that all the 18 loci were polymorphic (100%). A total of 210 alleles were obtained across the 18 loci with the average about 11.67 alleles per loci. The overall allele frequency ranging from 0.0093 to 0.5429. Based on the allelic richness that was obtained by using HP-Rare software, Parental population has the highest allelic richness mean (7.12) while the F1 have the lowest mean (5.54). The result showed the high degree of polymorphism in Parental population, which reflects the high degree of variability of the organisms. Selection for the parental population in Kampung Acheh Hatchery was done by choosing individuals that originated from the Sungai Perak Population and have bigger sizes. However, Wild Type Population has the second lowest allelic richness (5.87) which proved the hypothesis that this population is depleting and decreasing in numbers due to overfishing or environmental problems that lead to the reduction of genetic diversity as the same population showed highest polymorphisms by using RAPD and RAMS marker system studied done by Bhassu *et al*, in 2007. However, there are an increase numbers in allelic richness in F1, F2 and F3 population through the year. This could indicate hybridization occurring between two populations of Parental and Wild Type. From the example of the bar chart (Figure 4.12) in chapter result F1, F2 and F3 population, the result showed allele A, B, and C that were not originated from wild type population but only occurred in Parental Population. This situation seemed to occur in most of the loci studied.

The population structure produced by STRUCTURE software also showed the same result. It demonstrates hybridization that occurred between Wild Type and Parental population that produced F1, F2, and F3 populations that have mixed lineage. However, from the population structure, we could see that the F2 and F3 populations shared a high numbers of

alleles from parental populations. There are several theories behind this condition. One of the theories asserts that the declining of the wild type populations is causing the parental population to domesticate the river. Thus, the numbers of parental population have increased causing the number of wild type population in the river to decrease. It then will produce offspring that are more likely to have the same genetic diversity to parental populations. One question that we need to answer is why we cannot see the domestication phenomena in F1 generation. A study showed that the performance of cultured aquaculture species in the wild tends to be the poorest immediately after release, with observed differences diminishing as a result of developmental and genetic adaptation. Recent studies have shown that second-generation offsprings of cultured organisms spawned in the wild survive better than first-generation released, but remain less fit than wild parentage (Araki *et al.,* 2007; Araki, Cooper & Blouin, 2007). As the first generation (2007), a total of 184800 prawns released, tends to have low survivability, this has caused genetic drift to occur in the bottleneck population of wild type. The genetic drift had caused the wild type to inbreed with each other and produce F1 generation with low genetic diversity. This produced a less fit population as some of the genotype had been lost when inbreeding occurred. In the second year of restocking prawn (2008), a much more fitted cultured prawn had conquered the less fitted wild type prawn causing the domestication to occur. Luckily, the restocking prawn could hybridize with wild type prawns, which lead to an increase in the genetic diversity of the prawn in this river. Unfortunately, some of the rare alleles in wild type population are lost based on the allele frequencies result.

Heterozygosity is a reliable estimator of genetic diversity (MacHugh *et al.,* 1998). The highest mean observed heterozygosity found in the F1 population with a value of 0.7188 while the Wild Type Population had the lowest value of 0.5095. Such results could be explained by several hypotheses including inbreeding, methodology bias, null alleles (Jarne and Lagode,

1996), and founder effect during the introduction (Mei *et al.*, 2003). From this data, we can say that the population of wild type stock at Sungai Timun was very low in number as this showed that there was a possibility of "founder effect" occurring at Sungai Timun. The founder effect is the loss of genetic variation that occurs when a new population (F1) is established by a very small number of individuals from a larger population (Parental). The other explanation was explained by Population specific Inbreeding Coefficient (F_{IS}) value where it has shown that the Wild Type population had the highest positive value (0.1251). This means that inbreeding had occurred in high frequency in this population that led to a deficit of heterozygosity while the lowest in F1 population (-0.1063), wrong interpretation of data because of the null alleles is also one of the reasons for the condition. Heterozygous individuals might be mistyped as homozygotes. Mistyping of heterozygous individuals could explain why the observed heterozygosity in all of the population is significantly different with the expected heterozygosity.

F_{ST} is the value that measures the effect of population subdivision, which is the reduction in heterozygosity in a subpopulation due to genetic drift. It is the most inclusive measure of population substructure, and is useful for examining the overall genetic divergence among subpopulations. In Wright's definition (1978), populations may be classified into four levels of genetic differentiation (F_{ST}): low (0.00 to 0.05), moderate (0.05 to 0.15), high (0.15 to 0.25) and elevated (> 0.25). However, the overall F_{ST} indicated that genetic difference among the population is high, thus clearly showing that population of *Macrobrachium rosenbergii* at Sungai Timun had been divided into subpopulations. The genetic differences between populations might be promoted by the partial or total absence of genetic flow, and could occur due to different adaption environment between the species originated from Sungai Timun (Wild Type) and species originated from Kampung Acheh (Parental).

5.3 Mitochondria Analysis

The data shows some excess number of alleles, as would be expected from a recent population expansion or from genetic hitchhiking in wild type population and F1. This proves the hypothesis that the numbers of prawns were small in the river before the restocking. For the F1 generation, there were about one hundred thousand juveniles being released in April 2007, as the cycle to produce F1 generation between parental population and wild population often takes around 1 year. Unfortunately, a large number of these juveniles did not survive. That is the reason why both of the populations had negative Tajima's D.A positive value of Fu Test and Tajima is evidence for a deficiency of alleles, as would be expected from a recent population bottleneck or from over dominant selection. This situation occurred in parent population where certain phenotype, such as the size of the prawn, was choosen to make as a parent to produce good quality of juveniles. A positive value could also be seen in F2 and F3 population as in time, many of the juveniles had been released into the river. The total of about four hundred thousand juveniles had been released back into the river that made up around 31.25% of the prawns that had been caught within two years at Sungai Timun River. However, a positive Tajima's D value also showed that there was some over dominant problem in the river. This problem had to be controlled to avoid the domestication effect of hatchery stocks and the lost of wild stocks population. The mitochondria result also showed that the parental population had survived and hybridized with the wild type population.This was proven by looking at the haplotypes that had been shared by some of the population. From table 4.15, haplotype number 2 is shared between Wild type and F3 population while there also haplotype sharing between wild type and F1 and F2 population. It's interesting to discover that there was no haplotype of parental population in the F1 population. It is possible that the sample size for F1 population is not big enough to find all the haplotype for the F1 population. The other

reason could be a large number of the first batch of restocking juveniles did not survive in the river, and this supported the negative Tajima's D value in F1 population. However, in many ways the mitochondria result showed that some of the parental population had survived and had successfully increased the stock of prawns and genetic diversity at Sungai Timun, Negeri Sembilan.

5.4 Ecological Study

The Water Quality Index (WQI) and Interim National Water Quality Standards for Malaysia (INWQS) has classified Sungai Timun, Negeri Sembilan in as slightly polluted rivers. Rapid industrialization, urbanization, and infrastructure development at the upstream of Sungai Timun play an important role in the increase of pollution. Both results in benthic and water parameter studies suggested that only stations 4 and 5 served as a conducive condition for prawns to survive. This is because stations 4 and 5 had the suitable parameters, such as temperature and pH for the prawns to inhabit these parts of Sungai Timun, Negeri Sembilan. Unpublished data by Azwan in 2012 showed that a high number of giant freshwater prawns had been captured in this area compared to other station along the river. Station 4 and 5 also had a high number of benthos organisms that served as a food source for the prawns

5.5 General Conclusion

Wild populations remain the dominant source of broodstock for some of the most important aquaculture species. Therefore, sound management of the wild resource is vital for the aquaculture industry. Based on my observations, wild stocks of M. rosenbergii are declining in numbers because of overharvesting, habitat loss, and pollution. These factors may be responsible for the loss of genetic diversity in these populations. In consequence, the loss of genetic diversity through inbreeding can cause weakening of features of adaptation, such as the fitness of individuals and the probability of population survival. Therefore, measuring the genetic diversity in wild stocks is essential for the understanding and effective management of *M. rosenbergii*. Microsatellites and Mitochondria markers can detect polymorphisms in the freshwater prawn population. The dendograms constructed from the microsatellites marker data have revealed the relationship among and between the populations studied. All five populations showed some variations from one another with respect of the primers used. Based on the result obtained, we can conclude that stock enhancement program in Sungai Timun was a success. However, some results showed that the parental population from Kampung Acheh had started to domesticate the river, which lead to the degrading of wild type population of Sungai Timun. The finding is also supported by the unpublished fishermen's statements saying that there was a significant decrease in size of the prawns being captured recently compared to the previous prawns population. In conclusion, this study showed that the stock enhancement program at Sungai Timun, Negeri Sembilan should incorporate a different parental population that have the same or closer genetic content with the wild population of Sungai Timun to reduce the domestication effect and preventing the loss of rare allele of wild population of Sungai Timun.

BIBILIOGRAPHY

Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., ... & Moreno, R. F. (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. Science, 252(5013), 1651-1656.

Alba, M. M., Santibánez-Koref, M. F., & Hancock, J. M. (1999). Conservation of polyglutamine tract size between mice and humans depends on codon interruption. Mol Biol Evol, 16(11), 1641-1644.

Araki, H., Cooper, B., & Blouin, M. S. (2007). Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science, 318(5847), 100-103.

Araki, H., Ardren, W. R., Olsen, E., Cooper, B., & Blouin, M. S. (2007). Reproductive Success of Captive-Bred Steelhead Trout in the Wild: Evaluation of Three Hatchery Programs in the Hood River. Conservation Biology, 21(1), 181-190.

Armstrong, D.A., Stephenson, M.J. & Knight, A.W. (1976). Acute toxicity of nitrate to larvae of the giant Malaysian prawn, *Macrobrachium rosenbergii.* Aquaculture, 9, 39-46.

Arora, R . & Bhatia, S. (2004). Genetic structure of Muzzafarnagri sheep based on micro satellite analysis. Small Ruminant Res., 54, 227-230.

Barral. V., This, P., Imbert-Establet. D., Combes, C. & Delsney. M. (1993). Genetic variability and evolution of the Schistosoma genome analysed by using random amplified polymorphic DNA marker. Mol. Biochem. Parasitol, 59, 211-222.

Bentzen, P., Taggart, C. T., Ruzzante, D. E., & Cook, D. (1996). Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences, 53(12), 2706-2721.

Bhassu, S., See, L.M., R. Hassan, Siraj, S.S., Tan, S.G. (2008). Isolation and characterization of microsatellites loci in the Malaysian giant freshwater prawn, *Macrobrachium rosenbergii.* Molecular Ecology Resources, 8(5), 983-985.

Bhassu, S., R. Hassan, Chew, M.Y., Z. Jamari, See, L.M., & Tan, S.G. (2007). Assessing genetic variability among wild populations of Malaysia giant freshwater prawns, *Macrobrachium rosenbergii* using molecular markers. Malaysian Fisheries Journal, 6 (1), 26-37.

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American journal of human genetics, 32(3), 314.

Brown, W. M. (1983). Evolution of animal mitochondrial DNA. Evolution of Genes and Proteins, 1, 62-88.

Brown, W. M. (1985). The mitochondrial genome of animals. Molecular evolutionary genetics, *95*.

Brookfield, J.F.Y. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology, 5, 453-455.

Burns, C.D., Berrigan, M.E. & Hendelson. (1979). *Fusarium sp*. Infections in the freshwater prawn *Macrobrachium rosenbergii* (de Man). Aquaculture, 16, 193-198.

Callen, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C. & Sutherland, G.R. (1993). Incidence and origin of "null" alleles in the (AC)n microsatellite markers. American Journal of Human Genetics, 52, 922-927.

Castro, J., Bouza, C., Presa, P., Pino-Querido, A., Riaza, A., Ferreiro, I., Sanchez, L. & Martinez, L.M. (2004). Potential sources of error in parentage assessment of turbot (*Scophthalmus maximus*) using microsatellite loci. Aquaculture, 242(1), 119-135.

Chand, V., de Bruyn, M., & Mather, P. B. (2005). Microsatellite loci in the eastern form of the giant freshwater prawn (*Macrobrachium rosenbergii).*Mol. Ecol. Notes 5, 308 – 310.

Chareontawee, K., Poompuang, S., & Na-Nakorn, U. (2006). Isolation and characterization of microsatellites in giant freshwater prawn *Macrobrachium rosenbergii.* Mol. Ecol.Notes, 6, 823 – 825.

Chareontawee, K., Poompuang, S., Na-Nakorn, U., & Kamonrat, W., 2007. Genetic diversity of hatchery stocks of giant freshwater prawn (*Macrobrachium rosenbergii*) in Thailand. Aquaculture, 271(1), 121- 129.

Clement, M., Posada, D. C. K. A., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. Molecular ecology, 9(10), 1657-1659.

Cohen, D., Ra'nan, Z. & Barnes, A. (1983). Production of the freshwater prawn *Macrobrachium rosenbergii* in Israel: I. Integration into fish polyculture systems. Aquaculture, 31, 67-76.

Coloni, A. (1985). A study on the bacterial flora of giant prawn *Macrobrachium rosenbergii*, larvae fed with *Artemia salina naupili*. Aquaculture, 49, 1-10.

Correia, E. S., Suwannatous, S., & New, M. B. (2000). Flow-through hatchery systems and management. Freshwater prawn culture: the farming of *Macrobrachium rosenbergii*, 52-68.

Dakin, E. E., & Avise, J. C. (2004). Microsatellite null alleles in parentage analysis. Heredity, 93(5), 504-509.

D'Abramo, L. R., Ohs, C. L., Fondren, M. W., Steeby, J. A., & Posadas, B. C. (2003). Culture of freshwater prawns in temperate climates: management practices and economics. Office of Agricultural Communications, Division of Agriculture, Forestry, and Veterinary Medicine, Mississippi State University.

