CHAPTER-II

Genetic structure of *Panulirus homarus* collected in south peninsular coast by using Microsatellites

Genetic structure of *Panulirus homarus* collected in south peninsular coast by using Microsatellites

INTRODUCTION

Population genetics in itself can be defined as the science of how genetic variation is distributed among species, populations and individuals, and fundamentally, it is concerned with how the evolutionary forces of mutation, selection, random genetic drift and migration affect the distribution of genetic variability (Hansen, 2003). Patterns of genetic diversity or variation among populations can provide clues to the populations' life histories and degree of evolutionary isolation. Genetic differences are expressed as differences in the quantity and quality of alleles, genes, chromosomes, and gene arrangements on the chromosomes that are present within and among constituent populations (Williamson, 2001; Ciftci and Okumu, 2002). Populations stock is essential for interpretation, understanding and effective management of these populations or stocks. Genetic diversity has been measured indirectly and inferentially through controlled breeding and performances studies or by classical systematic analysis of phenotypic traits. Ecological tagging, parasite distribution, physiological and behavioural traits, morphometrics and meristics, calcified structures, cytogenetics, immunogenetics and blood pigments are among the diverse characteristics and methods used to analyze stock structure in fish populations (Ihssen *et al.,* 1981). Unfortunately the relationship between genes and their phenotypic expression is complex and often significantly interacted by environmental variables. Thus, the population geneticists mainly focused on Mendelian traits in species widely used in laboratory studies or on available pure breeds of few species. The methods used in these studies were not suitable

for wild populations and have found very limited applications in fisheries science and fisheries management (Hallerman, 2003). New methods were developed in the later twentieth century to identify, characterize, measure and analyze the genes. These are resulted from the discovery and accurate description of currently accepted model of DNA structure during the early 1950s and molecular genetics, the study of the structure, function, and dynamics of genes at the molecular level, has recognized as a powerful branch of genetics.

Initial studies in molecular genetics in the 1960's were limited with proteins such as haemoglobin and transferrin, but attention quickly turned to enzymatic proteins, allozymes (Ferguson *et al.,*1995), and allozymes was the dominant method employed during 1960s and the beginning of 1980s (Williamson, 2001). The development of DNA amplification using the PCR (Polymerase Chain Reaction) technique has opened up the possibility of examining genetic changes in fish populations over the past 10 years (Ferguson *et al.,* 1995). Today, many molecular methods are available for studying various aspects of wild populations, captive brood stocks and interactions between wild and cultured stocks of fish and other aquatic species.

A microsatellite is a simple DNA sequence that is repeated several times at various points in the organism's DNA. Such repeats are highly variable enabling that location (polymorphic locus or loci) to be tagged or used as a marker. Microsatellites have much more information than allozymes and mitochondrial DNA, yet offer the same advantages of analysis. Technical expertise required for detection and scoring/analysis once the polymorphic loci identified is similar for all the methods. Microsatellites are thought to occur approximately once every 10 kbp, while minisatellite loci occur once every 1500 kbp in fish

species (Wright, 1995), which suggests that microsatellites may be more useful for genome mapping studies (O'Connell and Wright, 1997). They are one of a class of highly variable, non-coding and considered to be selectively neutral, allowing for the assumption that the estimated amount of sequence divergence between units of interest is directly proportional to the length of time since separation (Brown and Epifanio, 2003). Microsatellites are co-dominant, inherited in a Mendelian fashion and tandem arrays of very short repeating motifs of 2-8 DNA bases that can be repeated upto ~100 times at a locus. They are among the fastest evolving genetic markers, with 10^{-3}-10^{-4} mutations/generation (Goldstein *et al.,* 1999).

Their high polymorphism, and PCR based analysis has made them one of the most popular genetic markers (Wright and Bentzen 1995). With current molecular methods it is feasible to score microsatellite length polymorphisms in large numbers of individuals for genetic analyses within and between populations. Some microsatellite loci have very high numbers of alleles per locus (>20), making them very useful for applications such as parent-offspring identification in mixed populations, while others have lower numbers of alleles and may be more suited for population genetics and phylogeny (O'Connell and Wright, 1997; Estoup and Angers, 1998). Primers developed for one species will often cross-amplify microsatellite loci in closely related species (Estoup and Angers, 1998).

