# CHAPTER 3

# GENETIC VARIATION BETWEEN POPULATIONS OF *L. russelli* IN GULF OF THAILAND AND ANDAMAN SEA

**Note: The bulk of the data in this chapter has been published as Klagnurak *et al.,* 2012**

## 3.1 INTRODUCTION

Populations of marine fishes have traditionally been regarded as "open" systems, i.e. local populations were expected to be replenished largely by larvae derived from elsewhere (Mora and Sale, 2002), although this view has been challenged by recent advances in larval research (Planes *et al*., 2009). Knowledge about population genetic structure and replenishment of marine organisms is thus important for the conservation and management of marine resources (Ovenden and Street, 2003; Robertson, 2000; Zhang *et al*., 2004). Recently, increasingly frequent reports indicate that some Caribbean lutjanids are threatened by overfishing, including such species as *Lutjanus campechanus* and *L. griseus, L. synagris* (Gillig *et al*., 2001; Renshaw *et al*., 2007). It is not currently known whether lutjanid species important for food security are threatened in the South-East Asian region, however; this is mostly because of a critical paucity in population-level studies. Many lutjanids in this region are heavily targeted and have vulnerable life history characteristics (slow growing, late maturing, near shore juvenile habitats allowing easy targeting). Consequently, increasing fishing pressure associated with increasing coastal populations increases the risk of overexploitation and population collapse.

In Thailand, *Lutjanus russelli* (Bleeker, 1849) is an important component of the local fisheries of the Andaman coast of Thailand and the Gulf of Thailand. Although it is not especially commercially valuable, it is important in terms of food security for local Thai fishers and the millions of coastal villagers who rely on marine ecosystem goods and services for their livelihood. Several authors (e.g. Allen, 1985; Satapoomin, 2011; Chapter 2, *this document*) have reported consistently different color

morphs in the Indian Ocean and the Pacific Ocean populations of *L. russelli*, hinting that the species may be partitioned into discrete stocks (or even races or subspecies). It is important, then, to establish estimates of population connectivity and stock resilience in order to appropriately manage local fisheries, since depletion of local stocks represents a serious threat to artisanal fishers' food security in Thailand.

Many molecular techniques for analyzing population structure have been used for population genetic studies within Lutjaninae, such as allozyme, mitochondrial DNA and microsatellites (Ovenden and Street, 2003; Pruett *et al*., 2005; Salini *et al*., 2006 and Zhang *et al*., 2006). Allozymes studies examine the frequency with which certain proteins are expressed in an organism, and were the cutting edge of molecular biogeography prior to the development of modern techniques for population genetics. Nowadays, allozymes are considered to be inappropriate markers for population genetic study because the expression of proteins depends on many (often environmental) factors that hinder accurate assessment of population structure (Baker, 2000). Mitochondrial DNA (mtDNA) has the property of maternal inheritance (mitochondria are possessed only by female-originated gametes) which may affect population study. Microsatellites have a high mutation rate (Hoelzel, 1992) and have been accepted as a major source of genetic variation (Baker, 2000), which is appropriate to investigate population-level genetic variation. Moreover, this marker does not require a large quantity of DNA and it has low expenditure. It is apparent, then that microsatellites present a suitable and useful tool for investigating genetic diversity and population genetics in the present context (Piyachokanakun, 2009; Selkoe and Toonen, 2006).

The fact that Peninsular Thailand's coasts front on two oceans means that the Gulf of Thailand (east coast) is the end of western Pacific Ocean, and the Andaman Sea (west coast) forms the eastern-most margin of the Indian Ocean. Superficially, the biotas of the two coasts should be not completely distinct, since the existence of a connection via a persistent northwesterly flow through the Straits of Malacca should allow dispersal (Kimura *et al*., 2009; Satapoomin, 2011). However, the water current from Malacca Straits does not flow throughout the entire Andaman Sea

(Wyrtki, 1961 (in Rizal *et al*., 2010)). The current from South China Sea would potentially influence only reefs in the southern part of the Andaman Sea; organisms in northern part of the Andaman Sea are likely to be influenced more by water currents from the northern Bay of Bengal. Surface water currents are one of the factors that control the dispersal of marine fish larvae and potentially inhabit genetic homogeneity between populations, particularly those with passively-dispersed larvae (Shulman and Bermingham, 1995). The convergence zone where the two opposing currents meet may represent a barrier to dispersal between northern and southern stocks and also the gentle current flowing westward from the South China Sea may effectually prevent eastward connectivity between the Gulf of Thailand and the Andaman Sea. The life history of *L. russelli* includes inshore-offshore migration, however; the stocks of this fish should therefore be able to circumvent any dispersal barrier caused by surface currents. Knowledge about the population structure of *L. russelli* may enhance understanding of the degree of larval migration between biogeographic provinces. It may also highlight differences in population genetics that imply or expose the existence of unsuspected geographical barriers.

This study focus on examination of differences in the population genetic structure of *L. russelli* between the Gulf of Thailand and the Andaman Sea, and also between northern and southern populations within the Andaman Sea. Genomic resources such as genome or transcriptome sequences or published microsatellite libraries make possible detailed studies of genetic connectivity. By looking at the variation of microsatellites loci in populations, inferences can be made about population structures and differences, genetic drift, the effects of population fragmentation and interaction of different populations. Microsatellites are also useful in identification of new and incipient subpopulations within a widely-dispersed metapopulation, partly because they tend to occur in non-coding regions, and are thus not subject to environmental culling. This study was undertaken using genetic microsatellite markers derived from overseas research into this species.

## 3.2 OBJECTIVE

This study aimed to establish whether *L. russelli* populations in the Gulf of Thailand, the northern and the southern Andaman Sea represent genetic stocks or components of a larger metapopulation. There are further consequences of stock isolation due to potential dispersal barriers for fisheries management and conservation of marine resources.

## 3.3 RESEARCH QUESTIONS

1) Are the populations of *L. russelli* in the Andaman Sea biogeographical regions genetically distinct from those in the Gulf of Thailand? Is some part of Andaman Sea similar exists with Gulf of Thailand?

2) Are the populations of *L. russelli* between northern and the southern Andaman Sea genetically different?

## 3.4 MATERIAL AND METHOD

### 3.4.1. Study site

The study area at each site in both the Andaman Sea and Gulf of Thailand (Figure16) consists of 3 distinct habitats: offshore coral reefs, inshore coral reefs and seagrass or estuary.