De Bruyn, M., Wilson, J. A., & Mather, P. B. (2004). Huxley's line demarcates extensive genetic divergence between eastern and western forms of the giant freshwater prawn *Macrobrachium rosenbergii.* Mol. Phylogen. Evol., 30, 251 – 257.

De Man, J.G. (1879). On some species of the genus *Palaemon Fabr.* with descriptions of two new forms. Notes Leden Museum, 1, 165-184.

De Vicente, M.C., Guzmán, F.A., Engels, J. & Rao, V. R. (2005). Genetic characterization and its use in decision making for the conservation of crop germplasm. The Role of Biotechnology, 121-128.

Divu, D., Khushiramani, R., Malathi, S., Iddya Karunasagar, Indrani Karunasagar. (2007). Isolation, characterization and evaluation of microsatellite DNA markers in giant freshwater prawn *Macrobrachium rosenbergii,* from South India. Aquaculture, 284(1), 281-284.

Dore, I., & Frimodt, C. (1987). An illustrated guide to shrimp of the world (p. 229). New York, New York: Van Nostrand Reinhold.

Doyle, R.W., Singholka, S. & New, M.B. (1983). "Indirect selection" for genetics change: A quantitative analysis illustrated with *Macrobrachium rosenbergii*. Aquaculture, 30, 237-247.

Drake, J. W., Charlesworth, B., Charlesworth, D., & Crow, J. F. (1998). Rates of spontaneous mutation. Genetics, 148(4), 1667-1686.

Edwards, A. L., Hammond, H. A., Jin, L., Caskey, C. T., & Chakraborty, R. (1992). Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics, 12(2), 241-253.

El-Gamal, A.A., Alderman, D.J., Rodger, C.J., Polylase, J.L. & Macintosh, D. (1986). A scanning electron microscope study of oxolinic acid treatment of burn spot lesions of *Macrobrachium rosenbergii.* Aquaculture, 52, 157-171.

Excoffier, L., Laval, G. and Schneider, S. 2005. ARLEQUIN version 3.01: an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47-50.

Folmer, O., Hoeh, W. R., Black, M. B., & Vrijenhoek, R. C. (1994). Conserved primers for PCR amplification of mitochondrial DNA from different invertebrate phyla. Molecular Marine Biology and Biotechnology, 3, 294-299.

FAO. (2000). Aquaculture production statistics 1989-1998. FAO Fisheries Circular 815 (Rev. 12). FAO, Rome.

FAO. (2001). Fishery statistics: aquaculture production 1999. FAO Fisheries Series No. 58/FAO Statistics Series No. 160.

FAO. (2002). Farming freshwater prawn. A manual for the culture of the giant river prawn (*Macrobrachium rosenbergii*). FAO Training Series, 428, 2.

FAO, F., & Food, U. N. (2000). Agricultural Organization (2010). Pyrrolizidine alkaloids in foods and animal feeds. FAO Consumer Protection Fact Sheets, (2), 1-6.

Ferguson, M.M., & Danzmann, R.G. (1998). Role of genetic markers in fisheries and aquaculture: Useful tools or stamp collecting?. Canadian Journal of Fisheries and Aquatic Sciences, 55(7), 1553–1563.

FitzSimmons, N.N., Moritz, C. & Moore, S.S. (1995). Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. Molecular Biology and Evolution, 12, 432-440.

Fujimura, T., & Okamoto, H. (1970). Notes on progress made in developing a mass culturing technique for *Macrobrachium rosenbergii* in Hawaii. Indo-Pacific Fisheries Council, FAO Regional Office for Asia and the Far East.

Garza, J. C., & Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. Molecular ecology, 10(2), 305-318.

George, M. J. (1969). VII Genus Macrobrachium 1868. Cent. Mar. Fish. Res. Inst., Bull. 14, 179-215.

Glaubitz, J. C. (2004). Convert: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Molecular Ecology Notes, 4(2), 309-310.

Gunderina, L.I. 2003. DNA polymorphism of Drosophila genes and factors determining it. Russian Journal of Genetics 39, 737-747.

Hansen, M. M., Kenchington, E., & Nielsen, E. E. (2001). Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries*, *2*(2), 93-112.

Hartman, O. (1939). *Polychaetous Annelids: Parts I-VIII* (Vol. 7). University of Southern California Press.

Hartman, O. (1941). *Some contributions to the biology and life history of Spionidae from California*. University of Southern California Press.

Hartman, O. (1944). *Polychaetous annelids*. Pt. VI. *Paraonidae, Magelonidae, Longosomidae, Ctenodrilidae* and *Sabellariidae*. Allan Hancock Pac. Exped., 10.

Hartman, O. (1947). *Polychaetous annelids*. University of Southern California Press.

Hartman, O. (1954). *Polychaetous annelids* of the Gulf of Mexico. United States Fish and Wildlife Service Bulletin, 55, 413-417.

Harpaz, S. & Schmalbach, E.A. (1986). Important growth and health of the Malaysian prawn *Macrobrachium rosenbergii*, by addition of fresh leaves to the artificial diet. Aquaculture, 55, 81-85.

Hedgecock, D., Stelmach, D.J., Nelson, K., Lindenfelser, M.E., & Malecha S.R. (1979). Genetic divergence and biogeography of natural populations of *Macrobrachium rosenbergii*. Proc. World Maricul. Soc., 10, 873-879.

Hebert, P. D., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270(Suppl. 1), S96-S99.

Herbinger, C.M., Doyle, R.W., Pitman, E.R., Paquet, D., Mesa, K.A., Morris, D.B., Wright, J.M. & Cook, D. (1995). DNA fingerprint based analysis of paternal and maternal effects on offspring growth and survival in communally reared rainbow trout. Aquaculture, 137, 245-256.

Holthuis, L.B. (1980). FAO species catalogue. Volume 1. Shrimps and prawns of the world (an annotated catalogue of species interest to fisheries). FAO Fisheries Synopsis, 125 (1), 103-104.

Hurlbert, S. H. (1971). The nonconcept of species diversity: a critique and alternative parameters. Ecology, 52(4), 577-586.

Ihssen, P. E., Booke, H. E., Casselman, J. M., McGlade, J. M., Payne, N. R., & Utter, F. M. (1981). Stock identification: materials and methods. Canadian Journal of Fisheries and Aquatic Sciences, 38(12), 1838-1855.

Jarne, P. & Lagode, P.J.L. (1996). Microsatellites from molecules to populations and back. Trends in Ecology and Evolution, 11, 424-429.

Karsi, A., Cao, D., Li, P., Patterson, A., Kocabas, A., Feng, J., & Liu, Z. (2002). Transcriptome analysis of channel catfish (*Ictalurus punctatus*): initial analysis of gene expression and microsatellite-containing cDNAs in the skin. Gene, 285(1), 157-168.

Kimura, M. & Weiss, G.H. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics, 49, 561-576.

Kincaid, H. L. (1976). Inbreeding in rainbow trout (*Salmo gairdneri*). Journal of the Fisheries Board of Canada, 33(11), 2420-2426.

Kincaid, H. L. (1983). Inbreeding in fish populations used for aquaculture. Aquaculture, 33(1), 215-227.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., & Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences, 86(16), 6196-6200.

Kochzius, M., Seidel, C., Hauschild, J., Kirchhoff, S., Mester, P., Meyer-Wachsmuth, I., & Timm, J. (2009). Genetic population structures of the blue starfish *Linckia laevigata* and its gastropod ectoparasite *Thyca crystallina*. Marine Ecology Progress Series, 396, 211-219.

Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C., & Sninsky, J. J. (1990). Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Research, 18(4), 999-1005.

Larkin, P. J., & Scowcroft, W. R. (1981). Somaclonal variation—a novel source of variability from cell cultures for plant improvement. Theoretical and Applied Genetics, 60(4), 197-214.

Leberg, P. L. (2002). Estimating allelic richness: effects of sample size and bottlenecks. Molecular ecology, 11(11), 2445-2449.

Levene, H. (1949). On a matching problem arising in genetics. The Annals of Mathematical Statistics, 20(1), 91-94.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25(11), 1451-1452.

Ling, S.W. (1969). Methods of rearing and culturing *Macrobrachium rosenbergii*. FAO Fisheries Reports, 57 (3), 607-619.

Liu, Z.J. & Cordes, J.F. (2004). DNA marker technologies and their application in aquaculture genetics. Aquaculture, 238, 1-37.

Lopez, J. V., Yuhki, N., Masuda, R., Modi, W., & O'Brien, S. J. (1994). Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. Journal of Molecular Evolution, 39(2), 174-190.

Lynch, M. (1995). Analysis of population genetic structure by DNA fingerprinting. DNA Fingerprinting Approach and Applications. 113-126.

Mather, P.B. & de Bruyn, M. (2003). Genetic diversity in wild stocks of the giant freshwater prawn (*Macrobrachium rosenbergii*): implications for aquaculture and conservation. *NAGA, WorldFish Center Quarterly*, 26, 4.

Mather, P. B. (2008). Aquaculture in the Asia-Pacific Region: Applications of molecular population genetics. Pertanika J. Trop. Agric. Sci. *31*(1), 117 – 125.

Machugh, D.E., Loftus, R.T., Cunningham, P. & Bradley, D.G. (1998). Genetic structure of seven European cattle breeds assessed using 20 microsatellite markers. Animal Genetics, 29 (5), 333-340.

McDonald, G.J., Danzmann, R.G. & Ferguson, M.M. (2004). Relatedness determination in the absence of pedigree information in three cultured stains of rainbow trout (*Oncorhynchus mykiss*). Aquaculture 233, 65–78.

McMillen-Jackson, A. L., & Bert, T. M. (2004). Genetic diversity in the mtDNA control region and population structure in the pink shrimp *Farfantepenaeus duorarum*. Journal of Crustacean Biology, 24(1), 101-109.

Mei, T.Z., Eric, S.H.T. & Alex, H.T.Y. (2003). Isolation and cross species amplification of microsatellites loci in the freshwater minnow *Zacco pachycephalus* (Teleostei: *Cyprinidae*) for diversity and conservation genetic analysis. Molecular Ecology Notes, 3, 567-569.

Mitchell, M. K., & Stapp, W. B. (1994). Field manual for water quality monitoring: An environmental education program for schools. Thomson-Shore. Inc, Dexter, Michigan, 277.

Moore, S.S., Whan, V., Davis, G.P., Byrne, K., Hetzel, D.J.S. & Preston, N. (1999). The development and application of genetic markers for the kuruma prawn *Penaeus japonicus*, Aquaculture, 173, 19-32.

Nachman, M. W., Brown, W. M., Stoneking, M., & Aquadro, C. F. (1996). Nonneutral mitochondrial DNA variation in humans and chimpanzees. Genetics, 142(3), 953-963.

Naylor, R. L., Goldburg, R. J., Primavera, J. H., Kautsky, N., Beveridge, M. C., Clay, J., ... & Troell, M. (2000). Effect of aquaculture on world fish supplies. Nature, 405(6790), 1017-1024.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci.USA, 70, 3321-3323.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, 89, 583-590.

Ng, P.K.L. (1997). The conservation status of freshwater prawns and crabs in Singapore with emphasis on the Nature Reserves. Proc. Nature Res. Surv. Sem. Gardens' Bull. Sing., 49, 267-272.

New, M.B. and Singholka, S. 1985. Freshwater prawn farming. A manual for the culture of *Macrobrachium rosenbergii*. FAO Fisheries Technical Paper 225, 118.

New, M.B. (2000). History and global status of freshwater prawn farming. In M.B. New & W.C. Valenti (eds.), Freshwater Prawn Culture: the Farming of *Macrobrachium rosenbergii* (pp. 1-11). Oxford: Blackwell Science.

New, M.B., Singholka, S., & Kutty, M.N. (2000). Prawn capture fisheries and enhancement. In M.B. New & W.C. Valenti (eds.), Freshwater Prawn Culture: the Farming of *Macrobrachium rosenbergii* (pp. 411-428). Oxford: Blackwell Science.

Nielsen, C. (1994). Larval and Adult Characters in Animal Phytogeny. American Zoologist, 34(4), 492-501.

Norris, A. T., Bradley, D. G., & Cunningham, E. P. (1999). Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) populations. Aquaculture, 180(3), 247-264.

O' Brien, S.J. 1991. Molecular genome mapping: lessons and prospects. Current Opinion in Genetics and Development (1), 105-111.

O'Connell, M. & Wright, J.M. (1997). Microsatellite DNA in fishes. Reviews in Fish Biology and Fisheries, 7,331-363.

Ohta, T., & Kimura, M. (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genetical research, *22*(02), 201-204.

Ong, B.V. (1983). Progress in selecting an appropriate culture system for a small scale *Macrobrachium rosenbergii* (de Man) hatchery. Aquaculture, 35, 267- 272.

O'Reilly, P.T., Herbinger, C. & Wright, J.M., (1988). Analysis of parentage determination in Atlantic salmon (*Salmo salar*) using microsatellites. Animal Genetics, 29, 363-370.

O'Reilly, P.T., & Wright, J.M. (1995). The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. J. Fish Biol. 47, 29–55.

Oosterhout C.V., C., Hutchinson, W. F., Wills, D. P., & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4(3), 535-538.

Pereira, S. L., Grau, E. T., & Wajntal, A. (2004). Molecular architecture and rates of DNA substitutions of the mitochondrial control region of cracid birds. Genome, 47(3), 535-545.

Perez, F., Ortiz, J., Zhinaula, M., Gonzabay, C., Caldero, J., Volckaert, F.A.M.J. (2005). Development of EST-SSR markers by data mining in three species of shrimp: *Litopenaeus vannamei*, *Litopenaeus stylirostris*, and *Trachypenaeus birdy*. Marine Biotechnology, 7, 554-569.

Philip G.L., Rodrick, G.E., Sodeman, W.A.Jr. & Blake, N.J. (1982). The giant Malaysian prawn, *Macrobrachium rosenbergii*, a potential predator for controlling the spread of schistosome vector snails in fish ponds. Aquaculture, 28, 293- 301.

