Genetic diversity of mud crabs, *Scylla tranquebarica* in Sabah, Malaysia based on Cytochrome C Oxidase (COI) gene sequence

Nurul Ain Mohd Sharif, Noor Amalia Shaiful Kahar, Kenneth Rodrigues, Julian Ransangan, and Annita Yong Seok Kian*

Borneo Marine Research Institute, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, 88400 Malaysia.

Received: 22 July 2015; Accepted: 15 December 2015

Abstract

Mud crabs genus *Scylla* are distributed across the Indo-West Pacific Oceans. Among the four species, *S. tranquebarica* dominates the mangrove areas in Sabah, Malaysia and constitutes the primary crustacean fishery resource. Overexploitation of this economically important fisheries resource can have a significant impact on population diversity. This study was conducted to evaluate the genetic diversity of *S. tranquebarica* from five important fishing grounds. The genetic diversity was estimated based on the cytochrome c oxidase (COI) gene sequence. A total of 143 individuals were sampled across the 5 fishing grounds. The findings revealed that the crabs in Sabah comprised 11 haplotypes with a mean haplotype diversity (*h*) of 0.5564 and a mean nucleotide diversity (*π*) of 0.0038. The molecular variance analysis (AMOVA) showed that the low genetic differentiation among crab individuals in the five fishing grounds. The low genetic diversity provides the basis for the establishment of a scientific breeding program to counteract the loss of genetic diversity which is the result of over-exploitation of this ecologically and economically important fisheries resource.

Keywords: genetic diversity, fisheries, *Scylla tranquebarica*, mtDNA, COI

1. Introduction

The mud crab which is commonly referred to as mangrove crab of the genus *Scylla* (Decapoda: Portunidae) is widely distributed in the mud flats and mangrove forests (Keenan *et al.*, 1998) throughout the tropical to warm temperature zone of Indo-West Pacific regions (Marichamy and Rajapackian, 2001; La Sara, 2010). The confusion in *Scylla* taxonomy has been a controversial issue especially in areas where more than one species has been reported. *S. serrata* is the only recognised species that has been documented in most scientific reports. However, in 1998, four new species of *Scylla* were described. The recategorisation of the species was achieved through morphometric characteristics and allosyme electrophoresis (Keenan *et al.*, 1998). These four species are *S. serrata* (Forskål, 1775), *S. olivacea* (Herbst, 1796), *S. tranquebarica* (Fabricius, 1798), and *S. paramamosain* Estampador, 1949.

Among these four species, *S. tranquebarica* has been reported to be present in Sabah coastal waters (Keenan *et al.*, 1998), where they prefer mangrove forests and coastlines inundated with reduced salinity (Keenan *et al.*, 1998). Unlike other mud crabs species, *S. tranquebarica* receives less attention in terms of biology, ecological distribution and genetic diversity, due to its limited distribution in the South China Sea, Indian Ocean and Western Pacific. Mud crabs are important for commercial fisheries and aquaculture production throughout their distribution (Shelly and Lovatelli, 2011). Besides that, mud crabs are important in trade markets as demand for them has been reported to be increasing on a yearly basis. Mud crab fishery in Malaysia is expected to continue to grow in the future because increasing demand for...
these crustaceans (Ikhwanuddin et al., 2011). However, even though Scylla mud crabs have been extensively exploited, little is known about these species (Ewel et al., 2009). The aim of this study is to determine genetic diversity of S. tranquebarica as it is essential for interpretation, understanding and effective management of this species.

Genetic variation refers to the diversity in gene frequencies resulting from mutation, but can often be modified by mechanisms such as selection and genetic drift (Lowe et al., 2004). If genetic variation exists within a species, any alterations due to environmental changes will allow some individuals to survive and reproduce (Lowe et al., 2004). The loss of genetic variation may reduce individual fitness and population adaptability (Lande, 1988), which becomes a great concern to conservation biologists due to the recent demographic bottlenecks in wild populations (Hedrick and Miller, 1992). Besides that, inbreeding may occur as the consequence of reduction in genetic diversity (Lacy, 1997; Matocq and Villablanca, 2001).