Microsatellite markers have a number of advantages over other molecular markers and have gradually replaced allozymes and mtDNA. Microsatellite loci are typically short, this makes it easy to amplify the loci using PCR, and the amplified products can subsequently be analysed on either "manual" sequencing gels or automated sequencing.

Microsatellites are relatively easy to isolate compared with minisatellites, sample DNA can be isolated quickly because labour-intensive phenol-chloroform steps can generally be eliminated in favour of a simpler form of DNA extraction. The much higher variability at microsatellites results in increased power for a number of applications (Luikart and England, 1999). Only small amounts of tissue are required for typing microsatellites and these markers can be assayed using non-lethal fin clips and archived scale samples, facilitating retrospective analyses and the study of depleted populations (McConnell *et al.,* 1995). Moreover, there is potential for significant increases in the number of samples that can be genotyped in a day using automated fluorescent sequencers. For applications where a large number of loci are required, such as genome mapping or identification of Quantitative Trait Loci (QTL), microsatellites offer a powerful alternative to other marker systems.

So, the aim of the present study was to test the genetic variation and phylogenitic structures of spiny lobsters represent different biogeographical range and genetic stocks linked by gene flow under a model of panmixis and isolation by distance.

MATERIALS AND METHODS

Lobster Samples and DNA Extraction

Five different chosen collection sites (as per in the chapter I) and adult specimens purchased from the landing sites from local fisherman were transferred to the laboratory alive or frozen. The tissue samples were dissected and transferred in ethanol. Then the samples were minced and were homogenized in 1.5 ml microcentrifuge tubes using a Teflon pestle. Crude DNA was extracted using a DNA extraction kit (Bio-Geno and Tissue DNA Isolation Kit, Amersham Bioscience).

Microsatellite primers

Five primers were used synthesized and used in this study and their repeatability were given in the table: 3.

Table: 3. Microsatellite primer sequences.

Code Name	Repeat Unit	Sequence
PH1	(TG)7 (GT)10	GGCGTGATGTCCGATTTTACTG AGTAATGCCATCGCCGCTGTAGG
PH2	(AAG)24 (AG)11	TGCTAAGCCTTCTGTCTGATAAGCTA ATTCTCGGAGGTAACCAGACTTTT
PH3	(GA)10	TACAACAAAGGGTCACTACAC CCATAGAGCTAACTAACGCAGG
PH4	(GT)6	TGTCGTCGTAACTACCCGCTG ATACATCGTCCACTGCTCCAC
PH5	(CT)30	ACTGTCCGGTCAGGTTGAGA TCAGTCAAGAGATTGGGGAG

PCR Amplification

PCR amplification was performed in a Mastercycler gradient (Eppendorf) thermal cycler and the mixtures contains (25 ng of genomic DNA, 0.5 µM forward and reverse primers, 16 mM (NH4)2SO4, 50 mM Tris-HCl pH 8.8 (at 25 °C), 0.1% Tween 20, 0.2 mM of each dNTP, 3 mM magnesium chloride and 1 U Taq Polymerase) at the volume of 50 µl. The amplification programme consisted of an initial denaturation at 94 °C for 1 min; 3 cycles of denaturation at 95 °C for 1 min, annealing at 32 °C for 1 min, and extension at 72 °C for 30 s. This was followed by 30 cycles of denaturation at 94 °C for 20 s, annealing at 48-58 °C for 15 s, and extension at 72 °C for 12 s; and a final extension at 72°C for 10 min. In order to reduce genotyping error, the PCR products, were treated with T4 DNA polymerase to

remove the extra bases added by the Taq polymerase. After PCR amplification, 0.4 U T4 polymerase (Promega) was added to the PCR products and the mixture was incubated at 37 °C for 30 min.

Agarose Gel Electrophoresis of PCR Amplified Products

1g of agarose and 50mL of 0.5xTBE were added in 250mL conical flask and stirred well. Then it was kept in microwave oven for a minute and added 1µL of ethidium bromide (10mg/mL) with the solution and mixed thoroughly. The gel was poured at the gel setup without bubbles at 60 ^{0}C and the combs were inserted. It was kept undisturbed for 30 minutes to allow the gel to solidify.