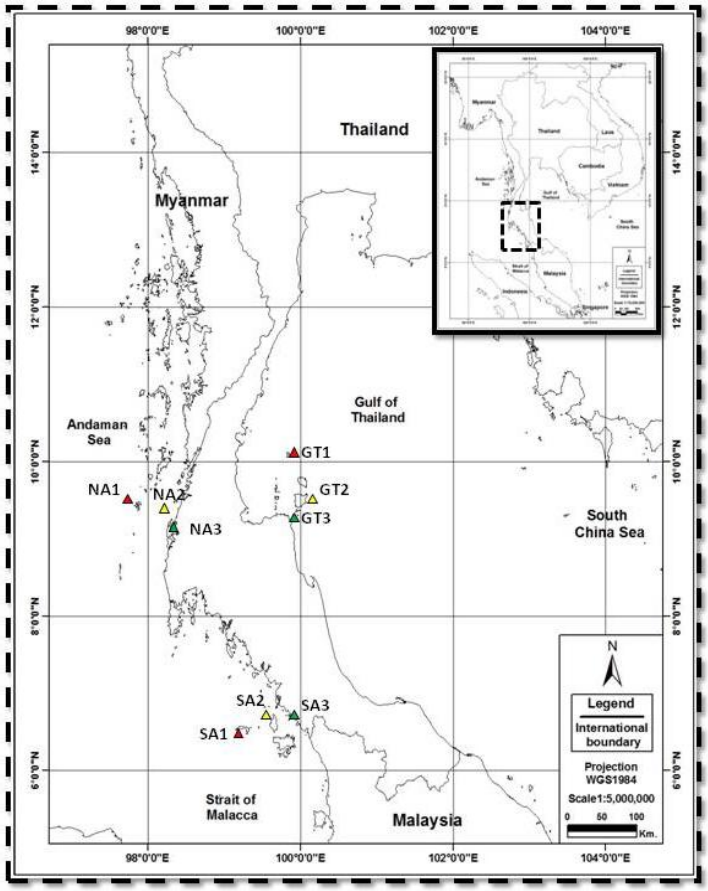


Figure 16 Study sites; NA is northern Andaman Sea, SA is southern Andaman Sea and GT is Gulf of Thailand. Number was represent type of collected habitat; 1 is offshore coral reef, 2 is inshore coral reef and 3 is seagrass bed or mangrove.

This fish are a common component of assemblages in many types of habitat. The samples for this study were collected from many habitats in order to obtain representative and comprehensive population samples. Moreover, by sampling thus, the

genetic variation in each habitat would be established in a way convenient for revealing connectivity between habitats.

The study sites in the Andaman sea are separated into 2 areas: northern part of the Andaman (NA) and southern part of the Andaman (SA); each area consisted of 3 sites reflecting the habitats described above. The six study sites in the Andaman Sea included:

1) Offshore coral reefs at Surin Island, Phangnga province (NA1).

2) Inshore coral reefs at Phayam Island, Ranong province (NA2).

3) Seagrass beds or mangrove at Ra Island, Phangnga province (NA3).

4) Offshore coral reefs at Adang – Rawi Island, Satun province (SA1).

5) Inshore coral reefs at Li De Island, Satun province (SA2).

6) Estuary near Pak Ba Ra port, Satun province (SA3).

Similarly, the sites in the Gulf of Thailand (GT) included:

1) Offshore coral reefs at Tao Island, Surat Thani province. (GT1)

2) Inshore coral reefs at Samui Island, Surat Thani province. (GT2)

3) Seagrass beds at Tha Rai Island, Nakhon Si Thammarat province (GT3)

### 3.4.2. Sample collection and preservation

Sample was collected from NA, SA and GT from May 2011 to March in 2012.

The target fishes (*L. russelli*) obtained from each habitat were represent in all size classes, the standard length of samples obtained for this study were between 8.7 – 32.9 cm. The samples were mostly collected by members of the local artisanal fishery, and some fishes were bought from local fish market. A population sample of 77 fishes from the GT, 106 samples from the NA, and 117 samples from the SA were obtained.

Pectoral or dorsal fins of each sample were immediately clipped and preserved in 100% acetone (Fukatsu, 1999) upon capture or purchase.

### 3.4.3. Sample analyses

The genetic diversity and population genetic structure of *L. russelli* were investigated using microsatellite markers examined by a polyacrylamide gel electrophoresis technique.

#### 1) DNA extraction

The tissue was homogenized in TNES buffer (50 mM tris(hydroxymethyl) aminomethane (Tris-HCl) (pH7.4), 100 mM Ethylenediaminetetraacetic acid ( EDTA) ((pH8.0), 400 mM Sodium chloride (NaCl) and 0.5% sodium dodecyl sulphate (SDS)). Genomic DNA was extracted using Phenol-Chloroform (Sambrook and Russell, 2001), before being precipitated in 100% iso-propanal, washed in 70% Ethanol, dried and resuspended in deionized water (DI) (Appendix E). The quantity and quality of DNA were investigated using a nanodrop spectrophotometer, and agarose gel electrophoresis, respectively.

#### 2) The polymerase chain reaction amplification

High quality genomic DNA extracted from the specimens and subsequently purified was subjected to polymerase chain reaction (PCR) amplification for microsatellite markers. Microsatellites are simple sequence tandem repeats, on each side of which are flanking regions that consist of "unordered" DNA. The flanking regions are critical because the probability of finding an identical stretch of unordered DNA 30-50 base pairs (bp) long more than once in the genome becomes vanishingly small and therefore allows development of locus-specific primers to amplify the microsatellites with PCR. The primers for PCR are sequences from these unique flanking regions. The

primers used in this study the amplification of loci for *L. russelli* were developed by Guo *et al*. (2007b) for a study in the South China Sea. Guo *et al*. (2007b) reported 43 microsatellite primer pairs for *L. russelli*. Of these 43, the ten loci that were in Hardy-Weinberg equilibrium, and presented high observed heterozygosity ($H_O$) and expected heterozygosity ($H_E$) in *Guo et al.* (2007b) study were chosen for this study. However, only 9 loci were successfully amplified in the course of this study. These included microsatellite loci designated Lru001, Lru002, Lru003, Lru010, Lru011, Lru014, Lru025, Lru030 and Lru043 (Appendix F) by Guo *et al.* (2007b).

In preparing samples of PCR-amplified DNA, it was discovered that the optimal annealing temperature for Lru014, Lru043 were 56 ºC, for Lru001, Lru002 were 58 ºC, for Lru003, Lru010, Lru011 were 60 ºC, for Lru030 was 62 ºC and for Lru025 was 66 ºC.

PCRs were performed on aliquots of 20microlitres (µL), containing at least 100nanograms (ng) of genomic DNA, 1µL of each primer (10pmol/ µL), 2 µL of 10×PCR buffers (10 mM Tris-HCl (pH 9.0)),10 mM KCl, 10 mM($(NH4)_2SO_4$, 2 mM $MgSO_4$), 0.4 µL each dNTPs (10ng/ µL) and 1 µL of I Taq polymerase (1Unit/ µL). PCR condition began with an initial denaturation step at 95°C for 3 minutes, followed by 30 cycles at 94°C for 30 seconds; 30 seconds at the annealing temperature (Ta); 72°C for 1 minute, and a final extension step at 72°C for 10 minutes.

### 3) Polyacrylamide gel electrophoresis (PAGE) examination

I examined PCR products using the technique of polyacrylamide gel electrophoresis and visualization by silver staining (Caetano-Anollés and Gresshoff, 1994) (Appendix G). Forty-two samples were loaded in seven lane intervals on each gel (Figure 17). For each gel, the allele bands (Figure 17) were manually scored against previously designated reference bands. The allele band that had progressed to the same position with any reference band was thus defined relative to that reference band, otherwise was defined as a new allele. The images of gel were captured using the GeneSnap program (Syngene). The Gene Tool program (Syngene) was used to

estimate the size of each band compared to the Low Molecular Weight DNA Ladder (BioLabs Inc.).

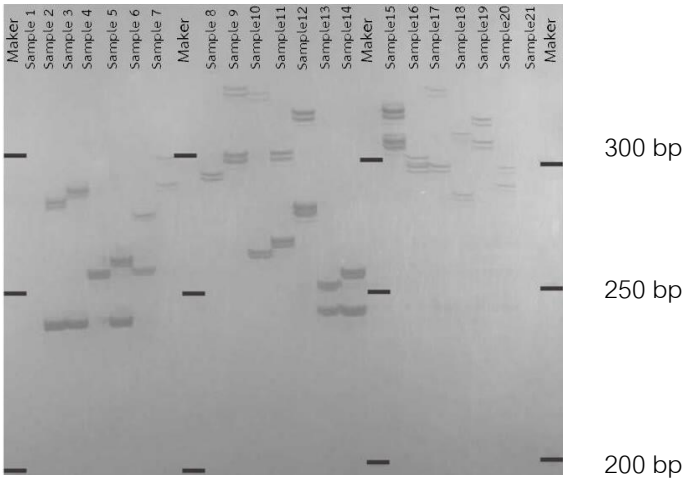


Figure 17 Visualization of Polyacrylamide gel of Lru43. Example allele band of twenty-one samples and four lanes of Low Molecular Weight DNA Ladder showed range of allele samples from 200 bp to upper 300 bp.