There are several ways to estimate the genetic variation, for instance allozyme gel electrophoresis, restriction fragment length polymorphism and mitochondrial DNA (mtDNA). MtDNA has an effective population size four times smaller than that of the nuclear genome (Birky et al., 1989), and is maternally inherited and haploid in nature. It has become a very effective marker (Lowe et al., 2004) to understand phylogeographic and population genetic in many taxa (Avise, 1994; Chu et al., 1999; Hillis 1996) as well as intraspecific analysis due to high degree of polymorphism (Hu et al., 2008). There are numerous studies on mtDNA that have been conducted in crustaceans; for example, coconut crab Birgus latro (Laverly et al., 1996), swimming crab Callinectes bellicosus (Pfeiler et al., 2005), mangrove crab Perisesarma guttatus (Silva et al., 2010), Scylla paramamosain (Ma et al., 2011), Scylla olivacea (Rosly et al., 2013), and Scylla serrata (Fratini and Vannini, 2002). While the genetic variations of other Scylla species have been well elucidated, the genetic diversity of S. tranquebarica in the coastal waters of Sabah, Malaysia is yet to be understood.

2. Materials and Methods

2.1 Sample collection

Samples of mud crabs, S. tranquebarica were identified following key identifications provided by Keenan et al., (1998). They were collected from five main mangrove areas in Sabah namely; Kota Marudu (06°33’N; 114°44’E), Menumbok (05°19’N; 115°22’E), Pitas (06°47’N; 117°01’E), Sandakan (05°50’N; 118°03’E) and Semporna (04°25’N; 118°34’E) (Figure 1). The mud crabs were caught using collapsible baited traps with big-eye fish (Priacanthus sp.) as used as bait. Samples collected from Kota Marudu (n=37), Menumbok (n=26), Pitas (n=25), Sandakan (n=28) and Semporna (n=27) were transported to the laboratory alive, for further analysis.

2.2 DNA extraction

Genomic DNA from muscle tissues of the 3rd leg of the mud crab was extracted using DTAB-CTAB method (Philips and Simon, 1995) and subjected to polymerase chain reaction (PCR).

2.3 Amplification of cytochrome oxidase I (COI)

The DNA region of mitochondrial COI gene was amplified by polymerase chain reaction (PCR) with in-house designed primers: COI-left: 5’- ACGCAGAGCTTTCGGTT GAT-3’ and COI-right: 5’-TGCAATGCACACAAATTAC-3’. PCR amplification consisted of 5X PCR buffer, 2.5 mM dNTPs, 25mM MgCl2, 1.0 µl each reverse and forward primers, 0.3 unit of Taq DNA polymerase (Promega), and 2.0 µl (~10.0 ng/µl) DNA. The PCR was performed in GeneAmp® PCR system 7900 with the following amplification conditions: 95°C for 3 minutes, followed by 30 cycles of 1 minute at 94°C, 1 minute at 55°C, and 1 minute at 72°C; with a final extension period of 5 minutes at 72°C and maintained at 4°C until further analysis. The PCR product was first separated on 1.5% agarose gel electrophoresis and later was purified using Qiagen Gel Extraction kit (Qiagen, USA). Purified PCR products were sent to AITBiotech, Singapore for direct sequencing.

2.4 Data analysis

The sequences were manually checked and trimmed using MEGA ver. 6.0 (Tamura et al., 2013) and the resulting sequences were aligned using CLUSTAL W ver. 1.6 (Thompson et al., 1994). The variable and parsimonious sites were identified using Software MEGA ver. 6.0. Haplotypes were identified using ARLEQUIN ver 3.5 (Excoffier and Lischer, 2010). The sequences were then deposited at the NCBI Genbank (www.ncbi.nih.gov) database and assigned the accession numbers KP300948 to KP301090.