Running the gel:

0.5X of TBE buffer was poured into the gel tank until the gel is submerged in a depth of 2mm. The Samples from Microsatellite PCR were mixed with 5 microlitre of loading buffer and each sample was loaded into adjacent wells along with HindIII lambda DNA digest marker. Initially run the gel at 5V for 5min before increasing to 100V and it was allowed to run until the tracking dye reaches the end of the gel. Finally the gel was taken from the setup and documented with BioRAD Gel documentation system.

Phylogenetic tree construction

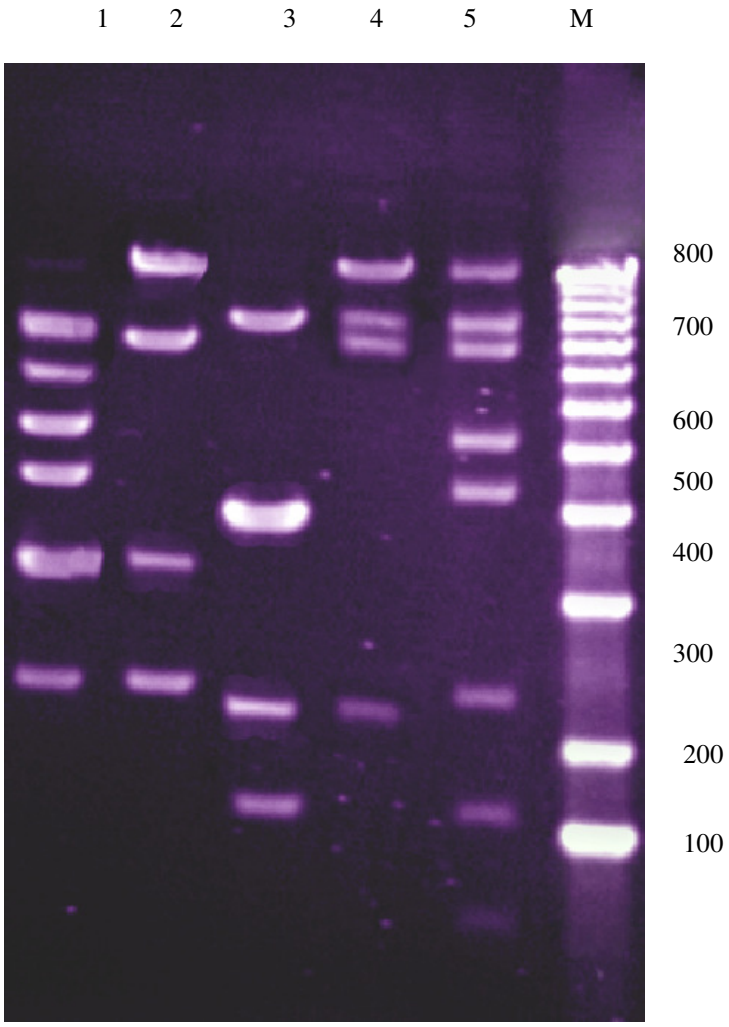
Genetic variation within each of the 5 populations Distance matrix can be calculated by using Jacard NJ Method. The genetic distance was used to construct a UPGMA dendrogram by using Neighbour Joining Algorithm.

RESULTS

The electropherogram summarise the positive and negative amplifications obtained with five different microsatellite primers (PH1, PH2, PH3, PH4 and PH5) for all five collection spots studied. The microsatellite markers PH 2 and PH5 were successful with all five collection sites of *P. homarus*. Distance Matrix produced by Jaccard NJ method for PH2 and PH5 primers amplified product were given in table: 4 and 5 the consequent phylogenetic tree also given in fig: 9 and fig: 11.

Based on the results the southern and west coast specimens has more close with each other and the distance between these animals are so close than others. In the mean time the east coast specimens such as Chinnamuttam and Tuticorin has very close with each other with significant similarity matrix. . Least bands are produced in Chinnamuttom, Cochin and Tuticorin. Highest numbers of fragments were produced in Vizhinjam populations. Allele polymorphic bands were produced in the Tuticorin and Vizhinjam populations. Primer (PH2) Microsatellite distance matrix were generated by using Jaccard NJ Method. The highest genetic distances were found Tuticorin population with Muttom and Chinnamuttom populations (0.48684).