Piyan, N.T., Law, A.T. & Cheah, S. H. (1985). Toxic levels of mercury for sequential larval stages of *Macrobrachium rosenbergii* (de Man). Aquaculture, 46, 353- 359.

Putney, S. D., Herlihy, W. C., Schimmel, P. (1983). A new troponin T and cDNA clones for 13 different muscle proteins, found by shotgun sequencing. Nature, 302, 718–721

Qi Bin, K. (2011). Global differential transcriptomic expression of IHHNV-infected *Macrobrachium rosenbergii* using high-throughput solexa sequencing (Master dissertation). Retrieved from University Malaya Database. (TP248.15 UM 2011 Kwo)

Raymond, M. & Rousset, F. (1995). GENEPOP (ver 1.2): A population genetics software for exact test and ecumenisms. J. Heredity, 86, 248-249.

Rozen, S. and Skaletsky, H. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Molecular Biology 132, 365-386.

Rico, C. Rico, I. & Hewitt, G. (1996). 470 million years of conservation of microsatellite loci among fish species. Proceedings of Royal Society of London B: Biological Sciences, 263, 549-557.

Sagi, A., Ra'nan, Z., Cohen, D. & Wax, Y. (1986). Production of *Macrobrachium rosenbergii* in monosex populations: Yield characteristic under intensive monoculture conditions in cages. Aquaculture, 51, 262-275.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T. & Erlich, H. A. (1988). Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science, 239(4839), 487-491.

Schierwater. B., Streit. B., Wagner. G.P. & Desalle., R. (1994). Molecular ecology and evolution: Approach and applications. Birkhauser Verlag, Basel, Switzerland, 495-507.

Schlotterer, C., Amos, B. & Tautz, D. (1991). Conservation of polymorphic simple sequence loci in cetacean species. Nature, 354, 63-65.

Schlotterer, C. and Tautz, D. (1992). Slippage synthesis of simple sequence DNA. Nucleic Acids Research, 20, 211-215.

Schneider, S., Roessli, D., & Excoffier, L. (2000). Arlequin: a software for population genetics data analysis. User manual ver. 2, 2496-2497.

Scribner, K. T., & Pearce, J. M. (2000). Microsatellites: evolutionary and methodological background and empirical applications at individual, population, and phylogenetic levels. Molecular methods in ecology, 235-273.

See, L.M., R. Hassan, Tan, S.G., Subha B. (2008). Genetic characterization of wild stocks of prawn *Macrobrachium rosenbergii* using random amplified polymorphic DNA markers. Biotechnology, 7 (2), 338-342.

See, L.M., Bhassu, S., R. Hassan, Siraj, S.S., Tan, S.Y. & Tan, S.G. (2007). A population genetics study on the Malaysian wild stocks, *Macrobrachium rosenbergii* using cross amplified microsatellites primers. Pertanika J. Trop. Agric. Sci., 30 (1), 71-82.

Serapion, J., Kucuktas, H., Feng, J., Liu, Z. (2004). Bioinformatic mining of Type 1 microsatelite from Expressed Sequence Tags of Channel Catfish (*Ictalurus punctatus*). Mar. Biotechnol, 6, 364-377.

Smith, S. B. (1971). Macroscopic benthos populations and taxonomy of the family Chironomidae in Lake Poinsett, South Dakota (Doctoral dissertation, South Dakota State University).

Smith, W., & Grassle, J. F. (1977). Sampling properties of a family of diversity measures. Biometrics, 283-292.

Smith, E. P., Zahran, A., Mahmoud, M., & Ye, K. (2003). Evaluation of water quality using acceptance sampling by variables. Environmetrics, 14(4), 373-386.

Sorenson, M. D., & Quinn, T. W. (1998). Numts: a challenge for avian systematics and population biology. The Auk, 214-221.

Strachan, T., & Read, A. P. (1999). Molecular pathology. In M.B. New & W.C. Valenti (eds.), Human Molecular Genetics. 2nd edition (pp. 211-228). New York: Wiley-Liss

Sudsuk, S., & Sodsuk, P, K. (1998). Genetic diversity of giant freshwater prawn from three locations in Thailand. Technical paper No. 18/1998. National Aquaculture Genetics Research Institute, Pathumthani, Thailand, 40 pp.

Sunier, A. L. J. (1925). Twee mededeelingen over Palaemoniden. Tijdschrift der Nederlandsche Dierkundige Vereeniging. Leiden (2), 19, 115-117.

Tamura, K., Dudley, J., Nei, M., & Kumar, S. (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular biology and evolution, 24(8), 1596-1599.

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. Genetics, 105(2), 437-460.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123(3), 585-595.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Research, 17, 6463-6471.

Testolin, R.T., Marrazzo, G. Cipriani, R. Quarta, I., Verde, M.T., Dettori, M., Pancaldi & s. Sansavini. (2000). Microsatellites DNA in peach (*Prunus persica*) and its use in fingerprinting and testing the genetic origin of cultivars. Genome, 43, 512-520.

Thanh, N. M., Ponzoni, R. W., Nguyen, N. H., Vu, N. T., Barnes, A., & Mather, P. B. (2009). Evaluation of growth performance in a diallel cross of three strains of giant freshwater prawn (*Macrobrachium rosenbergii*) in Vietnam. Aquaculture, 287(1), 75-83.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic acids research, 25(24), 4876-4882.

Thrower, F. P., Hard, J. J., & Joyce, J. E. (2004). Genetic architecture of growth and early life-history transitions in anadromous and derived freshwater populations of steelhead. Journal of Fish Biology, 65(s1), 286-307.

Uno, J. & Kwan, C. S. (1969). Larvae development of *Macrobrachium rosenbergii* (de Man) reared in the laboratory. Japan Tokyo Univ. Fish. 55 (2), 179-190.

UNFPA United Nations Population Fund. (2011). State of world population 2010.

Valdes, A. M., Slatkin, M., & Freimer, N. B. (1993). Allele frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics, 133(3), 737-749.

Waugh, J. (2007). DNA barcoding in animal species: progress, potential and pitfalls. BioEssays, 29(2), 188-197.

Weber, J. L., & Wong, C. (1993). Mutation of human short tandem repeats. Human molecular genetics, 2(8), 1123-1128.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358–1370.

Weir, B. S., & Hill, W. G. (2002). Estimating F-statistics. Annual Review of Genetics, 36(1), 721-750.

Yýlmaz Ç. & Brahim O. (2002). Fish population genetics and applications of molecular markers to fisheries and aquaculture: I- basic principles of fish population genetics. Turkish Journal of Fisheries and Aquatic Sciences, 2, 145-155.

APPENDIXES

Appendix A: Genotype data of F3 Population (54 Samples)

	MR 4		MR 6		MR 8		MR 13		MR 19		MR 20		MR 27		MR 38		MR 40		LM 13		LM 8		MBR 5		LM 16		LM 4		T917_7e		T917_7d		T917_9c		VJ2_ 10a	
1	311	313	240	200	270	261	204	204	255	245	244	250	271	283	265	265	288	303	257	257	335	335	326	340	250	250	244	252	278	278	288	288	157	157	164	172
2	309	311	236	196	270	270	204	195	255	245	247	250	283	283	255	255	288	300	260	248	287	302	340	366	235	235	246	252	284	344	288	288	136	157	156	164
3	295	295	240	196	294	261	204	195	255	255	235	250	283	286	250	250	303	303	293	269	302	302	340	366	238	238	248	260	296	320	296	308	124	157	164	168
4	313	315	240	200	270	270	204	195	255	245	235	244	283	286	250	250	303	303	299	305	311	311	326	366	244	244	236	260	269	320	288	320	136	157	160	162
5	311	321	244	204	270	270	204	195	255	245	235	244	262	283	250	250	288	300	284	257	302	302	300	366	250	250	244	252	272	338	288	332	124	157	162	162
6	313	315	244	204	270	261	204	195	255	245	235	247	280	280	250	250	276	300	275	299	323	323	326	366	244	244	244	248	272	272	284	340	151	157	162	172
7	297	317	248	208	276	261	204	204	255	245	235	235	283	283	250	250	288	303	284	257	323	323	326	366	238	238	244	244	272	332	284	344	157	157	172	172
8	297	321	252	208	270	261	204	186	255	245	247	247	271	283	245	245	288	303	260	260	299	317	326	366	244	244	242	252	296	296	284	284	157	157	168	172
9	313	315	260	216	270	270	204	195	255	245	235	244	283	286	245	245	303	303	284	284	323	302	326	366	244	244	252	260	344	344	288	288	157	157	160	172
10	297	317	268	216	270	270	204	204	255	245	235	244	280	283	245	245	288	300	317	248	302	302	296	366	244	244	244	260	278	332	288	288	157	157	160	162
11	311	327	220	220	270	270	204	195	255	245	235	244	271	283	240	240	288	303	281	236	317	317	310	350	244	244	236	236	272	272	308	308	151	157	162	162
12	295	311	228	228	270	261	204	195	255	255	235	247	271	283	250	250	285	288	317	305	311	311	326	366	247	247	244	248	278	344	288	320	151	157	164	176
13	291	291	284	232	261	261	207	204	255	255	244	247	271	274	240	240	282	288	305	305	302	302	296	366	244	244	244	244	284	305	296	296	157	157	168	172
14	311	307	284	236	270	261	204	189	255	245	235	235	283	289	240	240	285	288	317	260	302	302	310	310	253	253	230	260	278	332	288	332	157	157	160	172
15	293	307	240	200	270	261	204	195	245	255	244	247	271	283	310	310	282	303	290	269	302	302	300	310	244	253	236	260	278	332	284	344	151	157	160	168
16	313	313	240	200	291	261	204	195	275	255	235	235	262	283	305	305	285	288	275	275	302	302	288	310	250	250	240	260	305	338	284	284	124	157	160	164
17	313	313	240	200	291	261	204	195	255	245	235	244	280	283	300	300	288	303	317	269	302	302	326	366	253	253	236	236	296	344	288	320	124	157	160	176
18	311	319	228	192	261	261	204	195	255	255	235	250	271	283	300	300	288	303	293	269	311	311	310	350	247	250	236	242	278	332	288	288	157	157	168	172
19	311	317	228	188	270	261	204	195	255	245	235	244	262	283	290	290	276	303	299	305	323	323	310	350	244	244	240	256	272	296	284	344	145	157	168	168
20	293	297	224	188	261	261	204	204	255	245	235	244	271	283	285	285	288	300	263	299	329	329	310	310	238	238	236	236	284	338	344	344	124	157	160	168

21	313	319	192	192	276	261	204	195	255	245	235	247	262	283	285	285	273	276	299	269	296	296	288	288	250	250	230	244	278	329	336	344	145	157	162	162
22	313	315	192	192	279	261	204	204	255	245	235	235	262	283	280	280	285	303	311	293	302	302	280	300	250	250	236	244	278	278	288	308	124	157	160	172
23	311	321	232	192	261	261	204	195	255	245	235	250	283	286	275	275	288	303	275	275	302	302	280	326	244	244	236	242	284	338	320	344	130	157	160	172
24	313	315	236	196	261	261	204	204	255	245	235	244	262	265	280	280	282	303	305	311	302	302	310	310	244	244	236	244	278	329	296	296	151	157	160	172
25	297	317	228	208	276	261	210	210	255	245	235	247	283	286	280	270	279	303	275	260	299	299	310	350	244	244	236	260	305	305	320	344	145	157	160	162
26	311	327	224	200	291	261	210	210	255	245	244	244	280	280	280	270	288	303	260	260	302	302	280	376	244	244	248	260	278	332	288	308	124	157	156	160
27	295	311	216	192	270	270	213	213	255	245	247	247	265	286	280	270	288	288	317	317	323	323	326	366	250	250	230	244	284	284	292	332	145	157	160	172
28	291	291	212	188	270	261	210	210	260	245	235	235	283	289	285	270	282	303	281	284	323	323	326	340	244	244	240	256	272	338	292	332	124	157	160	168
29	311	321	216	192	270	270	210	210	260	245	235	247	289	289	280	270	288	303	275	275	317	317	326	366	253	253	230	244	305	317	288	288	130	157	156	172
30	311	317	216	192	270	270	213	213	255	245	244	247	271	283	280	270	288	300	284	257	302	302	326	340	247	244	244	244	278	332	288	308	151	157	156	176
31	311	313	216	192	276	261	213	213	255	245	235	235	274	280	280	270	288	303	269	290	302	302	340	366	244	244	260	260	278	296	288	344	145	151	160	168
32	311	321	216	192	282	261	213	213	255	245	235	244	277	280	280	270	288	303	236	269	311	311	340	366	250	250	248	252	296	344	296	336	124	157	160	168
33	311	315	220	200	270	270	216	216	255	245	235	235	280	280	280	270	300	300	305	311	311	311	326	366	238	238	236	248	311	311	288	288	145	157	160	168
34	311	315	220	196	270	270	213	213	255	245	235	247	265	286	280	270	288	303	269	248	296	296	310	350	244	244	244	256	284	284	284	344	130	130	160	162
35	307	311	220	192	270	270	210	210	245	255	244	247	283	289	280	270	300	300	248	236	317	317	310	350	253	250	242	256	272	344	284	332	130	157	156	176
36	311	321	220	200	276	261	210	210	255	245	235	244	289	289	280	270	288	303	275	281	302	302	296	350	253	253	256	256	284	284	320	320	124	157	156	168
37	313	315	224	200	294	261	210	210	255	245	235	244	262	283	280	270	279	303	284	275	311	311	310	350	244	244	230	230	284	344	296	320	124	157	168	168
38	297	317	220	196	276	261	213	213	255	245	235	247	283	286	280	270	303	303	275	275	302	302	300	366	244	244	240	256	278	332	284	320	124	130	164	168
39	311	327	220	200	261	261	207	207	255	245	235	247	262	283	280	270	288	300	248	245	323	323	310	366	250	250	230	244	272	338	208	336	130	157	156	168
40	291	307	216	196	270	270	207	207	255	245	244	247	283	283	280	270	273	276	260	248	311	311	310	366	250	250	256	260	278	272	288	296	130	136	156	162
41	311	313	196	196	270	261	210	210	255	245	244	247	283	286	280	270	285	303	275	281	299	317	296	350	235	235	248	260	272	338	284	344	157	157	168	172
42	311	311	212	188	261	261	210	210	255	245	235	244	271	283	280	270	288	303	248	263	302	302	310	350	238	238	242	242	284	284	320	332	124	157	160	164
43	311	319	204	188	270	270	210	210	255	245	235	235	271	283	280	270	285	288	281	260	329	338	296	350	244	244	230	236	296	305	288	344	130	157	156	156
44	313	315	200	188	270	270	210	210	255	245	241	247	262	283	280	270	282	303	299	299	308	323	300	310	250	250	236	236	278	332	292	332	151	151	156	168
45	313	313	200	180	276	261	213	213	255	245	235	250	283	286	280	270	288	303	275	260	302	302	310	310	244	244	236	246	272	305	288	288	145	145	160	160
46	293	311	200	180	270	261	213	213	255	245	238	244	271	283	280	270	303	303	260	260	317	317	296	296	244	244	236	236	278	329	288	344	130	157	160	172
47	309	311	196	176	270	270	213	213	255	245	235	247	271	283	280	260	288	300	236	269	317	317	296	300	238	238	230	242	272	278	296	308	136	157	172	172
48	311	313	200	176	270	270	213	213	255	245	244	247	262	283	280	270	288	303	275	275	302	302	300	366	244	244	230	260	296	296	296	320	124	136	162	172
49	311	321	196	176	270	261	210	210	255	245	235	244	271	283	280	270	288	303	248	245	323	323	310	366	250	250	236	260	296	329	284	336	130	157	156	168
50	311	313	192	172	276	261	213	213	255	245	235	244	271	283	280	270	285	303	275	275	317	317	326	366	253	253	230	244	284	344	284	296	136	136	156	172
51	307	315	188	176	270	261	216	216	255	245	235	244	271	274	280	270	288	303	284	257	302	302	326	340	247	244	236	244	284	338	320	320	130	157	156	172
52	311	319	188	168	276	261	219	219	255	245	235	244	283	286	280	270	285	288	281	260	329	338	296	350	244	244	230	244	278	332	284	344	130	130	160	162
53	311	313	188	168	276	261	219	219	255	245	247	250	271	274	280	270	282	303	299	299	308	323	300	310	250	250	236	242	278	332	284	336	130	157	156	168
54	311	313	192	168	261	261	225	225	255	245	247	250	262	262	280	270	288	303	275	275	302	302	300	366	244	244	240	256	317	338	284	296	136	136	156	172