Figure 1. Study sites for genetic diversity of Scylla tranquebarica in Sabah, Malaysia.
The haplotype number and molecular diversity indices (haplotype diversity, \( h \) and nucleotide diversity, \( \pi \)) were calculated using ARLEQUIN ver 3.5. The index \( h \) indicates the probability that two randomly chosen haplotypes were different in a single population, and \( \pi \) is the average percentage number of differences between all pairs of haplotypes in a single population. The phylogenetic tree of the haplotypes was constructed in neighbour joining with 10,000 bootstrap using MEGA ver. 6.0 software. Molecular variance analysis (AMOVA) was performed using ARLEQUIN ver 3.5 to investigate levels of population structure, and the fixation index \( F_{st} \) value was calculated to estimate the genetic differentiation between and within populations. The significant level of the test was assessed using 1000 permutations of each pairwise comparison.

To assess demographic history of the populations, Tajima’s \( D \) (Tajima, 1989) and Fu’s \( F_{s} \) (Fu, 1997) values were calculated using ARLEQUIN ver 3.5 with 1000 permutations. These tests assume that the population has been in mutation-drift balance for a long period of evolutionary times (Nei and Kumar, 2000). However, if the population is not under mutation-drift equilibrium due to sudden expansion will be resulted on negative values on neutrality tests. Mismatch analysis with 10000 bootstrap replicates was also performed using DnaSP ver 5.1 (Librado and Rozas, 2009) and ARLEQUIN ver 3.5. The population that has undergone rapid expansion will exhibit unimodal distribution, whereas a subdivided population or in demographic equilibrium is expected to have a multimodal distribution (Rogers and Harpending, 1992).

3. Results

3.1 Characteristics of COI sequence

A 409 bp partial COI length fragment of mtDNA COI gene was obtained from 143 individuals of *S. tranquebarica* from the five sampling sites. The average nucleotide frequencies of T, C, A, and G in all sequences were 37.54%, 15.17%, 30.31% and 16.98%, respectively. The AT content (67.85%) was higher than the CG content (32.15%).

### 3.2 Haplotypes and nucleotide diversity of *S. tranquebarica*

Alignment of the 143 sequences revealed 11 variable sites which indicated 4 parsimony informative sites (position: 108, 219, 288, 377), while other sites showed singleton mutations. Besides that, 11 haplotypes were determined from five different localities (Table 1). All five localities were characterised by two common haplotypes H1 (55.9%) and H2 (36.4%) shared among all populations. In addition, H6 was shared between two sites (Menumbok and Semporna) while the remaining 8 haplotypes were rare ones that only were found in few individuals. The phylogenetic analyses between the samples were constructed with other *Scylla* species and outgroup; *S. serrata* (Genbank accession no. AF097014.1), *S. paramamosain* (Genbank accession no. AY373351.1), *S. olivacea* (Genbank accession no. JX878351.1), *Thalamita crenata* (Genbank accession no. JX398104.1), *Portunus pelagicus* (Genbank accession no. KF604896) showed that the haplotypes samples were clearly grouped as *S. tranquebarica* (Figure 2). In addition, it showed that the mud crab haplotypes are grouped into one reciprocally monophyletic clade which distinguished from outgroup.

Haplotype diversity \( (h) \) and standard deviation per site ranged from 0.533±0.101 at Semporna to 0.582±0.059 at Menumbok with the mean for all localities of 0.558±0.026. Meanwhile, the mean nucleotide diversity \( (\pi) \) and standard deviation was 0.0039±0.0001, ranging from 0.003±0.002 (Semporna) to 0.004±0.003 (Kota Marudu, Menumbok, Pitas and Sandakan) (Table 2). It was found that the mud crab exhibited high haplotype diversity but low nucleotide diversity.