Fig: 8. Banding Patterns of PH2 Microsatellite primer



1 – Muttom

2 – Chinnamuttom

3 – Cochin

4 – Tuticorin

5 – Vizhinjam

M – 100bp marker

Table: 4. PH2 Microsatellite primer Distance Matrix (Jaccard NJ method)

	'Muttom '	'Chinnamuttom '	'Cochin '	'Tuticorin'	'Vizhijam '
'Muttom '	0	0.48276	0.42759	0.48684	0.47651
'Chinnamuttom '	0.48276	0	0.4698	0.48684	0.47651
'Cochin '	0.42759	0.4698	0	0.48408	0.46405
'Tuticorin'	0.48684	0.48684	0.48408	0	0.49045
'Vizhijam '	0.47651	0.47651	0.46405	0.49045	0

Fig: 9. PH2 Primer Resolved Phylogenetic tree constructed using Neighbour Joining Algorithm

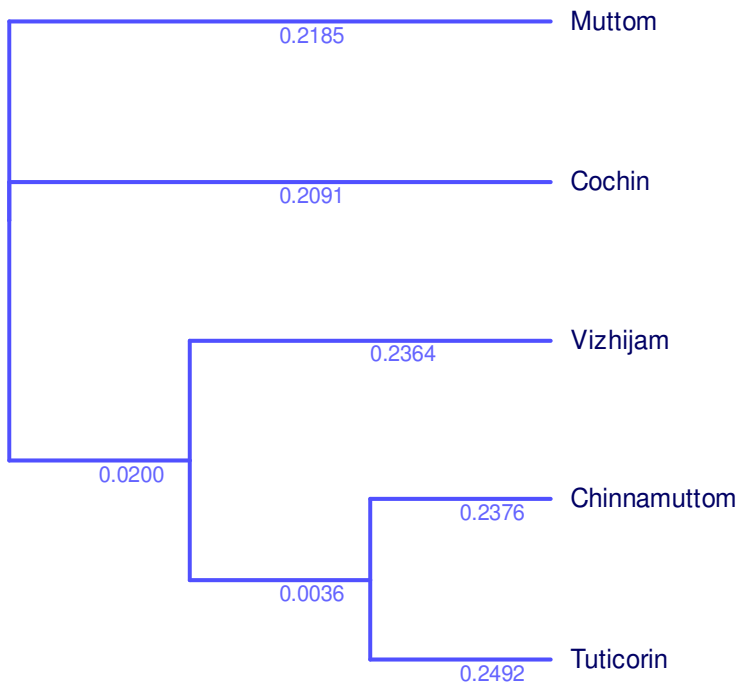
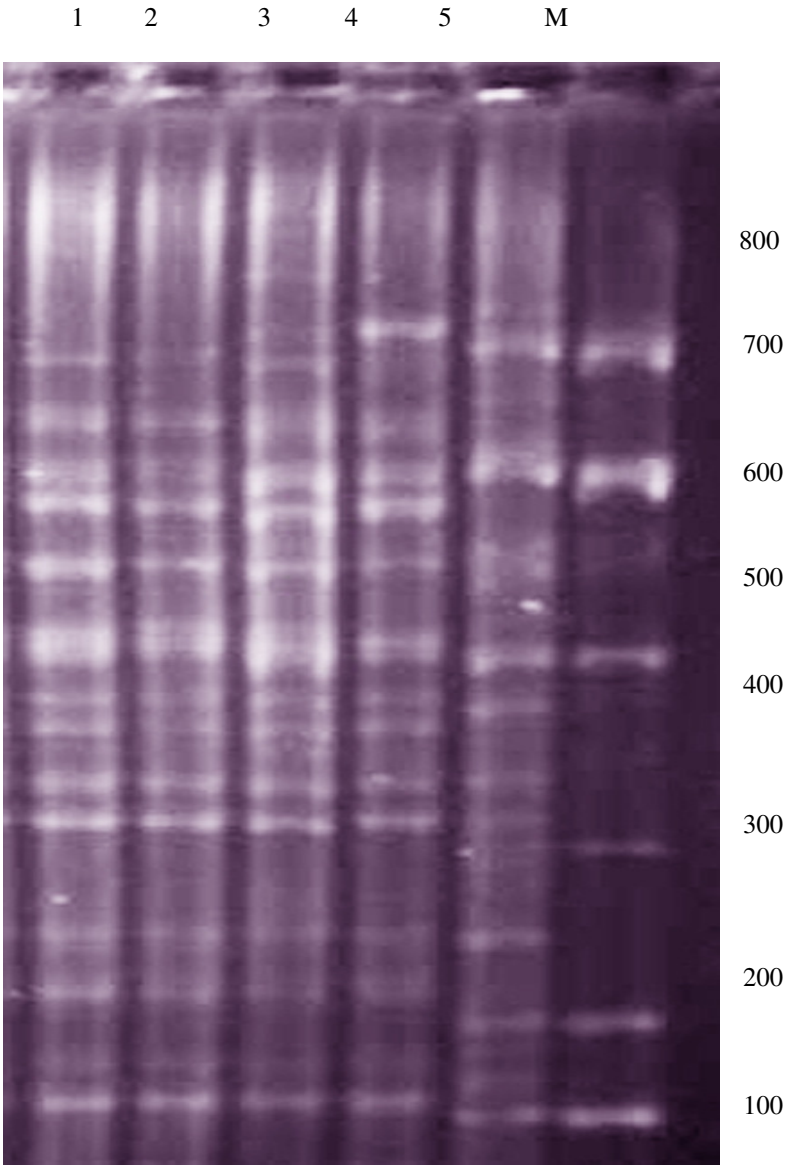