Appendix B: Genotype data of F2 Population (52 Samples)

	MR 4		MR 6		MR 8		MR 13		MR 19		MR 20		MR 27		MR 38		MR 40		LM 13		LM 8		MBR 5		LM 16		LM 4		T917_7e		T917_7d		T917_9c		VJ2_ 10a	
1	307	307	224	200	279	270	213	213	255	245	244	250	283	286	285	285	282	303	284	257	311	311	326	366	244	244	256	256	278	338	288	308	124	157	156	156
2	311	311	224	200	267	267	216	216	280	270	247	250	271	283	285	285	288	300	260	260	302	302	326	366	238	238	240	256	272	338	288	340	124	157	160	160
3	311	313	236	200	264	264	210	210	255	245	235	250	271	283	300	300	279	303	284	284	323	323	296	366	238	238	230	236	284	284	284	340	136	157	156	160
4	311	311	224	200	264	264	210	210	255	245	235	244	283	283	300	300	288	303	317	248	311	311	310	350	244	244	230	236	278	278	284	344	157	157	168	172
5	311	307	224	196	309	264	210	210	255	235	244	244	283	286	300	300	303	303	281	236	299	317	310	350	238	238	236	236	272	338	284	284	157	157	156	168
6	293	307	228	196	270	270	207	207	255	255	232	244	283	286	295	295	288	303	317	305	302	302	296	350	244	244	236	246	272	272	288	288	151	157	160	168
7	309	313	232	200	285	270	216	216	245	255	247	253	262	283	300	300	300	300	305	305	329	338	300	310	253	253	236	236	272	332	288	288	157	157	160	162
8	311	311	228	196	270	261	222	222	255	255	244	244	283	283	300	300	288	303	317	260	308	323	310	310	244	238	256	256	296	296	288	288	157	157	162	162
9	311	319	224	196	270	261	225	225	255	245	247	253	283	286	305	305	288	300	290	269	302	302	296	296	238	238	240	256	272	272	288	320	157	157	162	172
10	313	315	220	196	270	261	228	228	255	245	247	247	271	283	310	310	288	303	305	311	317	317	296	300	244	244	236	236	278	332	296	296	157	157	172	172
11	309	313	216	192	273	273	222	222	255	255	235	244	271	283	315	315	288	303	275	281	317	317	296	326	253	250	240	256	272	344	288	332	157	157	168	172
12	311	315	216	192	270	261	222	222	255	245	235	244	262	283	310	310	288	303	248	263	302	302	296	350	253	253	236	236	278	344	284	344	136	157	160	172
13	293	311	216	188	270	261	231	231	255	255	247	250	271	283	310	310	288	303	281	260	302	302	310	350	244	244	230	244	284	332	284	284	124	157	160	168
14	309	311	216	188	270	261	237	237	255	255	247	250	271	283	315	315	300	300	299	299	323	323	296	340	244	244	240	256	278	332	288	288	136	157	160	164
15	311	313	184	184	279	270	234	234	255	255	235	244	271	274	280	280	288	303	275	260	323	323	310	310	238	238	256	256	278	332	288	288	124	157	164	172
16	311	321	212	184	297	270	234	234	255	255	235	250	283	289	280	280	288	303	260	260	317	317	310	350	244	244	240	256	278	338	284	344	157	157	156	164
17	311	321	248	220	270	261	204	204	255	255	238	244	271	283	275	275	288	303	248	245	335	335	280	376	250	250	230	244	296	344	344	344	145	157	164	168
18	311	315	244	220	270	261	210	210	255	255	247	247	262	283	275	275	288	300	260	248	287	302	310	350	244	244	256	260	278	332	336	344	145	157	160	162
19	311	311	244	216	279	270	210	210	255	255	235	244	280	283	285	285	288	300	281	260	302	302	326	366	250	250	248	260	278	332	288	344	124	157	162	162
20	309	311	240	212	297	276	210	210	255	245	235	247	271	283	275	275	288	300	299	299	311	311	326	340	235	235	236	260	284	338	320	344	145	157	162	172

21	313	315	240	208	270	261	207	207	255	255	247	250	262	283	275	275	288	300	236	269	302	302	340	366	238	238	240	260	278	338	296	296	124	157	172	172
22	297	317	232	212	270	270	207	207	255	245	235	244	271	283	275	275	285	303	305	311	323	323	340	366	244	253	236	236	278	278	288	332	130	157	168	172
23	297	321	232	204	270	270	210	210	255	255	253	253	262	283	280	280	288	303	275	260	323	323	326	366	247	244	236	242	284	338	284	336	151	157	160	172
24	313	315	236	208	270	270	210	210	255	255	247	250	262	283	280	280	300	300	260	260	299	317	300	366	244	244	246	246	284	338	284	332	145	151	160	168
25	297	317	244	212	270	261	207	207	255	255	235	250	283	286	275	275	288	303	317	317	323	302	310	366	250	250	242	256	278	332	292	332	124	157	156	172
26	311	327	236	208	270	270	207	207	255	255	235	244	271	283	275	275	288	303	281	284	302	302	310	366	244	244	256	256	278	332	288	288	124	157	164	164
27	311	313	208	208	297	270	210	210	255	255	235	244	268	283	270	270	285	300	275	275	317	317	296	350	247	250	240	256	278	338	288	288	124	157	160	162
28	311	319	228	204	270	270	210	210	255	255	235	247	283	286	265	265	300	300	284	257	317	317	310	350	235	247	244	252	272	338	288	344	157	157	156	168
29	311	313	228	200	261	261	213	213	255	255	235	250	289	289	275	275	288	303	275	299	302	302	310	350	244	244	244	248	284	311	296	336	145	157	156	168
30	315	317	224	200	285	276	213	213	255	255	238	244	262	283	280	280	288	300	269	248	302	302	310	310	250	250	244	244	278	278	288	288	124	157	168	168
31	297	321	196	196	378	273	213	213	255	255	235	247	271	283	285	285	288	303	248	236	311	311	326	366	244	244	236	244	296	296	284	344	145	157	160	168
32	313	315	224	196	270	270	213	213	255	255	235	253	262	283	285	285	288	300	275	281	311	311	326	366	244	244	236	260	284	284	284	332	124	157	162	162
33	297	317	220	192	270	270	210	210	255	255	235	244	283	286	295	295	288	300	284	275	296	296	326	366	250	250	248	260	284	311	308	320	130	157	160	172
34	311	321	212	192	270	270	207	207	255	245	235	244	283	283	300	300	288	303	275	275	302	302	296	366	244	244	230	244	284	284	296	320	130	130	160	172
35	307	311	212	188	270	270	201	201	255	255	235	244	280	286	305	305	273	303	257	257	323	323	310	350	253	253	240	248	284	344	288	288	124	130	160	172
36	311	317	212	188	270	261	195	195	255	255	235	244	283	286	310	310	288	288	260	248	329	329	310	350	253	253	244	252	284	284	288	344	130	157	160	162
37	311	313	208	188	270	261	192	192	255	255	235	244	280	280	310	310	288	303	293	269	296	296	296	350	247	250	246	252	284	344	296	336	130	136	156	168
38	311	313	212	184	270	270	189	189	255	255	235	247	283	283	315	315	288	303	299	305	302	302	310	350	253	253	248	260	278	332	284	344	157	157	156	168
39	307	315	212	188	270	270	186	186	255	255	235	235	271	283	320	320	288	300	263	299	302	302	310	310	250	250	236	260	272	338	284	332	124	157	168	168
40	311	311	216	192	270	261	186	186	255	255	244	250	283	286	320	320	288	300	284	257	302	302	300	310	250	250	236	236	278	338	288	344	130	157	160	168
41	291	319	184	184	270	270	186	186	255	245	247	250	280	283	325	325	288	303	269	290	299	299	288	310	244	244	244	244	272	338	292	332	151	151	162	162
42	311	315	208	192	294	270	186	186	255	255	247	250	271	283	320	320	300	300	236	269	311	311	296	326	244	244	260	260	284	284	288	288	145	145	168	172
43	311	315	208	184	270	270	183	183	255	255	235	244	271	283	315	315	288	303	305	311	302	302	296	350	244	244	244	260	278	278	288	344	124	151	160	172
44	311	311	188	188	270	270	186	186	255	255	235	247	271	283	280	280	288	303	299	269	302	302	340	366	244	244	236	236	278	332	296	336	124	145	160	168
45	309	311	188	188	270	270	183	183	255	255	244	244	262	283	310	310	288	303	311	293	302	302	340	366	238	238	244	248	272	272	296	320	157	157	156	172
46	311	317	216	192	270	270	180	180	255	255	235	235	283	286	315	315	276	303	275	275	302	302	326	366	250	250	244	244	272	338	284	336	151	157	156	156
47	307	311	216	192	270	270	183	183	255	255	247	247	283	286	310	310	282	303	293	269	311	311	300	366	250	250	236	244	272	320	284	296	157	157	156	156
48	311	317	220	192	291	270	180	180	255	255	235	244	283	286	300	300	288	303	?	?	?	?	?	?	?	?	?	?	278	338	284	344	157	157	156	168
49	311	319	208	188	270	270	213	213	255	255	235	235	283	286	295	295	288	303	?	?	?	?	?	?	?	?	?	?	278	278	288	288	157	157	160	160
50	311	317	204	184	270	270	210	210	255	255	247	247	262	283	280	280	288	303	?	?	?	?	?	?	?	?	?	?	284	338	284	340	124	157	160	172
51	293	297	208	184	291	270	210	210	255	255	235	244	283	286	280	280	288	303	?	?	?	?	?	?	?	?	?	?	284	338	284	344	130	157	172	172
52	307	311	208	188	270	270	204	204	255	245	235	250	271	283	275	275	288	300	?	?	?	?	?	?	?	?	?	?	278	332	284	284	136	157	162	172

Appendix C: Genotype data of F1 Population (32 Samples)