**Table 1.** Variable sites among 11 COI haplotypes of *S. tranquebarica* from the five localities mangrove areas (n=143)

<table>
<thead>
<tr>
<th>Hap.</th>
<th>Nucleotide position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>72</td>
</tr>
<tr>
<td>H1</td>
<td>T</td>
</tr>
<tr>
<td>H2</td>
<td>C</td>
</tr>
<tr>
<td>H3</td>
<td>C</td>
</tr>
<tr>
<td>H7</td>
<td>C</td>
</tr>
<tr>
<td>H8</td>
<td>.</td>
</tr>
<tr>
<td>H11</td>
<td>C</td>
</tr>
</tbody>
</table>

Note: All haplotypes are compared with H1.
3.3 Genetic differentiation

The molecular variance analysis (AMOVA) indicated that 99.79% of total genetic variation was contributed by differences within the population (locally) and only 0.21% was contributed by the differences among population (localities) \((\text{Fst}=0.002)\) (Table 3). No significant genetic differentiation was observed among samples from different populations (localities). Results indicated that the genetic variation among samples from different localities was limited. This suggests that the \(S.\ tranquebarica\) in the coastal waters of Sabah were not genetically structured.

3.4 Neutrality tests and mismatch analysis

For entire samples, Tajima’s D was negative but no significant deviation from mutation-drift equilibrium was recorded \((D = -0.495, P>0.05)\), which was also the same for Fu’s F’s \((F’s = -1.983, P>0.05)\). The negative results of these neutrality tests support for slight population expansion, but the non-significant negative value may be caused by expansion that has only been restricted to certain sampling sites of the crab population (Liao et al., 2010; Rosly et al., 2013). Besides that, the mismatch distribution analysis showed unimodal patterns which fitted poorly with their corresponding distribution (Figure 3). This also suggests that the population underwent population expansion.

The sum of the square deviation (SSD) per locality ranged from 0.147 to 0.379, with an average of 0.222, \((\text{Table 4})\). The populations of Pitas and Semporna had mismatch frequency spectra that significantly deviated from what would be expected under the rapid population growth (sudden population expansion; Rogers and Harpending, 1992) indicated by significantly positive values of the SSD test statistics. Raggedness index is calculated similarly, the raggedness indexes \((\text{Rag})\) per locality were between 0.200 and 0.602, with an average 0.455. Menumbok, Pitas and Semporna have significant raggedness which also may suggest a population expansion.

### Table 2. Genetic diversity of \(S.\ tranquebarica\) from five localities in Sabah coastal waters, Malaysia

<table>
<thead>
<tr>
<th>Locality</th>
<th>Sample Size, (n)</th>
<th>No. of Haplotype</th>
<th>Haplotype Diversity, (h \pm SD)</th>
<th>Nucleotide Diversity, (\pi \pm SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kota Marudu</td>
<td>37</td>
<td>5</td>
<td>0.554±0.006</td>
<td>0.004±0.003</td>
</tr>
<tr>
<td>Menumbok</td>
<td>26</td>
<td>5</td>
<td>0.582±0.059</td>
<td>0.004±0.003</td>
</tr>
<tr>
<td>Pitas</td>
<td>25</td>
<td>4</td>
<td>0.560±0.044</td>
<td>0.004±0.003</td>
</tr>
<tr>
<td>Sandakan</td>
<td>28</td>
<td>4</td>
<td>0.553±0.041</td>
<td>0.004±0.003</td>
</tr>
<tr>
<td>Semporna</td>
<td>27</td>
<td>5</td>
<td>0.533±0.101</td>
<td>0.003±0.002</td>
</tr>
</tbody>
</table>

Note: SD= standard deviation

Figure 2. Neighbour-joining phylogram showing the relationships among cytochrome oxidase I (COI) haplotypes of \(Scylla\ tranquebarica\) with numbers in nodes indicate bootstrap values. KM=Marudu Bay; MN=Menumbok; PT=Pitas; SD=Sandakan; SM=Semporna
4. Discussion