### 3.4.4 Data analyses

#### 1) Microsatellite marker efficiency

Sources of error include poor amplification, variations in the primer binding site, incorrect interpretation of stutter patterns, or contamination. In many cases, knowing the sources of error in the genotype data can allow certain corrections to be applied. Stutter peaks (small peaks that occur immediately before or after a real allele band) occur as a by-product of PCR. However, it is sometimes difficult to distinguish stutter bands of the first allele from a secondary contributor that might exist in a close position; it can be a cause of errors in allele scoring (Bagley *et al.,* 1999). I avoided this error by counted the first band of every stutter allele and not counting any allele that had a more complicated stutter pattern. Variations in the primer binding site may inhibit the efficiency of amplification (Callen *et al*., 1993), leading to domination by only one allele

that is easy to prime (and probably causes the majority of “large allele drop out”). I therefore examined the relative frequency of missing genotype data of each locus due to unsuccessful amplification and/or uncounted alleles at PAGE step as a means to gauge microsatellite marker efficiency.

The term “null allele” refers to non-amplifying alleles (Bagley *et al.,* 1999) are problematic of priming or extension in Polymerase chain reaction (PCR) step. Using Micro-checker program (van Oosterhout *et al*., 2004), I tested the possibility of stutter bands, large allele drop out and null alleles which may cause a deviation from Hardy-Weinberg equilibrium. Gene frequencies and genotype numbers were estimated using the Expectation Maximization (EM) algorithm (Dempster *et al*., 1977, implemented in FreeNA program) to correct for null alleles (Chapuis and Estoup, 2007).The corrected frequencies were used for computing $F_{ST}$ (Appendix H).

**2) Genetic diversity**

The number of alleles, allelic richness, private allele and the estimation of observed heterozygosity ($H_o$) and expected heterozygosity ($H_e$) were computed using the FSTAT (Goudet, 2001) and GDA programs (Lewis and Zaykin, 2001). Hardy-Weinberg equilibrium was revealed by the inbreeding coefficients within population (*Fis*) and tested for significance using a Markov chain method with 10 000 dememorizations, 100 batches, 5000 iterations per batch implemented in Genepop program version 4.1(Ovenden and Street, 2003).

**3) Population genetic structure**

The analysis of molecular variance (AMOVA) was performed using ARLEQUIN version 3.5 (Excoffier *et al.,* 2001) for testing population differentiation between the GT and the Andaman Sea groups. Pairwise $F_{ST}$ of nine population following Weir and Cockerham’s equation – examine difference number of alleles between loci and population – were calculated using Genepop program version 4.1 (Ovenden and

Street, 2003). POWSIM was applied to determine the statistic power of the 3 population data set to detect population differentiation (Ryman and Palm, 2006). The tested $F_{ST}$ values were set corresponding to the empirical $F_{ST}$ values (see more Appendix I). Slatkin's $R_{ST}$ (Hardy *et al*., 2003) was calculated and used to test for differentiation between populations based on a stepwise mutational model of evolution (Wiesner *et al*, 2011) using RSTCAL (Goodman, 1997). Perform both $F_{ST}$ and $R_{ST}$ would show more precise of variation between population genetic examinations, guarantee this idea by many study (Bagley *et al*., 1999; Rocha-olivares, 2003; Wiesner *et al*., 2011). A standard Bonferroni correction was applied to correct for the level of significance for multiple tests (Salini *et al*., 2006). Tests of population differentiation using pairwise $F_{ST}$ before and after corrections for null allele frequencies were compared to reveal the effects of null alleles on population genetic structure examination.

Possible cluster of individuals based on genotypes at several loci was identified by a Bayesian approach using STRUCTURE (Pritchard *et al*., 2000). The potential numbers of cluster (K) were set from 1 to 9. The program was run for 10000000 replicates with 2500000 burn-in period. No Admixture model was assumed. Number of population (K) obtained from STRUCTURE software was detected by the maximal value of log likelihood L(K). However, Pritchard *et al.* (2000) who created the STRUCTURE suggested that while this software was suitable for estimating small numbers of discrete population, it is difficult to obtain accurate numbers of population and the projected result may not directly explain potential biological aspects of a K that was obtained from this software. Moreover, Evanno (2005) argued that log likelihood L(K) was biased if samples have complex dispersal among groups and in effect was mostly suggesting an uppermost of clusters. So log likelihood obtained from STRUCTURE software may be not suggested the real K. Recently, an ad hoc quantity ($\Delta K$), that presents the second change rate of the log probability of data between the potential numbers of cluster (K)" (Evanno, 2005), was used to reveal the most probable number of clusters. $\Delta K$ was use log likelihood L(K) value of 20 run in each K and calculated by:

$$\Delta K = m(|L(K + 1)- 2 L(K) + L(K-1)|) /s[L(K)]$$ (more detail in Appendix J).

Isolation by distance was examined by a Mantel test. Slatkin's linearize $F_{ST}$ was tested for correlation with the geographic distance between recent populations defined by the logarithms (log) of kilometers of coastal line. Geographic distance need to be referred as a pathway to connect of fish population in the present. I estimated geographic distance coastal line by Google Earth version 6.1.7601.1 (Google, Inc, 2011) from Surin Island, Phangnga province passed all sample sites in Andaman Sea to Strait of Malacca and went to the lower of Malay Peninsula after that went northwestward along the eastern coast to Tao Island, Surat Thani province. Mantel tests were twice performed on the data set, the first time for Gulf of Thailand and Andaman Sea population, it show genetic correlation between population that separate by 9,000 kilometer from Ko Adang - Rawi, Satun province to Tha Rai, Nakhon Si Thammarat province, the shorter pathway of connectivity between these population. The second time was between SA and NA population that was separated by 1,600 kilometer from the shorter pathway of connectivity between two populations, Ko Ra, Phangnga province to Pak Ba Ra port, Satun province.

## 3.5 RESULTS

### 3.5.1. Microsatellite marker efficiency

Allele sizes of all locus were found to be consistent with those reported by Guo *et al*. (2007b), confirming that the alleles of samples studied are the expected allele sizes (Table 4).

Percent of missing genotype data varied from 1.30% (Lru11) to 13.00% (Lru30). This range of missing percentage is acceptable – most of the samples were successfully amplified. There was no large allele drop out and most of the loci studied were not problematic with stuttering bands. The amplification step was achievable for most of the selected primers. The proportion of the substances in the PCR reactions and the temperatures that were employed in this study were suitable. The possibility of stutter bands was significant for Lru10, Lru14 and Lru30 (Table 4)

When data from all locations were pooled and analysed together, six loci were apparently affected by null alleles, however the effect disappeared when the samples were separated by locality. This apparent discrepancy might be the result of several factors. The presence of null alleles when analyzed all samples together was significant in many loci (Lru01, Lru10, Lru14, Lru25, Lru30, and Lru43). However, the presence of null alleles in Gulf of Thailand population (GT) was evident only in Lru14 and Lru30, for northern and southern Andaman Sea populations presented in Lru10, Lru14, Lru25, and Lru30. It is likely, therefore, that null alleles at some loci may not affect deviation from Hardy–Weinberg (HW) equilibrium of every population.