	MR 4		MR 6		MR 8		MR 13		MR 19		MR 20		MR 27		MR 38		MR 40		LM 13		LM 8		MBR 5		LM 16		LM 4		T917_7e		T917_7d		T917_9c		VJ2_ 10a	
1	293	307	192	200	273	282	195	204	245	260	244	250	283	289	270	285	288	303	284	275	311	311	310	340	244	244	236	236	284	332	284	344	124	136	160	160
2	307	311	200	200	276	282	195	204	245	260	247	250	271	283	270	285	288	303	269	269	302	302	326	350	253	253	244	244	278	332	288	344	130	157	160	164
3	311	313	212	192	291	291	216	216	245	245	235	250	283	286	280	280	288	303	275	275	329	329	310	310	238	238	236	236	272	272	284	344	157	157	160	172
4	311	311	200	200	300	273	216	216	245	245	235	244	283	286	280	280	303	303	284	257	329	329	326	366	244	244	256	260	272	338	296	296	124	124	160	162
5	311	311	188	196	267	279	204	204	245	255	235	244	271	283	270	285	288	300	284	257	299	317	310	366	244	244	236	236	272	272	284	340	124	124	160	168
6	311	315	196	200	267	267	195	204	245	255	247	250	262	283	265	285	279	303	317	317	323	302	310	376	253	250	244	248	278	332	284	344	130	157	160	168
7	311	313	196	196	261	261	216	216	240	240	235	235	283	286	290	290	285	288	284	284	311	311	296	366	244	244	236	244	278	332	284	284	124	136	168	168
8	311	313	212	192	276	261	216	216	245	245	247	247	262	283	295	295	282	288	260	290	311	311	310	366	244	244	236	260	305	305	288	288	130	157	160	168
9	313	313	196	196	309	270	213	213	250	250	235	244	283	283	295	295	285	288	293	269	323	323	296	340	244	244	256	256	278	338	288	288	136	136	162	162
10	293	307	192	192	270	270	210	210	250	250	235	244	283	286	305	305	282	303	260	260	317	317	296	350	244	244	240	256	278	311	284	344	130	157	168	172
11	309	313	192	192	309	270	210	210	250	250	247	250	271	283	305	305	285	288	281	236	302	302	310	350	238	238	236	236	305	338	344	344	124	130	160	172
12	311	315	212	192	270	270	207	207	250	250	247	250	271	283	300	300	288	303	284	257	302	302	310	350	250	250	236	236	284	338	336	344	136	136	160	168
13	313	313	196	196	270	261	201	201	260	260	235	244	262	283	290	290	288	303	260	260	299	299	310	350	235	235	244	244	278	332	288	344	130	130	156	172
14	293	311	212	188	270	261	198	198	260	260	235	235	283	286	285	285	285	303	275	275	299	299	296	350	238	238	260	260	278	332	288	288	130	157	156	156
15	309	311	192	192	276	270	192	192	260	260	235	247	271	283	290	290	303	303	275	275	296	296	310	350	244	244	244	260	278	338	284	344	124	124	156	156
16	311	313	212	188	276	270	192	192	260	260	235	235	271	283	290	290	288	300	317	284	302	302	310	350	250	250	236	236	311	311	288	344	124	157	156	168
17	311	321	208	188	276	270	231	204	255	255	235	244	283	283	290	290	288	300	284	275	317	317	310	310	244	244	248	260	284	284	288	332	136	157	160	160
18	311	315	192	192	270	261	204	204	260	260	235	244	283	286	295	295	285	303	269	269	317	317	326	366	244	244	230	244	278	278	284	336	124	157	160	172
19	311	311	212	192	261	261	207	207	255	255	235	244	271	283	300	300	288	303	284	275	311	311	326	366	250	250	248	260	305	311	284	296	157	157	172	172
20	309	311	212	192	270	270	204	204	250	250	235	247	271	283	270	280	300	300	290	269	299	299	326	366	244	244	236	260	284	284	284	344	157	157	162	172
21	311	313	192	200	279	288	195	204	245	255	247	250	283	286	270	285	288	288	293	269	299	299	280	366	250	250	240	256	284	284	292	332	124	157	162	172
22	291	307	200	204	267	270	195	204	245	255	244	247	271	283	265	285	300	300	299	305	296	296	280	350	244	244	230	236	284	284	288	288	130	157	172	172
23	311	313	188	196	267	273	195	204	260	280	244	247	271	283	270	285	288	303	263	299	302	302	280	376	253	253	230	236	284	344	288	344	130	130	168	172
24	311	311	196	200	270	273	195	204	245	255	244	247	262	268	270	285	300	300	281	299	302	302	376	326	247	250	236	236	284	284	296	336	124	130	160	172
25	311	319	192	200	273	282	195	204	245	255	244	247	283	286	285	285	288	303	248	263	296	296	376	340	244	244	256	260	284	344	296	296	130	157	160	168
26	313	319	212	188	270	270	198	198	250	250	244	244	271	283	290	290	288	303	269	248	296	296	310	350	244	244	240	244	278	332	284	344	124	157	160	162
27	311	313	216	188	270	261	192	192	245	245	247	247	271	283	290	290	288	303	245	245	299	299	310	350	244	244	236	236	272	338	284	332	124	157	162	162

28	311	319	188	188	270	270	192	192	245	245	235	235	271	283	285	285	300	300	248	245	302	302	310	350	244	244	230	244	305	305	292	332	124	157	162	172
29	311	313	196	200	267	273	195	204	245	255	235	244	283	286	270	285	303	303	311	293	302	302	340	376	238	238	240	256	272	332	284	344	130	157	160	164
30	311	313	220	192	270	270	204	204	270	270	235	244	262	283	285	285	300	300	260	248	302	302	310	350	244	244	248	260	296	296	288	288	124	157	160	160
31	311	315	188	200	276	279	204	204	245	255	235	247	283	286	280	285	276	300	293	269	296	296	300	310	244	244	240	256	272	272	344	344	124	157	168	168
32	291	327	196	200	273	276	195	204	255	255	235	244	271	283	270	285	288	303	275	260	299	317	300	300	244	244	230	244	278	332	336	344	124	157	156	156

Appendix D: Genotype data of Wild Type Population (35 Samples)

	MR 4		MR 6		MR 8		MR 13		MR 19		MR 20		MR 27		MR 38		MR 40		LM 13		LM 8		MBR 5		LM 16		LM 4		T917_7e		T917_7d		T917_9c		VJ2_ 10a	
1	307	307	216	196	288	288	210	210	250	250	244	250	271	283	280	280	285	288	293	269	329	329	310	310	244	244	230	236	305	332	284	336	130	157	160	172
2	311	311	220	200	324	291	210	210	245	245	247	250	283	286	275	275	282	303	317	248	323	323	310	376	238	238	230	236	278	332	284	296	130	136	160	172
3	311	313	212	192	291	291	216	216	245	245	235	250	283	286	280	280	288	303	275	275	329	329	310	310	238	238	236	236	272	272	284	344	157	157	160	172
4	311	311	200	200	300	273	216	216	245	245	235	244	283	286	280	280	303	303	284	257	329	329	326	366	244	244	256	260	272	338	296	296	124	124	160	162
5	311	311	220	196	270	270	219	219	245	245	235	244	283	286	280	280	288	300	317	260	323	323	326	366	238	238	244	248	272	272	288	332	130	157	156	168
6	307	311	196	196	270	270	216	216	240	240	235	247	262	283	290	290	288	303	317	275	323	323	326	366	244	244	244	244	305	332	284	344	136	157	156	168
7	311	313	196	196	261	261	216	216	240	240	235	235	283	286	290	290	285	288	284	284	311	311	296	366	244	244	236	244	278	332	284	284	124	136	168	168
8	311	313	212	192	276	261	216	216	245	245	247	247	262	283	295	295	282	288	260	290	311	311	310	366	244	244	236	260	305	305	288	288	130	157	160	168
9	313	313	196	196	309	270	213	213	250	250	235	244	283	283	295	295	285	288	293	269	323	323	296	340	244	244	256	256	278	338	288	288	136	136	162	162
10	293	307	192	192	270	270	210	210	250	250	235	244	283	286	305	305	282	303	260	260	317	317	296	350	244	244	240	256	278	311	284	344	130	157	168	172
11	309	313	192	192	309	270	210	210	250	250	247	250	271	283	305	305	285	288	281	236	302	302	310	350	238	238	236	236	305	338	344	344	124	130	160	172
12	311	315	212	192	270	270	207	207	250	250	247	250	271	283	300	300	288	303	284	257	302	302	310	350	250	250	236	236	284	338	336	344	136	136	160	168
13	313	313	196	196	270	261	201	201	260	260	235	244	262	283	290	290	288	303	260	260	299	299	310	350	235	235	244	244	278	332	288	344	130	130	156	172
14	293	311	212	188	270	261	198	198	260	260	235	235	283	286	285	285	285	303	275	275	299	299	296	350	238	238	260	260	278	332	288	288	130	157	156	156
15	309	311	192	192	276	270	192	192	260	260	235	247	271	283	290	290	303	303	275	275	296	296	310	350	244	244	244	260	278	338	284	344	124	124	156	156
16	311	313	212	188	276	270	192	192	260	260	235	235	271	283	290	290	288	300	317	284	302	302	310	350	250	250	236	236	311	311	288	344	124	157	156	168
17	311	321	208	188	276	270	231	204	255	255	235	244	283	283	290	290	288	300	284	275	317	317	310	310	244	244	248	260	284	284	288	332	136	157	160	160
18	311	315	192	192	270	261	204	204	260	260	235	244	283	286	295	295	285	303	269	269	317	317	326	366	244	244	230	244	278	278	284	336	124	157	160	172
19	311	311	212	192	261	261	207	207	255	255	235	244	271	283	300	300	288	303	284	275	311	311	326	366	250	250	248	260	305	311	284	296	157	157	172	172
20	309	311	212	192	270	270	204	204	250	250	235	247	271	283	270	280	300	300	290	269	299	299	326	366	244	244	236	260	284	284	284	344	157	157	162	172
21	311	313	212	192	273	270	186	186	250	250	244	247	283	286	305	305	288	303	290	248	299	299	296	366	250	250	256	256	284	284	288	340	130	157	156	168
22	307	311	212	188	300	273	186	186	250	250	235	235	283	286	305	305	288	288	317	260	299	299	310	350	250	250	240	256	311	311	284	340	130	157	156	172
23	311	313	192	192	270	270	186	186	250	250	235	250	283	286	315	315	282	303	293	275	302	302	310	350	253	253	236	236	284	311	288	344	124	157	160	168
24	311	311	216	188	270	270	186	186	250	250	235	244	283	283	295	295	288	303	275	260	302	302	296	350	250	250	236	236	284	284	292	292	124	157	160	168
25	311	319	192	192	270	261	201	201	250	250	235	247	283	286	295	295	288	300	260	260	302	302	310	350	250	250	244	244	305	305	284	284	136	136	160	168
26	313	319	212	188	270	270	198	198	250	250	244	244	271	283	290	290	288	303	269	248	296	296	310	350	244	244	240	244	278	332	284	344	124	157	160	162
27	311	313	216	188	270	261	192	192	245	245	247	247	271	283	290	290	288	303	245	245	299	299	310	350	244	244	236	236	272	338	284	332	124	157	162	162

28	311	319	188	188	270	270	192	192	245	245	235	235	271	283	285	285	300	300	248	245	302	302	310	350	244	244	230	244	305	305	292	332	124	157	162	172
29	311	313	184	184	309	270	231	204	280	280	235	250	262	283	280	280	288	303	257	257	311	311	296	350	244	244	230	244	272	332	288	288	124	124	156	156
30	311	313	220	192	270	270	204	204	270	270	235	244	262	283	285	285	300	300	260	248	302	302	310	350	244	244	248	260	296	296	288	288	124	157	160	160
31	307	315	216	192	309	270	204	204	270	270	244	247	283	286	285	285	288	303	248	257	323	323	310	350	250	250	236	260	272	272	288	344	157	157	156	160
32	291	311	212	192	270	270	204	204	260	260	235	244	271	283	285	285	288	303	269	290	311	311	296	350	244	244	240	260	278	332	296	336	124	157	168	172
33	313	315	204	184	300	273	216	216	260	260	235	247	271	283	280	270	288	300	317	317	302	302	310	350	244	244	236	236	311	311	288	288	130	157	156	168
34	311	321	188	188	300	273	213	213	260	260	244	244	271	283	280	270	288	300	281	284	302	302	310	350	250	250	256	260	272	338	288	288	130	130	160	160
35	313	315	200	184	270	270	210	210	260	260	244	244	262	283	270	270	288	303	269	290	296	296	296	366	244	244	256	256	272	272	288	288	157	157	168	168

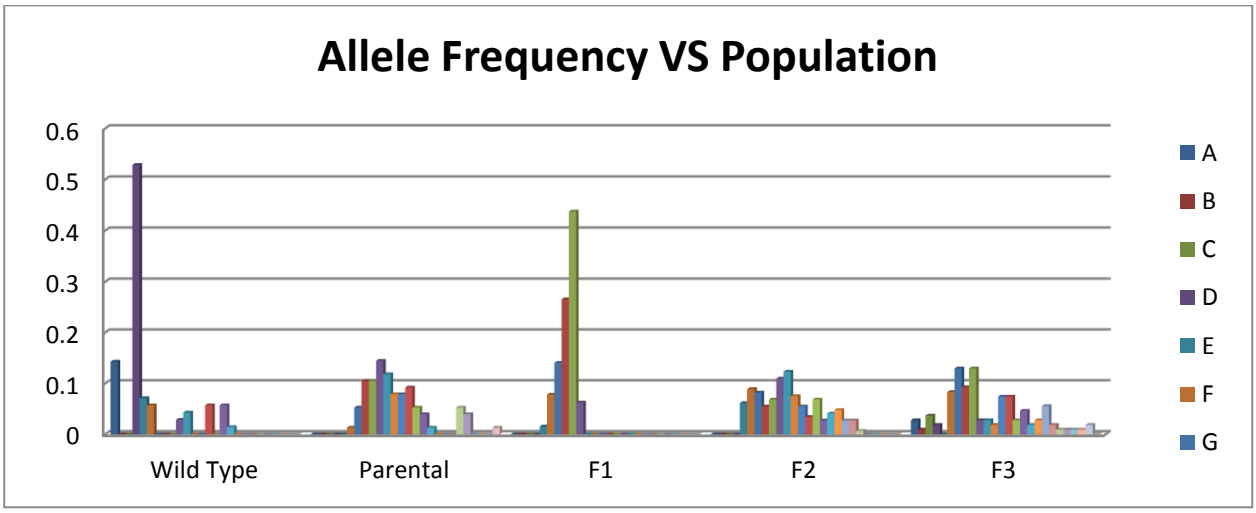
Appendix E: Genotype data of Parental Population (55 Samples)