This study has successfully defined 11 haplotype from 143 individuals of *S. tranquebarica* collected from five fishing grounds (Kota Marudu, Menumbok, Pitas, Semporna and Sandakan). This is lower compared to the haplotypes of other *Scylla* species; *S. olivacea* (Rosly et al., 2013), *S. paramamosain* (Ma et al., 2011; He et al., 2010), *S. serrata* (Fratini and Vannini, 2002). However, the difference between haplotype numbers of our samples and other previous *Scylla* studies could be due to the differences in sample sources, numbers and the length of COI gene sequences (Ma et al., 2011). The haplotypes (H1 and H2) were commonly encountered in the five fishing grounds which reflects the common origin, extensive gene flows, and migration of crabs between distant localities at a relatively evolutionary time scale (Naim et al., 2012; Horne et al., 2008).

The high haplotype diversity and low nucleotide diversity patterns, which have also been recorded in other marine species such as *S. serrata* (Fratini and Vannini, 2002), *S. paramamosain* (Ma et al., 2011), *S. olivacea* (Rosly et al., 2013), *Tympanoctomys barrerae* (Ojeda, 2010) and *Panulirus homarus* (Senevirathna and Munasinghe, 2014), may indicate that the *S. tranquebarica* populations are mainly composed of closely related individuals with similar haplotypes (Avise, 2000). There are a few factors that may contribute to this pattern, namely are the effect of high maternal effective population size (Lavery et al., 1996) high mutation rate, short existence of haplotypes that only a small number of base pair differences was obtained (Silva et al., 2010; Cassone and Boulding, 2006) and recent demographic events such as bottleneck (Avise, 2000) or a population expansion (Silva et al., 2010). Besides that, non-significant negative values of Tajima D and Fu’s Fs, unimodal pattern of mismatch analysis, significant raggedness and sum of square (SSD) values suggest that the *S. tranquebarica* populations in Sabah coastal waters might have experienced a rapid population expansion in the recent past, which recalls that traces of ancient demographic expansions have been observed in present mud crab population. Many studies have proposed that distribution and demographics of species worldwide (Hewitt, 1996) have been

---

Table 3. Analysis of molecular variance (AMOVA) of mtDNA COI sequence of *S. tranquebarica* in Sabah coastal waters, Malaysia.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum of square</th>
<th>Variance components</th>
<th>Percentage of variation</th>
<th>Fst</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among localities</td>
<td>4</td>
<td>3.369</td>
<td>0.002 Va</td>
<td>0.210</td>
<td>0.00214</td>
<td>0.368</td>
</tr>
<tr>
<td>Within localities</td>
<td>138</td>
<td>109.554</td>
<td>0.794 Vb</td>
<td>99.790</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>142</td>
<td>112.923</td>
<td>0.796</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: df= degree of freedom; Fst= F statistic; p = value of p

Table 4. Mismatch distribution parameters for *Scylla tranquebarica* in five localities

<table>
<thead>
<tr>
<th>KM</th>
<th>MN</th>
<th>PT</th>
<th>SD</th>
<th>SM</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tau</td>
<td>4.201</td>
<td>3.916</td>
<td>4.109</td>
<td>4.113</td>
<td>0.000</td>
<td>3.268</td>
</tr>
<tr>
<td>θ₀</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>θ₁</td>
<td>1.901</td>
<td>2.301</td>
<td>2.275</td>
<td>2.184</td>
<td>999999.00</td>
<td>2001.534</td>
</tr>
<tr>
<td>SSD</td>
<td>0.146</td>
<td>0.162</td>
<td>0.209*</td>
<td>0.213</td>
<td>0.379*</td>
<td>0.222</td>
</tr>
<tr>
<td>Rag</td>
<td>0.409</td>
<td>0.474*</td>
<td>0.587*</td>
<td>0.602*</td>
<td>0.200</td>
<td>0.455</td>
</tr>
</tbody>
</table>