Fig: 10. Band Patterns of PH5 Microsatellite primer



1 – Muttom

2 – Chinnamuttom

3 – Cochin

4 – Tuticorin

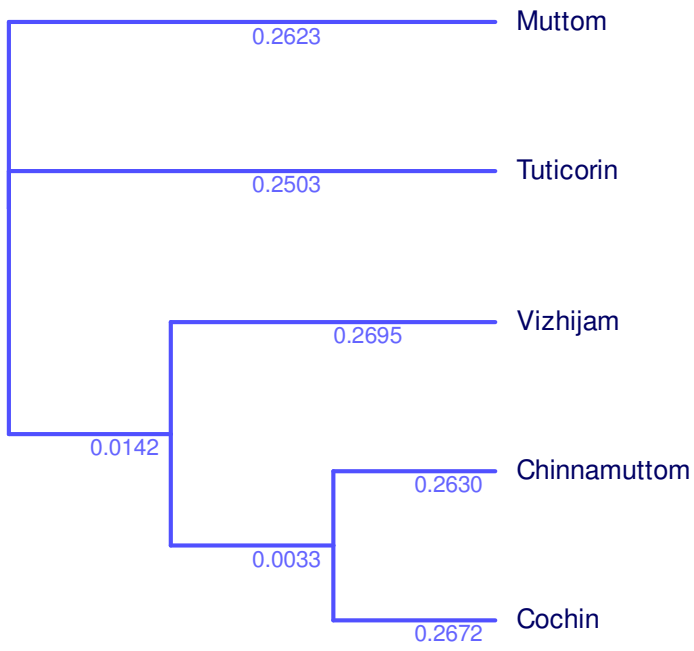
5 – Vizhinjam

M – 100bp marker

Table : 5. PH5 Microsatellite primer Distance Matrix (Jaccard NJ method)

	'Muttom '	'Chinnamuttom '	'Cochin '	'Tuticorin '	'Vizhijam '
'Muttom '	0	0.53628	0.54829	0.51266	0.5512
'Chinnamuttom '	0.53628	0	0.5302	0.53156	0.53871
'Cochin '	0.54829	0.5302	0	0.53947	0.53698
'Tuticorin '	0.51266	0.53156	0.53947	0	0.52885
'Vizhijam '	0.5512	0.53871	0.53698	0.52885	0

Fig: 11. PH5 primer resolved Phylogenetic tree constructed using Neighbour Joining Algorithm



DISCUSSION

The results above indicate that the genetic markers developed for lobster will be useful for a wide range of applications. Among these five primers (table: 3) two primers were produced amplified fragments in all populations. The primer PH2 (TGCTAAGCCTTCTGTCTGATAAGCTA & ATTCTCGGAGGTAACCAGACTTTT), Primer PH5 CTGTCCGGTCAGGTTGAGA & TCAGTCAAGAGATTGGGGAG) were chosen for further analysis. The primer PH2 produced 4 to 7 DNA fragments.