	MR 4		MR 6		MR 8		MR 13		MR 19		MR 20		MR 27		MR 38		MR 40		LM 13		LM 8		MBR 5		LM 16		LM 4		T917_7e		T917_7d		T917_9c		VJ2_ 10a	
1	311	313	224	196	270	261	186	186	255	245	244	244	283	283	285	285	300	300	284	269	296	296	326	366	250	250	246	252	272	278	284	296	145	151	160	160
2	311	311	192	192	291	270	186	186	255	255	232	244	283	286	285	285	273	303	284	263	311	311	326	366	238	247	252	260	272	320	288	288	157	157	160	176
3	291	319	216	188	270	270	186	186	255	245	247	253	262	265	285	285	300	300	290	275	311	302	326	366	238	238	230	252	320	320	320	320	130	157	160	160
4	311	315	216	192	276	276	186	186	255	245	244	244	283	286	290	290	285	288	281	257	317	317	310	300	250	244	244	248	272	338	284	288	136	157	164	168
5	291	327	232	200	270	261	207	207	255	245	247	253	280	280	285	285	273	279	290	269	317	317	310	350	244	238	244	244	296	338	344	344	151	151	156	164
6	311	317	224	196	270	270	207	207	255	245	241	250	265	286	285	285	288	303	290	269	308	317	326	366	238	238	236	244	284	332	288	308	145	145	156	160
7	307	311	220	196	291	270	207	207	255	245	241	241	283	289	285	285	279	303	290	269	329	323	326	366	244	244	236	246	284	332	284	332	145	157	164	176
8	311	317	220	192	270	261	204	204	255	245	232	244	289	289	285	285	303	303	317	305	320	329	300	310	238	238	236	236	296	296	284	336	151	151	160	160
9	311	313	216	196	270	270	195	195	255	245	235	241	271	283	280	280	288	300	305	305	338	338	310	376	238	238	244	244	320	320	284	296	157	157	162	176
10	291	291	220	196	270	270	198	198	255	245	241	247	274	280	280	280	273	276	275	299	335	335	310	376	253	250	240	244	272	305	284	284	157	157	168	168
11	311	313	196	196	270	270	198	198	255	250	241	247	277	280	275	275	285	303	269	248	296	302	310	376	250	250	236	236	305	311	296	296	124	157	160	160
12	297	307	248	212	270	270	186	186	255	250	235	250	271	283	280	280	288	303	248	236	323	323	310	366	253	250	230	242	311	311	308	332	151	151	160	168
13	297	317	248	220	270	270	189	189	255	245	235	244	277	280	310	310	300	300	311	293	311	311	280	326	247	244	230	236	344	344	308	308	157	157	156	172
14	311	327	220	200	270	261	189	189	255	245	253	253	283	289	310	310	288	303	275	275	302	302	300	300	250	250	230	236	329	344	336	288	130	145	156	164
15	295	311	200	200	291	270	189	189	255	245	235	250	283	286	320	320	288	303	293	269	311	311	280	300	244	244	236	236	284	284	288	288	151	157	160	172
16	291	291	200	200	270	261	195	195	255	245	235	235	271	274	320	320	276	276	281	281	302	302	288	288	244	247	256	260	284	296	288	344	145	157	160	176
17	311	307	204	204	270	261	201	201	255	245	235	250	262	262	320	320	276	279	275	275	329	311	280	300	250	250	248	260	320	344	284	344	130	157	164	172
18	293	307	208	208	270	270	201	201	280	260	232	244	271	277	320	320	276	288	284	260	317	317	288	326	250	250	236	260	320	338	284	308	157	157	160	164
19	309	313	212	212	270	270	201	201	245	255	244	247	283	286	325	325	279	288	269	269	299	296	280	326	244	244	240	260	338	344	284	336	151	157	162	172
20	295	313	208	208	270	270	201	201	255	255	235	250	271	271	315	315	288	303	290	290	299	299	280	340	253	253	236	236	329	344	284	296	130	157	162	172
21	313	327	204	204	270	270	201	201	255	245	238	244	271	283	325	325	279	279	281	263	302	302	296	340	253	253	236	242	272	329	288	344	130	157	156	172
22	317	317	208	208	270	270	198	198	255	245	235	247	271	280	305	305	288	300	260	260	311	311	296	340	247	250	246	246	278	329	308	340	151	157	164	168
23	295	313	228	212	270	270	192	192	255	255	235	253	271	289	300	300	285	288	299	299	311	287	296	340	250	250	242	256	278	317	292	340	145	151	156	172
24	311	313	220	192	270	270	204	204	270	270	235	244	262	283	285	285	300	300	260	248	302	302	310	350	244	244	248	260	296	296	288	288	124	157	160	160
25	307	315	216	192	309	270	204	204	270	270	244	247	283	286	285	285	288	303	248	257	323	323	310	350	250	250	236	260	272	272	288	344	157	157	156	160

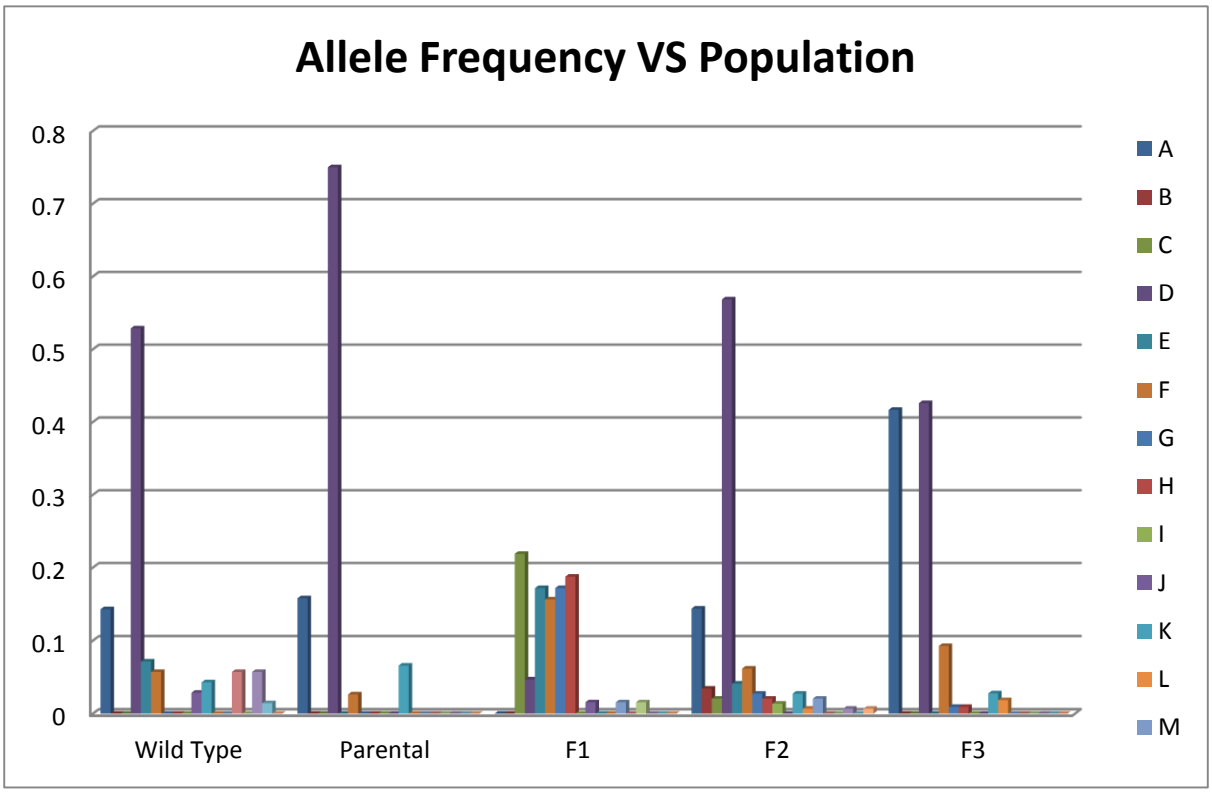
26	307	311	212	188	300	273	186	186	250	250	235	235	283	286	305	305	288	288	317	260	299	299	310	350	250	250	240	256	311	311	284	340	130	157	156	172
27	311	313	192	192	270	270	186	186	250	250	235	250	283	286	315	315	282	303	293	275	302	302	310	350	253	253	236	236	284	311	288	344	124	157	160	168
28	311	311	216	188	270	270	186	186	250	250	235	244	283	283	295	295	288	303	275	260	302	302	296	350	250	250	236	236	284	284	292	292	124	157	160	168
29	311	313	208	208	270	270	192	192	255	245	238	250	268	283	290	290	300	300	299	269	302	302	310	366	250	250	256	256	305	317	308	308	124	157	168	168
30	311	313	212	192	291	291	216	216	245	245	235	250	283	286	280	280	288	303	275	275	329	329	310	310	238	238	236	236	272	272	284	344	157	157	160	172
31	311	311	200	200	300	273	216	216	245	245	235	244	283	286	280	280	303	303	284	257	329	329	326	366	244	244	256	260	272	338	296	296	124	124	160	162
32	311	311	220	196	270	270	219	219	245	245	235	244	283	286	280	280	288	300	317	260	323	323	326	366	238	238	244	248	272	272	288	332	130	157	156	168
33	313	319	212	188	270	270	198	198	250	250	244	244	271	283	290	290	288	303	269	248	296	296	310	350	244	244	240	244	278	332	284	344	124	157	160	162
34	311	313	216	188	270	261	192	192	245	245	247	247	271	283	290	290	288	303	245	245	299	299	310	350	244	244	236	236	272	338	284	332	124	157	162	162
35	311	319	188	188	270	270	192	192	245	245	235	235	271	283	285	285	300	300	248	245	302	302	310	350	244	244	230	244	305	305	292	332	124	157	162	172
36	311	313	184	184	309	270	231	204	280	280	235	250	262	283	280	280	288	303	257	257	311	311	296	350	244	244	230	244	272	332	288	288	124	124	156	156
37	291	311	248	216	270	261	186	186	255	245	238	244	283	283	280	280	282	303	275	260	323	323	310	376	253	253	252	252	272	278	320	344	130	157	172	176
38	311	315	252	216	261	261	186	186	255	245	247	247	271	283	280	280	279	303	260	248	299	317	326	366	244	244	244	244	329	329	308	284	130	157	156	176
39	315	317	248	220	291	270	186	186	255	245	235	250	271	283	290	290	303	303	281	260	323	302	288	288	247	250	260	260	344	344	288	288	136	157	176	176
40	297	321	252	224	291	270	189	189	255	245	244	247	283	286	295	295	276	300	299	299	302	302	326	376	244	244	244	260	344	344	288	344	124	145	164	176
41	313	315	252	220	270	261	186	186	255	235	241	250	283	289	300	300	288	303	299	269	302	302	280	326	253	253	236	236	344	344	284	344	157	157	162	168
42	311	313	220	192	270	270	204	204	270	270	235	244	262	283	285	285	300	300	260	248	302	302	310	350	244	244	248	260	296	296	288	288	124	157	160	160
43	307	315	216	192	309	270	204	204	270	270	244	247	283	286	285	285	288	303	248	257	323	323	310	350	250	250	236	260	272	272	288	344	157	157	156	160
44	307	311	212	188	300	273	186	186	250	250	235	235	283	286	305	305	288	288	317	260	299	299	310	350	250	250	240	256	311	311	284	340	130	157	156	172
45	311	313	192	192	270	270	186	186	250	250	235	250	283	286	315	315	282	303	293	275	302	302	310	350	253	253	236	236	284	311	288	344	124	157	160	168
46	311	311	220	196	270	270	219	219	245	245	235	244	283	286	280	280	288	300	317	260	323	323	326	366	238	238	244	248	272	272	288	332	130	157	156	168
47	313	319	212	188	270	270	198	198	250	250	244	244	271	283	290	290	288	303	269	248	296	296	310	350	244	244	240	244	278	332	284	344	124	157	160	162
48	311	313	216	188	270	261	192	192	245	245	247	247	271	283	290	290	288	303	245	245	299	299	310	350	244	244	236	236	272	338	284	332	124	157	162	162
49	311	319	188	188	270	270	192	192	245	245	235	235	271	283	285	285	300	300	248	245	302	302	310	350	244	244	230	244	305	305	292	332	124	157	162	172
50	311	313	184	184	309	270	231	204	280	280	235	250	262	283	280	280	288	303	257	257	311	311	296	350	244	244	230	244	272	332	288	288	124	124	156	156
51	311	327	220	200	261	261	207	207	255	245	235	247	262	283	280	270	288	300	248	245	323	323	310	366	250	250	230	244	272	338	208	336	130	157	156	168
52	291	307	216	196	270	270	207	207	255	245	244	247	283	283	280	270	273	276	260	248	311	311	310	366	250	250	256	260	278	272	288	296	130	136	156	162
53	311	313	196	196	270	261	210	210	255	245	244	247	283	286	280	270	285	303	275	281	299	317	296	350	235	235	248	260	272	338	284	344	157	157	168	172
54	311	311	212	188	261	261	210	210	255	245	235	244	271	283	280	270	288	303	248	263	302	302	310	350	238	238	242	242	284	284	320	332	124	157	160	164
55	311	319	204	188	270	270	210	210	255	245	235	235	271	283	280	270	285	288	281	260	329	338	296	350	244	244	230	236	296	305	288	344	130	157	156	156