Table 4 Allele size (bp), missing genotype, and results from Microchecker.

| Locus name | Northern Andaman sea (NA) | | | Southern Andaman sea (SA) | | | Gulf of Thailand (GT) | | | Total | Size (bp) | Homozygote excess | stuttering band | Large allele drop out | null allele | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NA1 | NA2 | NA3 | SA1 | SA2 | SA3 | GT1 | GT2 | GT3 | | | | | | NA | SA | GT | Total |
| Lru01 | 4.9 | 12.5 | 0.0 | 21.6 | 2.3 | 0.0 | 20.0 | 0.0 | 9.4 | 7.0 | 183-113 | Yes | No | No | No | No | No | Yes |
| Lru02 | 14.6 | 4.2 | 12.2 | 0.0 | 4.5 | 2.8 | 10.0 | 0.0 | 3.1 | 6.0 | 205-169 | No | No | No | No | No | No | No |
| Lru03 | 0.0 | 4.2 | 2.4 | 10.8 | 4.5 | 8.3 | 5.0 | 12.0 | 3.1 | 5.3 | 172-96 | Yes | No | No | No | No | No | Yes |
| Lru10 | 2.4 | 0.0 | 0.0 | 2.7 | 4.5 | 8.3 | 5.0 | 0.0 | 9.4 | 3.7 | 208-120 | Yes | Yes | No | Yes | Yes | No | Yes |
| Lru11 | 2.4 | 4.2 | 0.0 | 2.7 | 0.0 | 2.8 | 0.0 | 0.0 | 0.0 | 1.3 | 208-142 | No | No | No | No | No | No | No |
| Lru14 | 0.0 | 8.3 | 0.0 | 8.1 | 4.5 | 2.8 | 20.0 | 12.0 | 6.3 | 5.7 | 276-178 | Yes | Yes | No | Yes | Yes | Yes | Yes |
| Lru25 | 2.4 | 4.2 | 0.0 | 18.9 | 6.8 | 5.6 | 10.0 | 0.0 | 9.4 | 6.3 | 198-140 | Yes | No | No | Yes | Yes | No | Yes |
| Lru30 | 14.6 | 4.2 | 4.9 | 10.8 | 9.1 | 5.6 | 40.0 | 40.0 | 6.3 | 13.0 | 290-184 | Yes | Yes | No | Yes | Yes | Yes | Yes |
| Lru43 | 7.3 | 4.2 | 2.4 | 32.4 | 0.0 | 0.0 | 35.0 | 8.0 | 9.4 | 9.7 | 356-242 | Yes | No | No | No | No | No | Yes |

### 3.5.2. Genetic diversity and tests of Hardy–Weinberg (HW) equilibrium

The genetic descriptive data was described in Table 5. The number of alleles of the microsatellite loci studied ranged from 12 (Lru02) to 42 (Lru14). Allelic richness varied between population; that is, from 15.238 to 29.538 for the GT population, 10.779 to 35.883 for NA population, and 11.050 to 37.415 for SA population. The average gene diversity across loci is similar in all three populations; GT = 0.933, NA = 0.909 and SA = 0.916. The average gene diversity of each locus was varied; Lru02 locus showed the lowest gene diversity (0.795) while Lru30 locus showed the highest gene diversity (0.957). The highest number of private allele in all loci was found in the GT population.

Averaged across samples, six of the nine microsatellite loci studied (Lru01, Lru10, Lru14, Lru25, Lru30, and Lru43) showed significant heterozygote deficiencies after sequential Bonferroni correction (Rice, 1989) except for Lru02, Lru03, Lru011 (Table 5). However, deviation from Hardy–Weinberg equilibrium was not present in all those loci when samples were separated into three putative subpopulations (GT, NA, and SA) (Table 5). Only loci Lru10, Lru14, and Lru30 showed significant heterozygote deficiencies in all three population samples. Moreover, the Gulf of Thailand population (GT) did not deviate from Hardy–Weinberg equilibrium at Lru43. None of the subpopulation samples (from any of GT, NA, and SA) exhibited consistent deviation from Hardy-Weinberg equilibrium at every locus studied. However, when I calculated each subpopulation sample across all loci, it was evident that there were departures from Hardy-Weinberg equilibrium.

Table 5 Descriptive data of genetic diversity of *L. russelli* population in Thailand

