Genetic variability and estimation of effective population sizes of the natural populations of green arowana, *Scleropages formosus* in Peninsular Malaysia

1Shafiqur Rahman, 1Mohd Zakaria-Ismail, 3Pek Yee Tang, 1Mohd Rizman-Idid and 4Sekeran Muniandy

1Bangladesh Fisheries Research Institute, Mymensingh-2201, Bangladesh, 7Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603, Kuala Lumpur, Malaysia, 3Faculty of Engineering and Science, University Tunku Abdul Rahman 4Department of Molecular Medicine, Faculty of medicine, University of Malaya, 50603, Kuala Lumpur, Malaysia,

Shafiqur Rahman, Mohd Zakaria-Ismail, Pek Yee Tang, Mohd Rizman-Idid and Sekeran Muniandy,

**ABSTRACT**

Microsatellite markers and mitochondrial DNA (mtDNA) sequences were used to assess the genetic variability and to estimate genetic effective population sizes of the natural populations of green arowana in Tasik Bera Lake and Endau River. Nine microsatellite loci were polymorphic in both the Tasek Bera and Endau River populations. Results showed evidence of inbreeding in Tasik Bera and Endau River populations. These two populations also exhibited significant heterozygosity excess (p<0.05), suggesting that these two populations have experienced a reduction in effective population size (N_e). Using mtDNA, haplotype diversity was estimated 1.000 for both the populations and nucleotide diversity was low: 0.004 for Tasik Bera and 0.006 for Endau River. The significantly (p< 0.05) negative Fu’s F_s in Tasik Bera (-4.263) indicated this population has undergone historical expansion and this is possibly the cause of the high mitochondrial gene diversity noticed. The historical and contemporary N_e of both Tasik Bera and Endau River populations based on microsatellite data suggested that the size of the arowana populations reduced drastically but these two populations has not experienced genetic bottleneck yet. Genetic differentiation based on microsatellite and mtDNA data between the arowana from Tasik Bera and Endau River was low. This observation may be resulted from isolations of these two populations due to changes of land mass configurations during glaciation and deglaciation and the accompanying lowering and rising of sea levels during the late Pleistocene era.

**Key words:** genetic variability, effective population sizes, natural populations, green arowana

**Introduction**

The Asian arowana (*Scleropages formosus*) is also known as dragonfish, Asian Bonytongue, kelisa or baju rantai. This ancient osteoglossid fish is one of the most expensive and sought after fish in aquatic world. It is related to the other two Australian arowana in the genus *Scleropages* and is the only representative of this species in Southeast Asia [7]. The Asian arowana’s, range was known to encompassed Cambodia, Thailand, Malaysia, Borneo and Sumatra [21]. There are three main colour varieties (green, gold and red) of Asian arowana. These varieties diverged from the late Pliocene to the late Pleistocene eras[28]. In Malaysia the distribution of the green arowana in Peninsular Malaysia is very wide. The green variety is relatively common in some areas such as Tasik Bera in Pahang, Endau

**Corresponding Author**

Shafiqur Rahman, Bangladesh Fisheries Research Institute, Mymensingh-2201, Bangladesh, Phone +8801730302661; E mail: shafiqbfri@yahoo.com
River in Johor[19,25] and Trengganu drainage[6] (Fig. 1).

Genetic analysis of arowana is of conservation interest, because it is a declining species. Habitat alteration, over exploitation and other human activities are threatening arowana populations in Malaysia [33]. Low fecundity rate, oralbrooding habit and openwater spawning further threaten the survivability of the arowana [13]. The species is considered to be vulnerable by the International Union for Conservation of Natural Resources (IUCN)[1,2], and is currently listed in Appendix I of the Convention on International Trade of Endangered Species of Wild Fauna and Flora (CITES) as endangered species.

Genetic assessment of the arowana has been undertaken using various kinds of molecular markers. The capability of Restriction fragment length polymorphism (RFLP) and random amplified polymorphic DNA (RAPD) to study the genetic variability of arowana is low [10]. Analysis using amplified fragment length polymorphism (AFLP) showed significant genetic differentiation between the green, red and gold arowana [35,36]. The use of microsatellites is well documented in fish [24,30,8]. It is a powerful technique to detect genetic variability between and within strains of arowana [35,36].

Molecular markers have been used successfully for monitoring genetic diversity in captive stocks of arowana. However, the lack of consensus regarding the population structure of the wild arowana suggests the need for population genetic studies to provide insight into the status of the wild populations. In this study, we used microsatellite markers and mitochondrial sequences to assess the genetic variability and to estimate genetic effective population sizes of two wild populations of green arowana in Tasik Bera and Endau River. This study will provide a genetic background for the conservation of wild arowana in the Malaysian river systems.

Materials and methods

Samples

Two populations of green arowana were used in this study. Ninety-seven green arowana were collected from Tasik Bera, Pahang (Fig. 1) and thirty-eight green arowana were sampled from the Endau River (Fig. 1). Tasik Bera, in central Peninsular Malaysia, lies within the catchments of Pahang River. This wetland is the first RAMSAR site in Malaysia. The Endau-Rompin area is situated at the boundary of Johor and Pahang States. It encompasses the watershed of the Endau and Rompin rivers. Green individuals were caught by fishermen using net. The fish were then transferred the Ecology and Biology laboratory, Faculty of Science, University of Malaya and were reared individually in aquarium. Scales, opercula and fin clips were collected form live specimens while muscles were collected from dead fish. These samples were preserved in 70% ethanol at 4°C until genomic DNA was extracted.