Appendix F: Alleles Frequency of all Loci

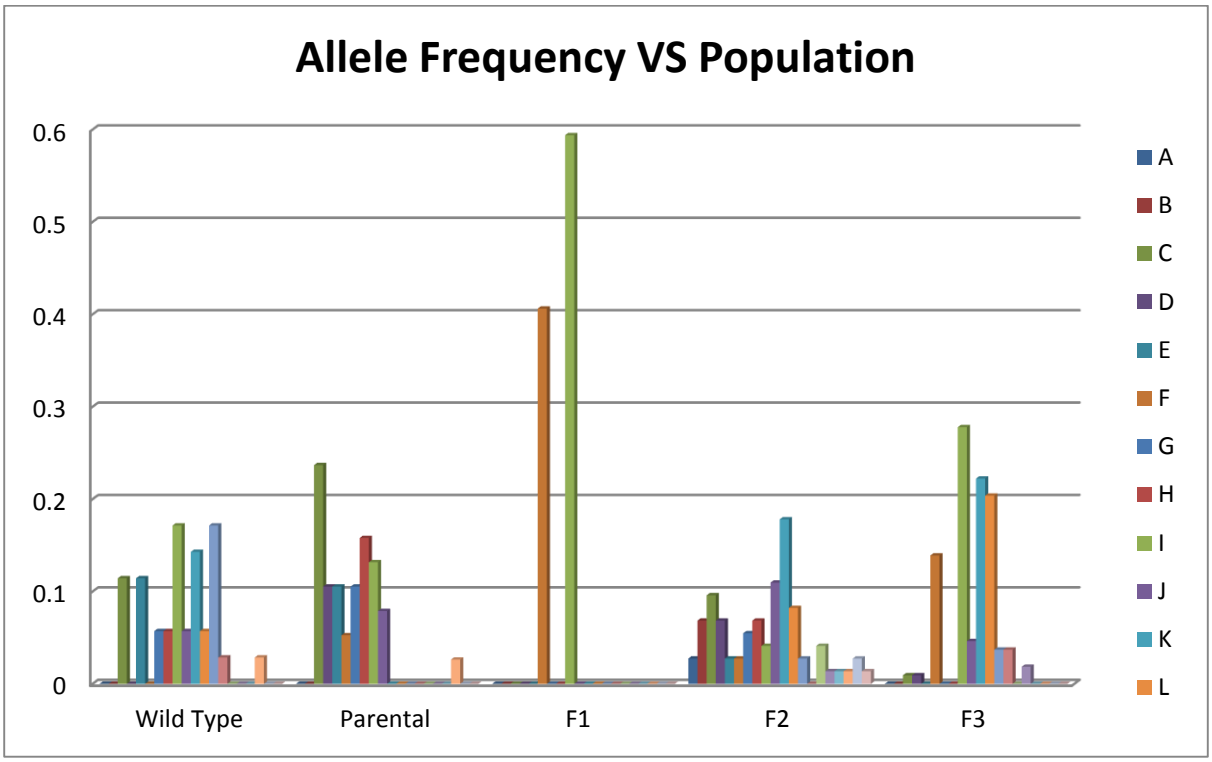
Loci: MR6



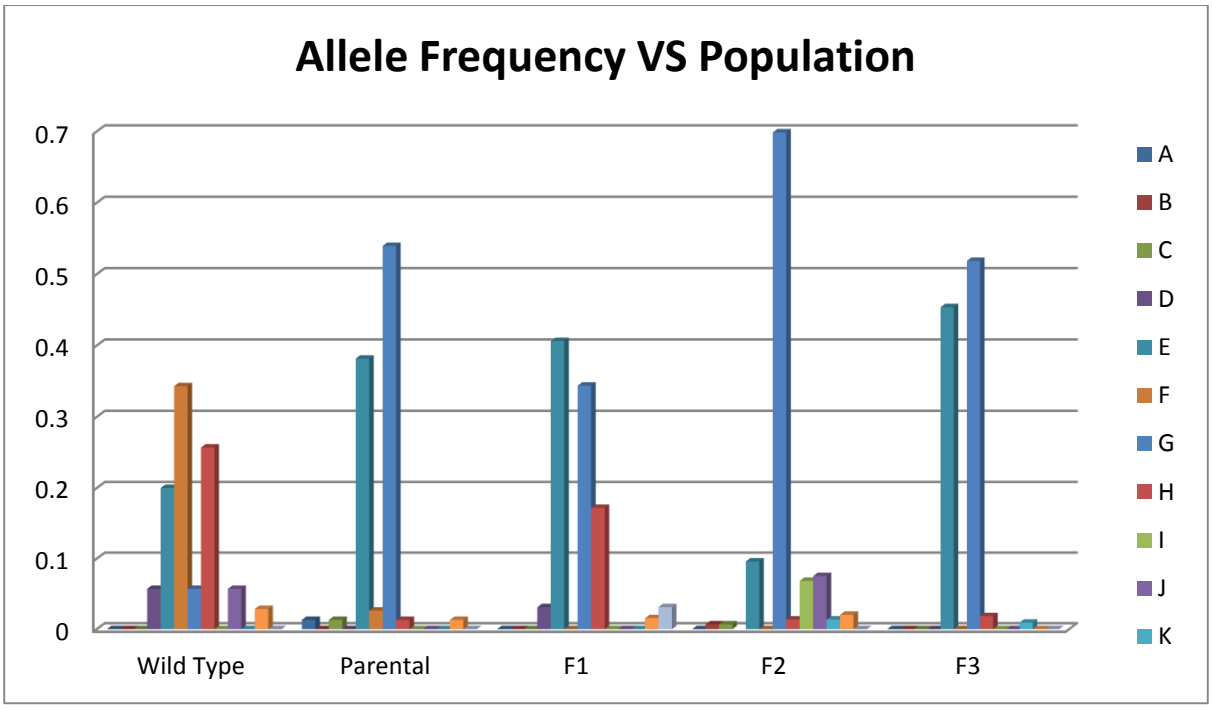
Loci: MR8



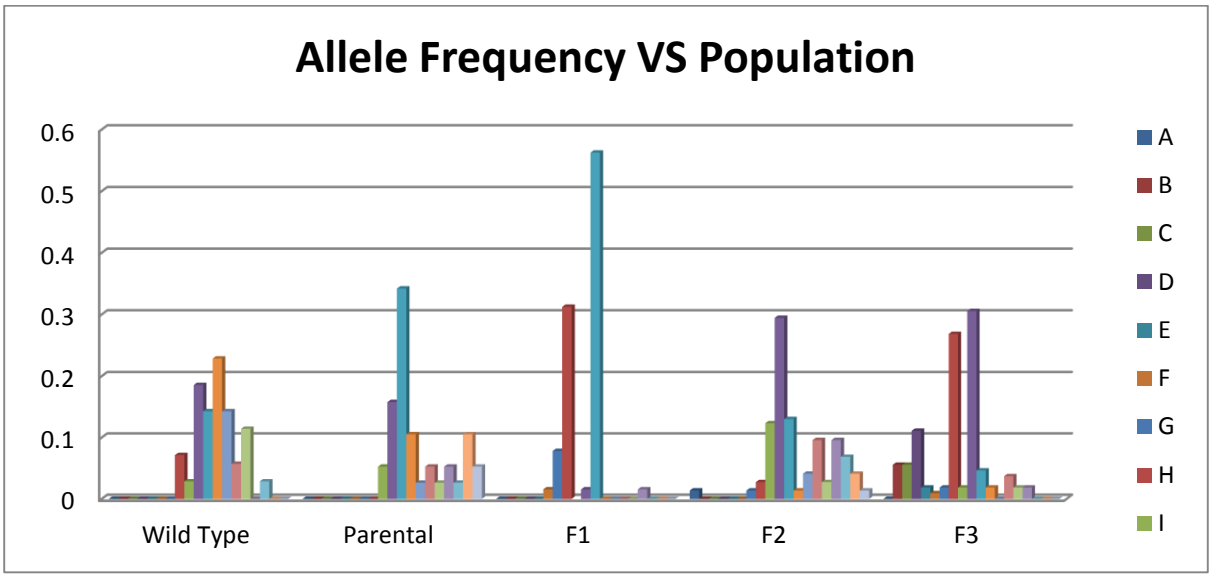
Loci: MR13



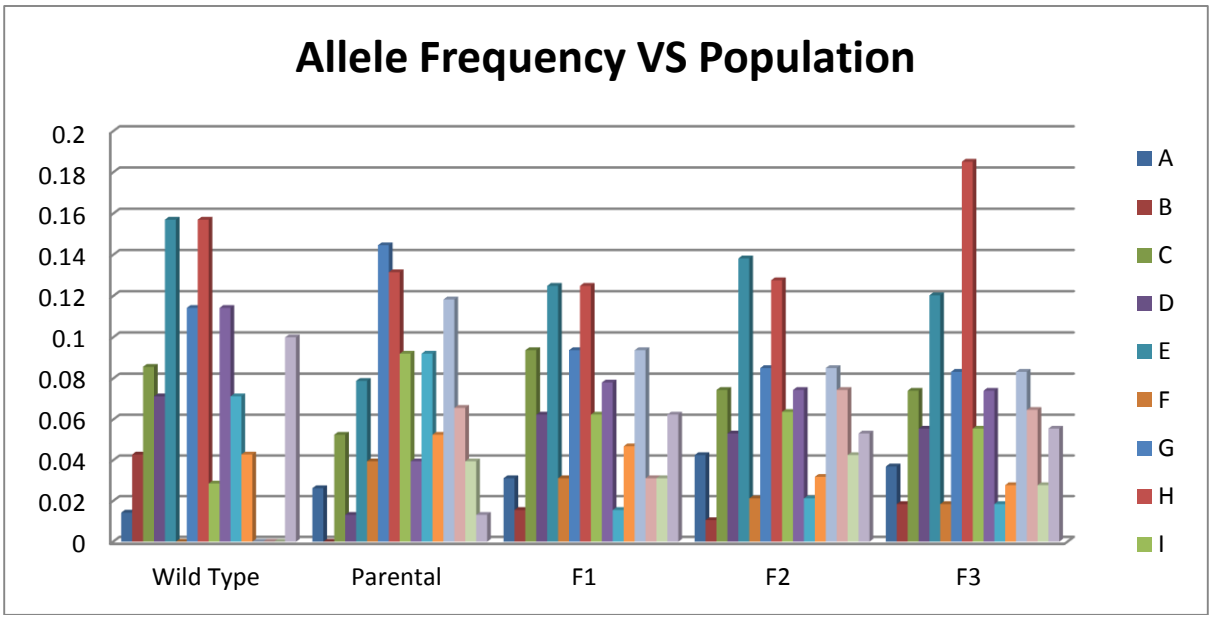
Loci: MR19



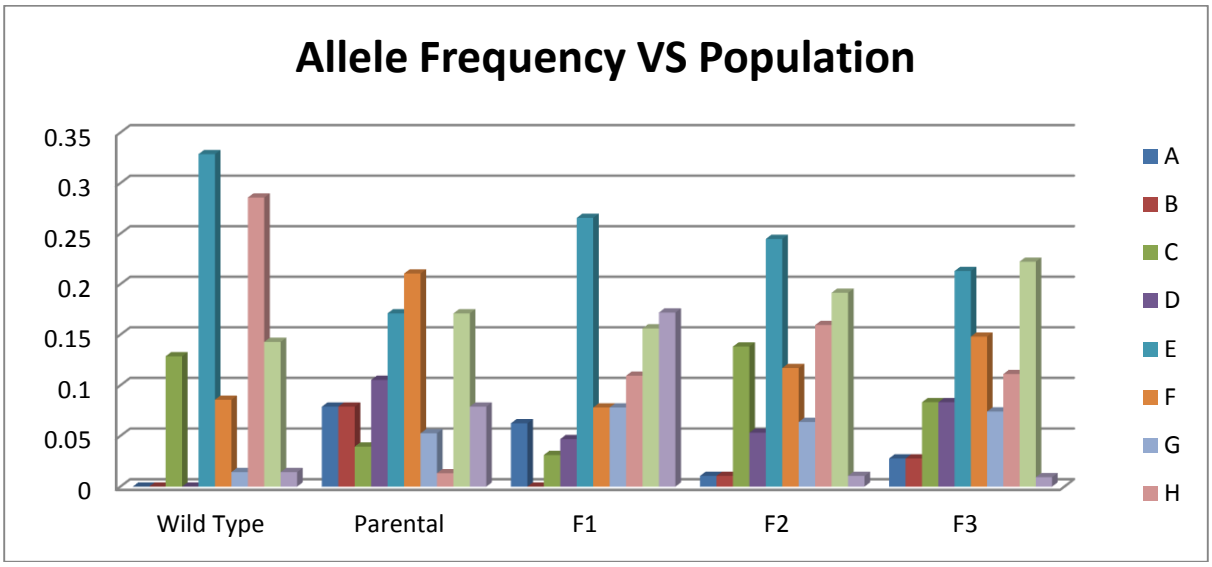
Loci: MR38



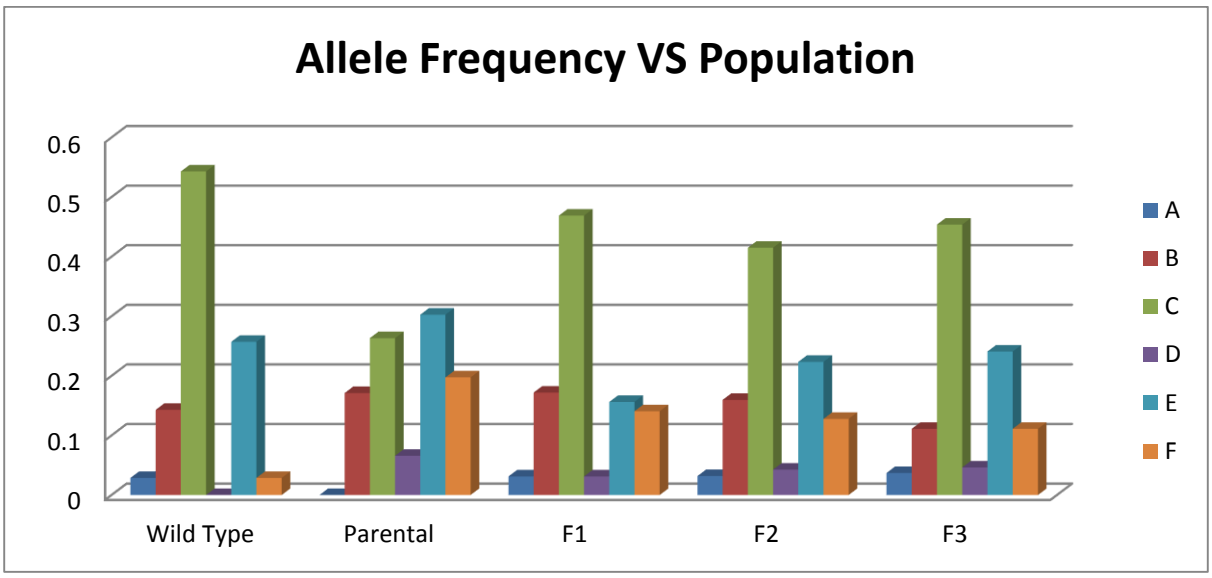
Loci: MR13



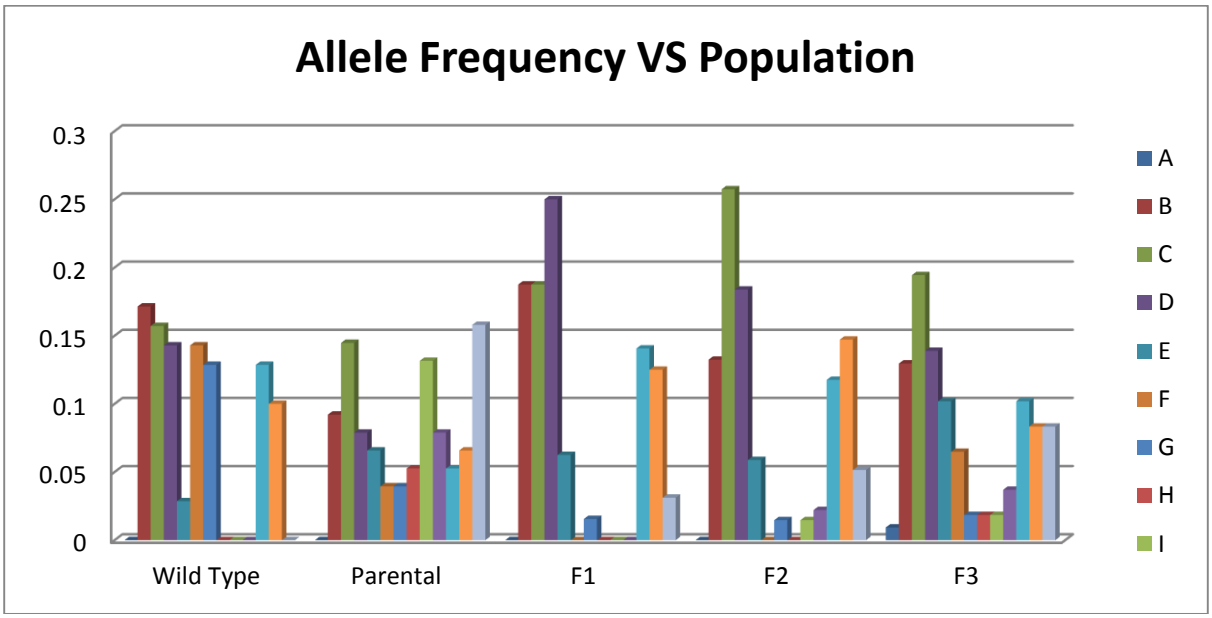
Loci: MBR5



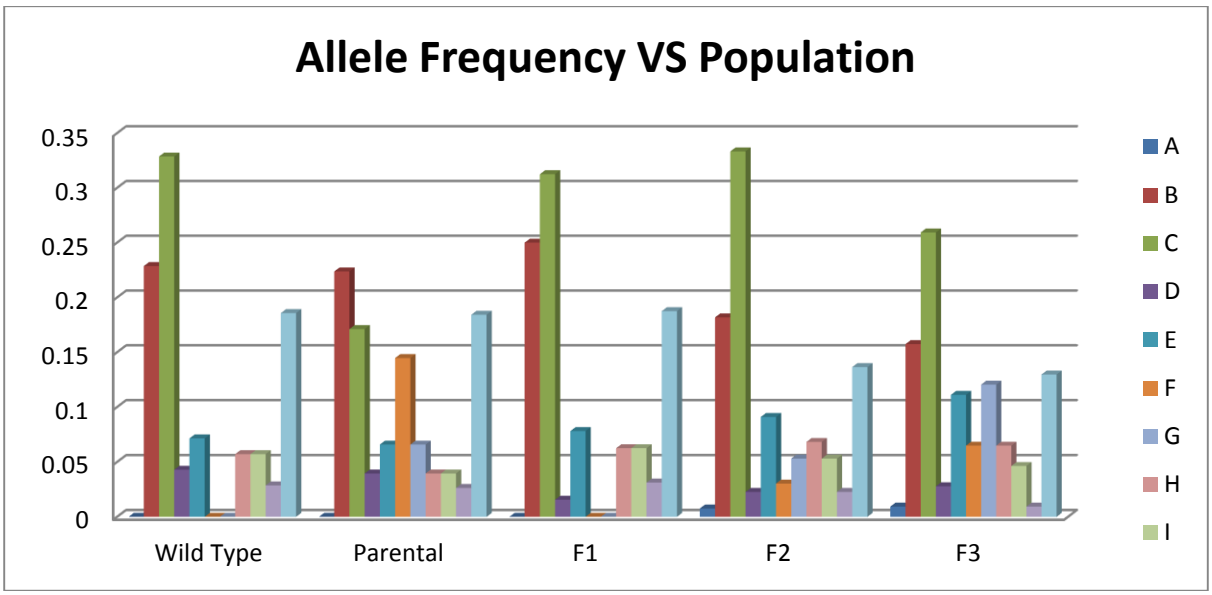
Loci: LM16



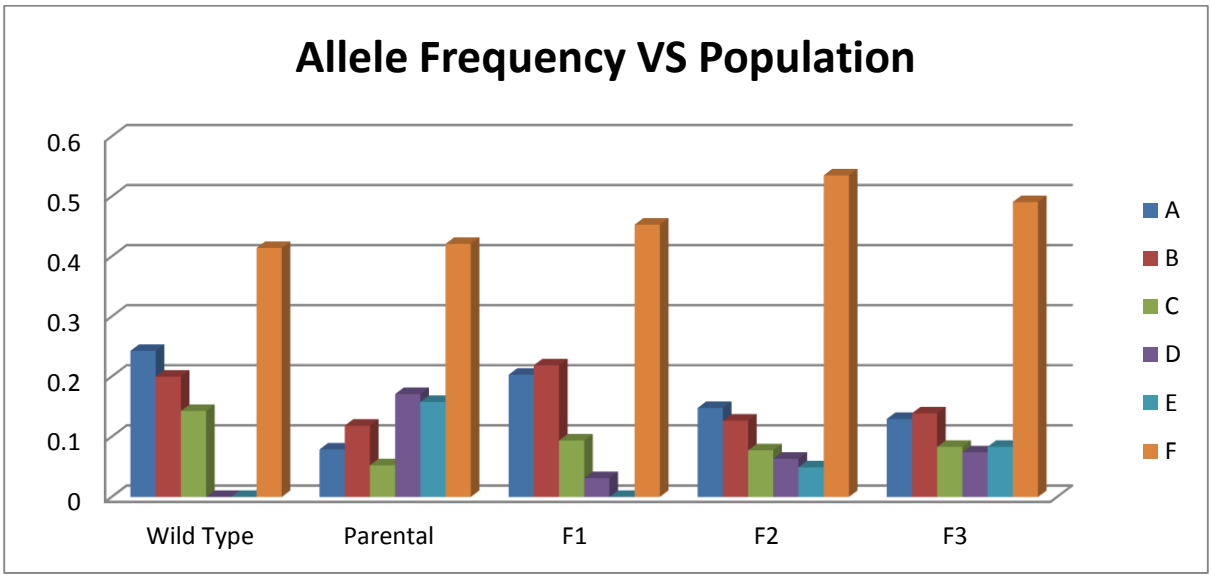
Loci: T917_7e



Loci: T917_7d



Loci: T917_9c


Allele Frequency VS Population
0.6
0.5
0.4
0.3
0.2
0.1
0
Wild Type
Parental
F1
F2
F3
A
B
C
D
E
F

Appendix G: 27 Haplotypes Sequence

#1F(2)
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#5F(2)
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCTGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCG
GCGGGAGGAGGGGACCCT

#9F(2)
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGGAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTAATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#10F(2)
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAACCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCAATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT

TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGAGACCCT

#13F(2)
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTACCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#14F(2)
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGGAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#20F(2)
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGTATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCTT
CTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCCA
CCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTTT
CCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#22F(2)

ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTAATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#26F(2)

ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGCAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCTGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCG
GCGGGAGGAGGGGACCCT

#28F(2)

ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGAGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#35F(2)

ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTGATAATTTTCTTCATAGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCTCCATCTCTAACACTTCTT
CTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCCA
CCACTAGCAGCTGGTACCGCCCACGCCGGG

GCATCAGTAGATCTAGGTATTTTTTCCCTCCACCTAGCAGGAGTTTCTTCAATCTT
AGGGGCTGTCAACTTTATTACCACAGTAATTAACATACGAGCCCCAGGAATAACT
ATAGATCGACTGCCCCTATTCGTATGAGCCGTATTTCTAACAGCCATCCTGCTTCT
TCTCTCACTACCAGTTTTAGCCGGAGCCATTACCATGCTCTTAACTGATCGAAAC
CTAAATACATCCTTTTTCGACCCAGCGGGAGGAGGGGACCCT

#506326_2_MITO_F
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCGCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#506332_5_MITO_F
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTGATAATTTTCTTCATAGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGTATAAACAACATAAGATTCTGACTCCTGCCTCCATCTCTAACACTTCTT
CTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCCA
CCACTAGCAGCTGGTACCGCCCACGCCGGGGCATCAGTAGATCTAGGTATTTTTT
CCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTAATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATGCTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGAGACCCT

#506336_7_MITO_F
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCGCTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#506337_8_MITO_R
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCAGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#508898_15_MITO_R
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTGATAATTTTCTTCATAGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCTCCATCTCTAACACTTCTT
CTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCCA
CCACTAGCAGCTGGTACCGCCCACGCCGGGGCATCAGTAGATCTAGGTATTTTTT
CCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTAATTAACATACGAGCCCCAGGAATAACTATAGATCGACTACCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATGCTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#506358_19_MITO_F