| | | | NA | | | SA | | | GT | | | NA | SA | GT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Locus | Total | NA1 | NA2 | NA3 | SA1 | SA2 | SA3 | GT1 | GT2 | GT3 | | | | |
| All | | | | | | | | | | | | | | |
| N | 300 | 41 | 24 | 41 | 37 | 44 | 36 | 20 | 25 | 32 | 106 | 117 | 77 | |
| Na | 31 | 20 | 17 | 20 | 19 | 22 | 20 | 15 | 19 | 20 | 25 | 27 | 25 | |
| Pa | - | 2 | 0 | 1 | 4 | 6 | 1 | 1 | 1 | 2 | 2 | 11 | 17 | |
| Ar | 13.705 | 12.698 | 12.236 | 12.629 | 12.461 | 13.097 | 12.965 | 12.979 | 13.857 | 13.505 | 22.362 | 24.020 | 23.697 | |
| Ho | 0.854 | 0.821 | 0.831 | 0.811 | 0.856 | 0.886 | 0.864 | 0.854 | 0.915 | 0.842 | 0.819 | 0.871 | 0.867 | |
| He | 0.917 | 0.910 | 0.900 | 0.908 | 0.913 | 0.911 | 0.914 | 0.923 | 0.938 | 0.934 | 0.909 | 0.916 | 0.933 | |
| Fis | | **0.099 | **0.078 | **0.108 | **0.064 | **0.027 | **0.055 | **0.075 | 0.025 | **0.101 | **0.819 | **0.871 | **0.867 | |
| Lru01 | | | | | | | | | | | | | | |
| N | 279 | 39 | 21 | 41 | 29 | 43 | 36 | 16 | 25 | 29 | 101 | 108 | 70 | |
| Na | 31 | 19 | 16 | 19 | 18 | 23 | 19 | 15 | 19 | 19 | 23 | 25 | 22 | |
| Pa | - | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 6 | |
| Ar | 13.760 | 12.223 | 12.283 | 13.456 | 12.689 | 13.391 | 12.804 | 13.054 | 14.048 | 14.245 | 21.605 | 22.826 | 21.493 | |
| Ho | 0.882 | 0.872 | 0.857 | 0.878 | 0.896 | 0.837 | 0.972 | 0.812 | 0.88 | 0.931 | 0.871 | 0.938 | 0.886 | |
| He | 0.934 | 0.916 | 0.921 | 0.943 | 0.926 | 0.935 | 0.932 | 0.938 | 0.940 | 0.950 | 0.930 | 0.898 | 0.945 | |
| Fis | **0.055 | 0.049 | 0.071 | 0.070 | 0.032 | 0.106 | -0.044 | 0.137 | 0.065 | 0.020 | 0.063 | 0.042 | 0.063 | |
| Lru02 | | | | | | | | | | | | | | |
| N | 294 | 40 | 24 | 41 | 37 | 42 | 35 | 18 | 25 | 32 | 105 | 114 | 75 | |
| Na | 16 | 9 | 10 | 10 | 8 | 9 | 10 | 11 | 13 | 13 | 12 | 13 | 16 | |
| Pa | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| Ar | 8.320 | 6.500 | 7.127 | 5.677 | 5.755 | 6.451 | 7.235 | 9.093 | 10.383 | 9.610 | 10.779 | 11.050 | 15.238 | |
| Ho | 0.820 | 0.800 | 0.750 | 0.805 | 0.703 | 0.833 | 0.857 | 0.889 | 0.960 | 0.781 | 0.790 | 0.798 | 0.867 | |
| He | 0.795 | 0.752 | 0.756 | 0.739 | 0.773 | 0.726 | 0.782 | 0.852 | 0.896 | 0.885 | 0.746 | 0.765 | 0.889 | |
| Fis | -0.0293 | -0.064 | 0.008 | -0.090 | *0.092 | -0.149 | -0.098 | -0.044 | -0.073 | 0.119 | -0.059 | -0.044 | 0.025 | |
| Lru03 | | | | | | | | | | | | | | |
| N | 285 | 41 | 23 | 41 | 33 | 42 | 33 | 19 | 22 | 31 | 105 | 108 | 72 | |
| Na | 30 | 20 | 14 | 17 | 19 | 22 | 23 | 19 | 22 | 23 | 21 | 29 | 28 | |
| Pa | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | |
| Ar | 14.197 | 13.119 | 11.576 | 12.782 | 12.679 | 13.265 | 14.940 | 14.655 | 15.928 | 15.771 | 19.257 | 24.967 | 27.384 | |
| Ho | 0.904 | 0.878 | 0.870 | 0.878 | 0.939 | 0.952 | 0.970 | 0.684 | 1.000 | 0.968 | 0.876 | 0.954 | 0.903 | |
| He | 0.941 | 0.928 | 0.918 | 0.937 | 0.932 | 0.930 | 0.952 | 0.943 | 0.960 | 0.962 | 0.928 | 0.938 | 0.957 | |
| Fis | 0.029 | 0.054 | 0.054 | 0.064 | -0.008 | -0.024 | -0.019 | **0.280 | -0.043 | -0.006 | 0.056 | -0.017 | 0.057 | |
| Lru10 | | | | | | | | | | | | | | |
| N | 289 | 40 | 24 | 41 | 36 | 42 | 33 | 19 | 25 | 29 | 105 | 111 | 73 | |
| Na | 36 | 23 | 20 | 24 | 26 | 26 | 23 | 19 | 21 | 26 | 29 | 32 | 28 | |
| Pa | - | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 3 | 1 | |
| Ar | 15.336 | 14.021 | 14.512 | 14.346 | 14.396 | 15.482 | 14.69 | 15.206 | 14.654 | 16.44 | 26.712 | 28.224 | 26.957 | |
| Ho | 0.854 | 0.800 | 0.833 | 0.732 | 0.861 | 0.809 | 0.848 | 0.947 | 0.920 | 0.931 | 0.781 | 0.838 | 0.932 | |
| He | 0.950 | 0.944 | 0.944 | 0.944 | 0.929 | 0.955 | 0.946 | 0.959 | 0.951 | 0.964 | 0.948 | 0.949 | 0.956 | |
| Fis | **0.116 | *0.154 | 0.120 | **0.227 | *0.074 | **0.154 | **0.104 | 0.012 | 0.033 | 0.035 | **0.177 | **0.118 | **0.025 | |

Table 5 Descriptive data of genetic diversity of *L. russelli* population in Thailand (continued)

| | | | NA | | | SA | | | GT | | | NA | SA | GT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Locus | Total | NA1 | NA2 | NA3 | SA1 | SA2 | SA3 | GT1 | GT2 | GT3 | | | | |
| Lru11 | | | | | | | | | | | | | | |
| N | 296 | 40 | 23 | 41 | 36 | 44 | 35 | 20 | 25 | 32 | 104 | 115 | 77 | |
| Na | 28 | 23 | 19 | 20 | 25 | 22 | 21 | 19 | 19 | 20 | 25 | 28 | 26 | |
| Pa | - | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 2 | 0 | 0 | 1 | 5 | |
| Ar | 14.315 | 13.408 | 14.133 | 13.216 | 14.507 | 14.001 | 14.079 | 14.835 | 13.847 | 14.097 | 22.19 | 24.879 | 24.482 | |
| Ho | 0.945 | 0.900 | 1.000 | 0.854 | 0.917 | 0.954 | 0.971 | 1.000 | 1.000 | 0.906 | 0.904 | 0.948 | 0.961 | |
| He | 0.944 | 0.936 | 0.941 | 0.938 | 0.944 | 0.948 | 0.948 | 0.947 | 0.947 | 0.945 | 0.938 | 0.947 | 0.946 | |
| Fis | 0.008 | 0.039 | -0.064 | 0.091 | 0.030 | -0.006 | -0.025 | -0.057 | -0.057 | 0.042 | 0.037 | -0.001 | -0.016 | |
| Lru14 | | | | | | | | | | | | | | |
| N | 282 | 41 | 22 | 41 | 34 | 42 | 34 | 16 | 22 | 30 | 104 | 110 | 68 | |
| Na | 43 | 32 | 25 | 32 | 27 | 35 | 30 | 16 | 23 | 24 | 39 | 42 | 31 | |
| Pa | - | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 3 | |
| Ar | 17.017 | 17.3 | 16.767 | 17.343 | 16.212 | 18.222 | 16.46 | 13.467 | 14.975 | 14.522 | 35.883 | 37.415 | 29.538 | |
| Ho | 0.846 | 0.829 | 0.909 | 0.927 | 0.941 | 0.952 | 0.824 | 0.750 | 0.818 | 0.667 | 0.889 | 0.909 | 0.735 | |
| He | 0.954 | 0.970 | 0.959 | 0.969 | 0.962 | 0.975 | 0.961 | 0.913 | 0.924 | 0.934 | 0.970 | 0.969 | 0.926 | |
| Fis | **0.105 | **0.147 | 0.053 | 0.044 | 0.022 | 0.024 | **0.145 | 0.184 | 0.117 | **0.290 | **0.088 | **0.062 | **0.208 | |
| Lru25 | | | | | | | | | | | | | | |
| N | 282 | 40 | 23 | 41 | 30 | 42 | 34 | 18 | 25 | 29 | 104 | 106 | 72 | |
| Na | 19 | 10 | 12 | 14 | 13 | 11 | 12 | 13 | 15 | 13 | 17 | 17 | 19 | |
| Pa | - | 1 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 4 | 0 | |
| Ar | 9.516 | 7.409 | 8.147 | 8.695 | 9.546 | 7.891 | 8.747 | 10.978 | 11.713 | 9.817 | 14.609 | 14.518 | 18.052 | |
| Ho | 0.800 | 0.750 | 0.783 | 0.634 | 0.867 | 0.810 | 0.735 | 0.944 | 0.880 | 0.793 | 0.712 | 0.802 | 0.861 | |
| He | 0.861 | 0.836 | 0.807 | 0.825 | 0.874 | 0.83 | 0.831 | 0.910 | 0.926 | 0.901 | 0.824 | 0.851 | 0.909 | |
| Fis | *0.083 | 0.104 | 0.031 | 0.233 | 0.009 | 0.026 | 0.117 | -0.040 | 0.050 | 0.122 | 0.137 | 0.058 | 0.053 | |
| Lru30 | | | | | | | | | | | | | | |
| N | 261 | 35 | 23 | 39 | 33 | 40 | 34 | 12 | 15 | 30 | 97 | 107 | 57 | |
| Na | 45 | 29 | 18 | 25 | 21 | 29 | 27 | 13 | 17 | 24 | 32 | 35 | 28 | |
| Pa | - | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | |
| Ar | 17.231 | 16.860 | 13.609 | 14.992 | 14.552 | 15.924 | 15.713 | 13.000 | 15.047 | 14.998 | 29.471 | 30.752 | 28.000 | |
| Ho | 0.759 | 0.743 | 0.565 | 0.718 | 0.818 | 0.875 | 0.794 | 0.750 | 0.867 | 0.700 | 0.691 | 0.832 | 0.754 | |
| He | 0.957 | 0.966 | 0.940 | 0.949 | 0.950 | 0.958 | 0.956 | 0.949 | 0.956 | 0.952 | 0.958 | 0.958 | 0.954 | |
| Fis | **0.203 | **0.234 | **0.404 | **0.246 | 0.140 | *0.088 | *0.171 | 0.217 | 0.097 | **0.268 | **0.280 | **0.132 | **0.210 | |