Note: T = units of mutational time; θ₀ = θ before population growth; θ₁ = θ after population growth; SSD = sum of the square deviations between the observed and expected mismatch; P<sub>SSD</sub> = the probability of SSD; Rag = raggedness index; P<sub>Rag</sub> = probability of raggedness; SD = standard deviation; KM = Kota Marudu; MN = Menumbok; PT = Pitas; SD = Sandakan; SM = Semporna. * indicates a significant difference (p<0.05)

---

Figure 3. Unimodal pattern of mismatch analysis of *S. tranquebarica* in Sabah coastal waters, Malaysia.
directly affected by the fall in sea level during the Pleistocene, which likely to have removed shelf habitats suitable for many marine species such as *Perisesarma guttatum* (Silva et al., 2010), coconut crab *Birgus latro* (Lavery et al., 1996), swimming crab *Callinectes bellicosus* (Pfeiler et al., 2005) and *S. olivacea* (Rosly et al., 2013), and is suggested to have shaped the modern population structure of many marine organisms (Imron et al., 2007; Hewitt, 1999; Jacob et al., 2004). These may resulted in the geographical ranges and population size of these *Scylla* mud crab that fluctuated extensively in magnitude, which has also been observed in the pink shrimp (McMillen-Jackson and Bert, 2004). The approaches to enhance the genetic diversity of these mud crab, selective breeding of the species can be one of the solutions for the fisheries management and conservation efforts.

Comparison between mtDNA COI sequences among *S. tranquebarica* in the five fishing grounds revealed no genetic structuring, as evidenced by the absence of genealogical divergence. Other *Scylla* species have been reported to exhibit low genetic differentiation despite a wide geographical distribution (Ma et al., 2011; Rosly et al., 2013). This may be explained by the migration of gravid female mud crabs to the offshore during spawning season (Ikhwanuddin et al., 2011; Jirapunpipat et al., 2009; Hill, 1994) and larval dispersion in the open sea which move along with water currents (Hill, 1994; Ong, 1994). In addition, floods that occur seasonally during Northwest monsoon also affect the movement of crabs from different fishing grounds (Naim et al., 2012). Marine organisms with high dispersal capabilities including other mud crab species, *S. paramamosain* (Ma et al., 2011) and *S. olivacea* (Rosly et al., 2013), typically show high levels of genetic homogeneity and lack genetic structuring.

Knowledge on genetic diversity and population differentiation is important for commercial species (Allendorf and Leary, 1988) such as this mud crab species to ensure sustainable exploitation. This is because the distribution and amount of genetic diversity enable mud crabs to adapt with changes caused by natural or anthropogenic factors (Nelson and Soulé, 1987). However, most of the reduction in genetic diversity may due to various human activities such as in pollution, overfishing exploitation, destruction of habitat, barriers in migration routes and other human developments (Ferguson, 1995; Çiftci and Okumus, 2002). This preliminary study was important to provide basic information on genetic diversity of *S. tranquebarica* species in Sabah coastal waters. The proper management of mud crabs fisheries and aquaculture has to be taken seriously as the demand for these species is increasing.

**5. Conclusions**

Despite the heavily exploitation of mud crab species genus *Scylla* in Sabah coastal waters, little is known about their genetic status. This preliminary study provides basic information on genetic diversity, population structure, and genetic differentiation of mud crabs *S. tranquebarica* in Sabah coastal waters, Malaysia. Results from this study showed that mud crab *S. tranquebarica* in Sabah coastal waters were not genetically structured even in the wide geographical area. Other molecular marker such as microsatellite and allozyme electrophoresis can be used to gain more information on these species. Besides that, more studies on biological, ecological and genetic of *Scylla* mud crabs have to done as the information of these species, especially for *S. tranquebarica*, is limited.

**Acknowledgements**

This study was financially supported by Ministry of Science, Technology and Innovation, under the eScienceFund Research Grant No. SCF0077-22-2012.

**References**


Çiftci, Y and Okumus, I. 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.