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTGATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCTGGTACCGCCCACGCCGGGGCATCAGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATGCTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#506360_23_MITO_F
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTGATAATTTTCTTCATAGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGTGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCTCCATCTCTAACACTTCTT
CTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCCA
CCACTAGCAGCTGGTACCGCCCACGCCGGGGCATCAGTAGATCTAGGTATTTTTT

CCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTAATTAACATACGAGCCCCAGGAATAACTATAGATCGACTACCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATGCTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#506372_35_MITO_F
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTGATAATTTTCTTCATAGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGTATAAACAACATAAGATTCTGACTCCTGCCTCCATCTCTAACACTTCTT
CTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCCA
CCACTAGCAGCTGGTACCGCCCACGCCGGGGCATCAGTAGATCTAGGTATTTTTT
CCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGAGCTGTCAACTTTATTACCAC
AGTAATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATGCTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGAGACCCT

#506388_47_MITO_F
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAGATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTGATAATTTTCTTCATAGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCTCCATCTCTAACACTTCTT
CTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCCA
CCACTAGCAGCTGGTACCGCCCACGCCGGGGCATCAGTAGATCTAGGTATTTTTT
CCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTAATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATGCTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#D01_2011-04-27_4F
ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCTCACGCATTCGTGATAATTTTCTTCATAGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCTCCATCTCTAACACTTCTT
CTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACCGTCTATCCA
CCACTAGCAGCTGGTACCGCCCACGCCGGGGCATCAGGAGATCTAGGTATTTTTT
CCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTAATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCCTCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATGCTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#H03_2011-04-27_16F

ATCTTGGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGA
GCAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTA
ATTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAAT
TGGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCA
TTCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTC
TTCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATC
CACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTT
TTCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCA
CAGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGT
ATGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCG
GAGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCC
AGCGGGAGGAGGGGACCCT

#M4F

ATTTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#MR45F

ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATGGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#2F

ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
GAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGAGTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT

TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGGGACCCT

#12F

ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACGTCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCCGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCA
GCGGGAGGAGGAGACCCT

#21F

ATCTTCGGAGCGTGAGCAGGCATGGTAGGTACATCACTAAGACTCTTAATTCGAG
CAGAATTAGGGCAGCCGGGCAGACTGATCGGAAATGACCAAATCTACAACGTAA
TTGTCACTGCCCACGCATTCGTAATAATTTTTTTCATGGTTATACCGATCATAATT
GGTGGTTTCGGTAATTGACTAGTACCCCTAATATTAGGGGCCCCAGACATAGCAT
TCCCACGCATAAACAACATAAGATTCTGACTCCTACCCCCATCTCTAACACTTCT
TCTCTCCAGAGGAATAGTAGAAAGAGGGGTTGGCACAGGATGAACTGTTTATCC
ACCACTAGCGGCTGGTACCGCCCACGCCGGGGCATCGGTAGATCTAGGTATTTTT
TCCCTCCACCTAGCAGGAGTTTCTTCAATCTTAGGGGCTGTCAACTTTATTACCAC
AGTGATTAACATACGAGCCCCAGGAATAACTATAGATCGACTGCCCCTATTCGTA
TGAGCCGTATTTCTAACAGCCATCCTGCTTCTTCTCTCACTACCAGTTTTAGCCGG
AGCCATTACCATACTCTTAACTGATCGAAACCTAAATACATCCTTTTTCGACCCG
GCGGGAGGAGGGGACCCT

**************THE END**************