Table 5 Descriptive data of genetic diversity of *L. russelli* population in Thailand (continued)

| | | NA | | | SA | | | GT | | | NA | SA | GT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Locus | Total | NA1 | NA2 | NA3 | SA1 | SA2 | SA3 | GT1 | GT2 | GT3 | | | |
| Lru43 | | | | | | | | | | | | | |
| N | 271 | 38 | 23 | 40 | 25 | 44 | 36 | 13 | 23 | 29 | 101 | 105 | 65 |
| Na | 27 | 19 | 15 | 20 | 15 | 19 | 18 | 13 | 19 | 18 | 23 | 24 | 23 |
| Pa | - | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Ar | 13.652 | 13.44 | 11.971 | 13.153 | 11.815 | 13.248 | 12.021 | 12.52 | 14.12 | 12.041 | 20.748 | 21.552 | 22.132 |
| Ho | 0.873 | 0.816 | 0.913 | 0.875 | 0.760 | 0.954 | 0.806 | 0.923 | 0.913 | 0.896 | 0.861 | 0.857 | 0.908 |
| He | 0.926 | 0.941 | 0.911 | 0.932 | 0.929 | 0.931 | 0.919 | 0.895 | 0.945 | 0.912 | 0.936 | 0.931 | 0.918 |
| Fis | **0.062 | *0.134 | -0.002 | 0.063 | *0.185 | -0.017 | **0.125 | -0.032 | 0.034 | 0.018 | **0.080 | **0.080 | 0.011 |

Note: **N** is number of individuals, **Na** is number of alleles, **Pa** is number of private alleles, **Ar** is allelic richness, **Ho** is observed heterozygosity, **He** is expected heterozygosity, **Fis** is inbreeding coefficient within samples

* indicated statistical significance after Borferroni correction when p-value < 0.05/ number of test

** indicated statistical significance after Borferroni correction when p-value < 0.01/ number of test

### 3.5.3 Population genetic differentiation

The AMOVA test showed that there are significant genetic differentiation between the Gulf of Thailand and the Andaman Sea population groups ($F_{ST}$ = 0.037, p = 0.019), which contributes to 3.73 % of the total genetic variation (Table 6). However, there is no significant differentiation between populations within each region ($F_{ST}$ =0.002, p = 0.129). Pairwise $F_{ST}$ statistic theoretically ranges from zero to one, by which near-zero values indicate the lack of genetic differentiation, and sometimes can also be found as negative values indicating a greater heterozygosity within than between populations (Shulzitski *et al.*, 2009). Consistent with the AMOVA results, the pairwise $F_{ST}$ and $R_{ST}$ were the Gulf of Thailand significant difference with NA and SA populations, and there was no significant population differentiation between the NA and the SA populations ($F_{ST}$ >0.001, p=0.284) (Table 7).

Table 6 AMOVA results testing population differentiation between the Gulf of Thailand and the Andaman Sea population groups

| Source of variation | d.f. | Sum of squares | Variance components | Percentage of variation | P-value |
|---|---|---|---|---|---|
| Among groups | 1 | 16.98 | 0.07 | 3.73 | 0.019 |
| Among populations within groups | 7 | 13.18 | >0.01 | 0.15 | 0.129 |
| Within populations | 591 | 1005.29 | >0.01 | 96.12 | 0.000 |
| Total | 599 | 1035.45 | 1.770 | | |

An unexpected result was found in the comparison between the offshore coral reefs and the inshore coral reef in the southern Andaman Sea and between the offshore coral reef and the coastal population samples in the same region. This result showed significant population differentiations with $F_{ST}$ = 0.0008 (p <0.0001) and 0.009 (p <0.0001), respectively (Table 8). However, there was some discrepancy within the results between $F_{ST}$ and $R_{ST}$. Pairwise. $R_{ST}$ did not show any difference within the southern Andaman Sea population.

Table 7 Pairwise $F_{ST}$ (below diagonal) and $R_{ST}$ (upper diagonal) between 3 populations of *L. russelli* in Thailand

| | NA | SA | GT |
|---|---|---|---|
| NA | - | -0.003<br>(0.972) | 0.206<br>(0.000)** |
| SA | >0.001<br>(0.284) | - | 0.181<br>(0.000)** |
| GT | 0.032<br>(0.000)** | 0.028<br>(0.000)** | - |

* Indicated statistical significance; P<0.017, after Borferroni correction = 0.05/ 3

** Indicated statistical significance; P<0.003, after Borferroni correction = 0.01/ 3

The presence of null alleles was found to have negligible effect on the estimation of $F_{ST}$ values. Comparing between the $F_{ST}$ values from the original data and those after correcting for null alleles found only slight differences (Appendix H). Therefore, the original data of all 9 microsatellite loci were used to test for significance of population differentiation.

Table 8 Pairwise $F_{ST}$ (below diagonal) and $R_{ST}$ (upper diagonal) between 9 populations of *L. russelli* in Thailand

| | NA1 | NA2 | NA3 | SA1 | SA2 | SA3 | GT1 | GT2 | GT3 |
|---|---|---|---|---|---|---|---|---|---|
| NA1 | - | 0.011 (0.140) | -0.005 (0.786) | 0.001 (0.450) | -0.002 (0.584) | 0.006 (0.204) | 0.194** (0.000) | 0.201** (0.000) | 0.195** (0.000) |
| NA2 | 0.001 (0.457) | - | 0.002 (0.373) | 0.009 (0.152) | 0.012 (0.084) | 0.003 (0.358) | 0.212** (0.000) | 0.216** (0.000) | 0.210** (0.000) |
| NA3 | -0.001 (0.670) | 0.005 (0.128) | - | 0.007 (0.144) | -0.001 (0.516) | -0.002 (0.552) | 0.195** (0.000) | 0.203** (0.000) | 0.195** (0.000) |
| SA1 | 0.004 (0.052) | 0.011 (0.002) | 0.005 (0.005) | - | 0.005 (0.200) | 0.014 (0.024) | 0.170** (0.000) | 0.179** (0.000) | 0.184** (0.000) |
| SA2 | -0.001 (0.268) | 0.003 (0.233) | 0.001 (0.117) | 0.008** (0.000) | - | 0.016 (0.018) | 0.188** (0.000) | 0.186** (0.000) | 0..179** (0.000) |
| SA3 | 0.002 (o.444) | 0.003 (0.508) | 0.001 (0.218) | 0.009** (0.000) | 0.001 (0.401) | - | 0.184** (0.000) | 0.198** (0.000) | 0.194** (0.000) |
| GT1 | 0.035** (0.000) | 0.040** (0.000) | 0.037** (0.000) | 0.036** (0.000) | 0.037** (0.000) | 0.034** (0.000) | - | -0.005 (0.546) | 0.027 (0.048) |
| GT2 | 0.030** (0.000) | 0.034** (0.000) | 0.031** (0.000) | 0.031** (0.000) | 0.035** (0.000) | 0.027** (0.000) | -0.002 (0.900) | - | -0.004 (0.616) |
| GT3 | 0.027** (0.000) | 0.032** (0.000) | 0.029** (0.000) | 0.027** (0.000) | 0.030** (0.000) | 0.022** (0.000) | 0.002 (0.666) | -0.002 (0.844) | - |