The primer PH5 produced 10 to 12 DNA fragments. Allele polymorphic bands were produced in all populations. Primer (PH5) Microsatellite distance matrix were generated by using Jaccard NJ Method. The genetic distance was found between Cochin and Muttom populations are 0.54829. Lowest genetic distance was found between Muttom and Tutucorin populations (0.51266). Phylogenic tree constructed using Neighbour joining algorithm has three clusters. First cluster contains Muttom populations, second cluster contains Tuticorin populations and third cluster contains Vizhinjam, Chinnamuttom and Cochin populations.

Earlier finding displayed that American lobster has much higher levels of genetic variation than European lobster is consistent with the results of Tam and Kornfield (1996), providing further evidence that American lobster has had a much greater long -term effective population size than European lobster. The clearly distinguishable alleles at all of the microsatellite loci characterised infers that they can be sized with minimal error in

population analyses and highlights the benefit of selecting microsatellite loci with longer repeat unit lengths. Loci composed of dinucleotide repeat motifs, while more frequent in the genome than loci with longer repeat motifs are more prone to PCR-induced stutter (Edwards *et al.,* 1991). This can make the differentiation between homozygotes from heterozygotes with alleles differing by one repeat length difficult.

The detection of alleles that differ by 1 or 2 bp underscores the importance of having a precise means of assessing allele size and the need for sample redundancy to ensure consistent allele designation. Although multiplexing can produce data more rapidly, it should be used cautiously for analyses of individuals of unknown potential genotypes (e.g., population analyses) and only after thorough optimization and genotype verification. Amplifying more than one locus in a PCR invariably leads to greater background and often creates artefact bands (Donini *et al.,* 1998). An alternative for population analyses that still achieves high throughput without necessitating as intensive an optimization process or compromising data quality is to amplify each locus separately and then combining 2–3 loci before electrophoresis. The linked loci (Ham53 and Ham54) have great potential for identifying genetic admixtures (Estoup *et al.,* 1998). Insignificant linkage disequilibrium in the Crowell Basin samples may reflect the mixing of several subpopulations or a sufficiently large effective population size being maintained longs enough for multiple lineages to develop through the accumulation of mutations at both loci. The possibility of lab error being a contributing factor can be rejected as the genotypes were verified with an initial 12% redundancy in these samples (i.e., 6 of the 50 American samples were randomly duplicated), as well as by regenotyping 32% of the samples at both loci; identical genotypes were

obtained in all cases. This finding highlights the importance of conducting linkage and segregation analyses with novel loci; had no a prior knowledge that these loci were physically linked, the use of population data alone would have resulted in the erroneous treatment of these as independent loci. The significant linkage disequilibrium detected in European lobster is consistent with a lower effective population size in that species.

Previous genotyping of embryos from berried females either only inferred multiple paternities based on unexpected allele frequencies at one allozyme locus (Nelson and Hedgecock 1977) or found no evidence of multiple paternities (Tam and Kornfield 1996). The detection of multiple paternity in the progeny of one of the two females (n = 48 for each female) used to test the inheritance demonstrates the utility of these loci to address questions about mating patterns. The four-locus multiplex developed provided the ability to rapidly assess parentage; the occurrence of additional male fertilization was detected at three of the four loci in two embryos and two of the four in the third embryo. The combined paternity exclusion probability of 97.8% can be increased to 99.6% with the addition of only one more locus (Ham6 or Ham53). These markers should thus prove highly amenable to rapidly address questions relating to mating structure and individual reproductive success.

Tri and tetranucleotide microsatellites have been demonstrated to be highly polymorphic and are stably inherited in the human genome (28 – TS 6%). They are becoming increasingly popular markers because the allele differences are easier to distinguish than those of dinucleotide repeats and less stutter bands have been observed in the amplification products. In this study, the abundance of different tri- and tetranucleotide

repeats in the *P. homarus* microsatellite were examined. *Xu et al.,* have found various sequence types of tri- and tetranucleotide microsatellites in *P. monodon* by direct sequencing of recombinant clones without probe screening. The simple repeat (GATA)n was first identified and isolated from snake satellite DNA and subsequently it has been found throughout the eukaryotic genomes. The low frequency of positive clones found in the screening suggests that enrichment of microsatellite sequences prior to screening would be useful.