* indicated statistical significance; P<0.001, after Borferroni correction = 0.05/ 35

** indicated statistical significance; P<0.000, after Borferroni correction = 0.01/ 35

Since there were nine locational groupings amongst the data set, the maximum number of populations possible is nine. Results from the POWSIM software indicated that the population data categorized *a priori* into 3 nominal groups has a high statistical power to detect population differentiation (83.1%- 85% for $F_{ST}$ = 0.001, and 100% for $F_{ST}$ =0.05), and the Type I error that was obtained by omitting the drift step ($F_{ST}$ =0) was low (4.7%-6.3%) (Appendix I). Statistical power of the test of data was not checked for the case where data were categorized into 9 putative populations.

Number of population structures that could be obtained from maximal value of log likelihood L(K) by STRUCTURE software was 5 clusters (Figure 19) but the Q plots (Figure 18) and the $\Delta$K (Figure 20) suggested only 2 population groups. Q plots showed homogeneity of populations in Andaman Sea (red area) and strong separation from the Gulf of Thailand population (green area). Apart from a few scattered individuals (for instance, the Gulf of Thailand (green area) had three samples (descending red bars) that show high probability for the sample to be a member of Andaman Sea group, and a similar number of Andaman samples display opposite putative affiliations). It is likely that these represent divergences from population-level allele prevalence's, rather than misclassifications (i.e. the polymorphisms that distinguish the populations represent a statistical proportion of each population, rather than complete replacement; this is consistent with biogeographic theory). $\Delta$K is the ratio of the likelihood of the largest possible number of groups and a nominal number; Evanno (2005) argued that it is more likely to show the correct number of clusters in multiple models of population structure than log likelihood. Using $\Delta$K as a criterion, therefore, it is possible that the *L. russelli* populations are divided into two separated populations which are the GT and the Andaman Sea populations.

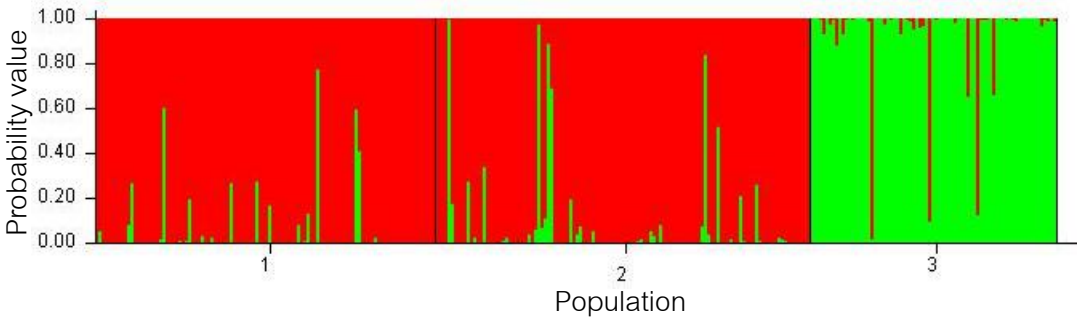


Figure 18 The probabilities of each sample in a single vertical line were defined to any given population. Red and green were the possibility of samples correctly assigned to Andaman Sea population and GT population, respectively. The numbers corresponds to the predefined population groups; 1is NA population; 2 is SA population and 3 is GT population.

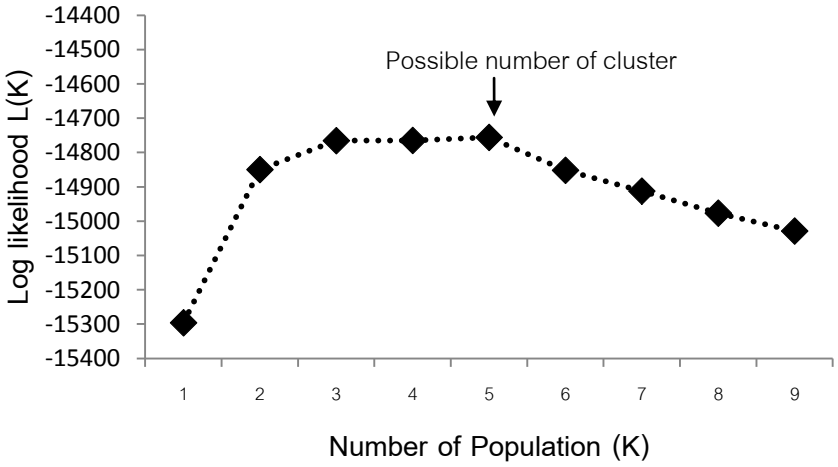


Figure 19 Mean likelihood L (K) of data over 20 runs for each K obtained from STRUCTURE software reveals the possible number of clusters.

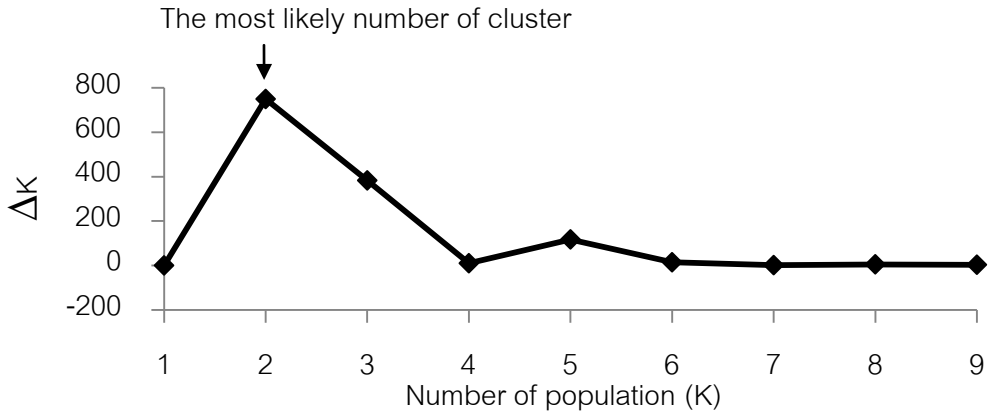


Figure 20 An ad hoc statistic $\Delta$K reveals the most likely number of clusters.

Correlation of Slatkin linear $F_{ST}$ values against the logarithm of the geographic distances revealed significant isolation by distance between GT population and the larger Andaman sea population ($R^2$= 0.812, P= 0.007). However within Andaman Sea revealed no significant correlation of genetic and geographic distance ($R^2$= -0.34, P= 0.947). It is possible that the geographical barrier of the Thai-Malay peninsula is the main factor that prohibits gene flow of *L. russelli* between the two oceans.

## 3.6 PRELIMINARY DISCUSSION

### 3.6.1 Genetic diversity and Hardy–Weinberg Equilibrium

The number of alleles per locus (allelic richness) in a population is a fundamental measure of genetic variation. The present study revealed a larger number of alleles and a higher range of allele sizes (measured in base-pairs: bp) than reported by Guo *et al* (2007b). However, allelic richness can be conflated with sample size (or sampling intensity: Kalinowski, 2004). Thus, the higher numbers are probably because this study involved a larger number of samples. Guo *et al.* (2007b) assessed the variability of microsatellite markers using 20 samples, whereas this study employed 300

samples, giving a higher chance to observe a larger number of alleles with more variable sizes. The numbers of alleles identified for *L. russelli* are higher than have been reported for other species within *Lutjanus* genus, (Carson *et al*., 2011; Karlsson *et al*., 2009; Muths *et al*., 2012; Ovenden and Street, 2003; and Shulzitski *et al*., 2009). In addition, this study may have included at least two genetically distinct stocks; the number of alleles at various loci is therefore the combination of alleles present in the two groups. These results can confirm partitioning of genetic stocks by evidence of the high number of private alleles in the Gulf of Thailand population sample (36 alleles). The genetic diversity can be different based on life history and demographic history of the species. For the most part, previously reported allelic diversity in *Lutjanus* spp. includes fewer than 40 alleles, whereas this study resulted in identification of 42 alleles in Lru 14. Moreover, allelic richness and the number of private allele (the alleles unique to a population) in *L. russelli* are also higher than those reported for other species.

Deviations from Hardy-Weinberg equilibrium of six of the nine microsatellite loci studied in pooled population samples were not apparent when individual samples were separated according to locality (GT, NA and SA). The heterozygote deficiency was more possible occurrence when analyzed across obviously variant population samples or genetically distinct stocks. Hardy-Weinberg deviations have also been reported in many other lutjanid fishes such as *L. analis* (Carson *et al*., 2011; Shulzitski *et al*., 2009), *L. griseus* (Renshow *et al*., 2007), *L. kasmira* (Muths *et al*., 2012), and *L. synagris* (Karlsson *et al*., 2009).

Artifacts of the screening process and amplification efficiency of studied markers may cause apparent homozygote excess and significant departure from Hardy-Weinberg equilibrium similar to that was found in loci Lru10, Lru14, and Lru30. Biological phenomena may not be responsible for the perceived homozygote excess of these three loci, because only these three loci deviated from Hardy-Weinberg equilibrium in all three subpopulations (GT, NA, SA). Other loci did not exhibit homozygote excess across all regions, so it may be assumed that homozygote excess is not specific to particular

subpopulation samples or biological perturbation of subpopulation samples but, rather, because of amplification efficiency of the markers.

Stutter bands are one type of error of allele scoring; it is sometimes difficult to distinguish stutter bands of the first allele from a secondary contributor that might exist in a closely adjacent position (Bagley *et al.,* 1999). In these results, there is strong evidence that stutter bands contribute to discrepancies in allele scoring for loci Lru10 and Lru14; there is somewhat less compelling evidence that stutter bands were significant for Lru30. In addition, apparent homozygote excess can be present because of a failure of PCR amplification and/or null allele. A failure of amplification can potentially result from many factors, such as DNA template degradation, unsuitability of incubation temperature, and the presence of a PCR inhibitor. A researcher would detect such anomalies by assessing percentage of missing genotype in each locus (Phinchongsakuldit *et al*., 2013). The percent of missing genotype was low at loci Lru 10 and Lru14 (3.7% and 5.7% respectively); it confirms that apparent homozygote excess at Lru 10 and Lru14 is due to another cause. Therefore, these two loci are potentially affected by the presence of null alleles potentially caused by nucleotide variation in the primer binding sites. In this case, one of the two alleles was not amplified; it therefore appeared as a homozygote instead of a heterozygote. However, Lru 30 has a high percentage of missing genotype and null alleles were present in every population sample, so the cause of heterozygote deficiencies in this locus can be attributed to both the failure of PCR amplification and null alleles.

### 3.6.2 Population genetic structure and biogeographic implications

Based on the results of this study, it is apparent that the population sample of *L. russelli* in the Gulf of Thailand is genetically separated from the population samples in both the Northern Andaman and the Southern Andaman Sea. The Mantel test also revealed correlation between genetic distance and geographic distance over 9,000 kilometers from Adang – Rawi Island, Satun province to Tha Rai Island, Nakhon Si Thammarat province, the shortest pathway of connectivity between eastern and western

population. The distinction between South China Sea (GT) and Indian Ocean (AN) stocks may have important biogeographic implications.

Variations of sea levels across Southeast - Asia and loss of maritime connectivity during eustatic low sea level episodes during the Pleistocene may be an important factor in restricting gene flow of organisms between the Indian Ocean to the Western Pacific Ocean (Voris, 2000). The Indo-Pacific Barrier has been reported to have affected heterogeneity in marine fish such as *Myripristies berndti* (Craig *et al*., 2007), and *Lutjanus fulvus* (Gaither *et al*., 2010). The late Holocene intrusion from South China Sea water mass into the southern Andaman Sea does not seem to have greatly diluted genetic differences between the eastern and the western populations. The population genetic differences and differences in color morph of *L. russelli* (Chapter 2) between Indian and Pacific Ocean population (e.g. Allen, 1985; Satapoomin, 2011) presents an argument of uncertain classification within this species.

There is no strong indication of population samples disjunction of this fish species between the SA and the NA. Genetic homogeneity of coastal populations of the SA and the NA populations implies connectivity of fish within these regions. Both regions may be potentially sources of larval recruits for each other. The flow of water from the South China Sea, through Malacca Straights (Wyrtki, 1961 (in Rizal *et al*., 2010)) converges with the water mass from the Bay of Bengal in the northeastern monsoon, but does not appear to affect genetic homogeneity of this fish within the Andaman Sea. This water current possibly assists recruitment of the southern larvae into the northern region in some monsoon periods; however, this does not need to be an annual or even decadal event to maintain connectivity given that individuals of this species may be reproductively competent for 20 or more years. Unfortunately, information regarding larval behavior and dispersal, biological, or life-history factors of this species is not well known for the Andaman Bioregion; it is difficult to speculate exactly how this water current affects dispersal of active swimming larvae – if at all. However, I can assume that this species may spawn many times in a year; this is a common phenomenon amongst snapper fishes in the tropics (Allen, 1985). Adults can potentially move all

around the Andaman Sea, and exploit multiple localities as nursery grounds of their stocks. Importantly, although I collected samples many times in a year; it was possible that samples were collected from the same source of recruitment, and so under-sampled the regional stock.

Pairwise-$F_{ST}$ and pairwise-$R_{ST}$ should be similar trend but discrepancy of them presented in the SA population sample. It is potential cause of insufficient sample number. Small numbers of samples from sites SA2 (44 samples) and SA3 (36 samples) may reduce the statistical power of the test, as can differences in the scale of sampling at each location. More fish samples and other marine organism would be required to test for the presence of an unexpected barrier between offshore coral reef and coastal habitat within SA. Moreover, examination of population genetic down into around the Malacca Strait is require for qualified source of fishes in SA.

It is apparent that the stock of Russell's snapper – a fish that supports the livelihoods of many thousands of local fishers in coastal Thailand – is not simply a homogeneous and ubiquitous resource. In many ways, this research points out how little is really known about non-commercial fishes in South East Asia, more research is needed especially at larger scales to determine the additional management needs to be implemented to safeguard resource security for coastal fishers. This includes knowledge about larval behavior and vagility and whether implementation of closed seasons and marine reserves are justified to optimize fisheries productivity and resource conservation. Moreover, the explicit partitioning of *L. russelli* into discrete Gulf of Thailand and Andaman populations indicated by this study suggests that this species (and perhaps genus) may be overdue for taxonomic re-evaluation.