DNA Extraction

Approximately 20-30 mg of sample was lysed with 260 µl TNES-Urea buffer (Asahida et al., 1996) and 0.2 mg to 0.8 mg of Proteinase K. The mixture was incubated for 12 hours to 15 hours at 55°C following conventional phenol-chloroform extraction.

Microsatellite Analyses

Genetic variation within and between these populations was assessed using nine microsatellite loci. Among these markers four were described by Yue et al. [34], while five loci were isolated by Tang et al.[28]. Details of all the microsatellite loci and PCR conditions are given in Table 1. The Polymerase chain reaction (PCR) procedure was performed on a Hybaid Omnigene thermal cycler in total a volume of 25 µl. Reactions contained 1x PCR buffer (Promega), 1.5 mM MgCl2, 200mM of each dNTP, 0.2 mM of each primer (Table 1), 1U Taq polymerase (Promega) and 20 ng of genomic DNA. Amplifications for the six loci described by Yue et al.[34], were carried out using 4 min of initial denaturation followed by 30 cycles of 30 s of denaturation at 94.5°C, 30 s annealing at the temperature detailed in Table 1, and 30 s extensions at 72°C, with a final extension of 5 min at 72°C. The other five loci were amplified with 3 min of initial denaturation followed by 40 cycles of 10 s of denaturation at 95°C, annealing at specific temperature (Table 2) for 10 s and 30 s extensions at 72°C with a final extension of 5 min. at 72°C. PCR products were run on a 10% nondenaturing polyacrylamide gel (16cm x 20cm) at 250 V for 4-5 hours. A 20-bp DNA marker (Cambrex BioScience) was used to estimate the PCR fragment size. Gels were visualized using a DNA silver staining system (Promega) and analyzed using GelCompar the II Software (Applied Maths).

Mitochondrial DNA Sequencing

Ten individuals from Tasik Bera and 5 individuals from Endau River were chosen for mitochondrial DNA sequence analysis. Three target regions of the arowana mitochondrial DNA were selected for sequencing: 1044bp encompassing the complete ND2, 628 bp of cytb gene region and 72 bp of rRNA-Thr. The positions for the amplifications are shown in Fig. 2 and Fig. 3 respectively.
**Fig. 1:** Major River Drainages and Major Habitats of Arowana (*Scleropages formosus*) in Peninsular Malaysia.

**Fig. 2:** Position of primers for amplification and/or sequencing of ND2 gene. The tRNA genes are indicated by the single letter amino acid codes (I= Isoleucine tRNA; Q=Glutamine tRNA; M= Methionine; W=Tryptophan tRNA; A=Alanine tRNA; N=Asparagine tRNA; C= Cysteine tRNA; Y= Tyrosine tRNA).

**Fig. 3:** Position of primers for amplification and/or sequencing of cyt-b gene. The tRNA genes are indicated by the single letter amino acid codes (E= Glutamic Acid tRNA; T= Threonine tRNA; P= Proline tRNA).

PCR amplification was carried out in a 25-µl containing 1x PCR buffer (Promega), 1.5 mM MgCl₂, 200mM of each dNTP, 0.2 mM of each primer (Table 2), 1U Taq polymerase (Promega) and 20 ng of genomic DNA, on a Hybaid Omnigene Thermocycler. Amplification conditions were as follows: initial denaturation at 94°C for 4 min, followed by 40 cycles of denaturation at 94°C for 40 s, annealing at specific temperatures (Table 2) for 1, extension at 72°C for 2 min followed by a final extension at 72°C for 10 min.
Table 1: Primers sequences of nine microsatellite loci.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Repeat motif</th>
<th>Primer sequences (5’-3’)</th>
<th>GenBank accession nos.</th>
<th>Optimized Annealing Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D92</td>
<td>(GT)13</td>
<td>F AGTCGCCACACCAACCTTCAGG</td>
<td>AF219969</td>
<td>55°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R GCAGCAGTTAACCTGTTACCAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D27</td>
<td>(CA)17</td>
<td>F GGTGTCAGTATGTTAAGCTGAG</td>
<td>AF219958</td>
<td>55°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R TGGAGCACAGGAACTTTTTGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D85</td>
<td>(CA)10</td>
<td>F GTGCCACAGGGGCTTCTGAAAT</td>
<td>AF219967</td>
<td>55°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R GAGGACGACAAAAACGTCATTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D11</td>
<td>(GT)16</td>
<td>F TGTTTCCACCTAGTCACCTAAAAGA</td>
<td>AF219953</td>
<td>55°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R GTTTACGATATGGTGTCAGCTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K10</td>
<td>(CA)20</td>
<td>F GCACCCACGTTACGGGAACTT</td>
<td>AY173130</td>
<td>57°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R AAATATACCTGCTTCAGCTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K13</td>
<td>(CA)12, (CA)16</td>
<td>F GCACCTGTTAACAGCTCTGTC</td>
<td>AY173131</td>
<td>51°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R GATACGGCATGATCTCTGTCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K16</td>
<td>(TG)14</td>
<td>F CAGTTGCTGACCTACCTACAG</td>
<td>AY173136</td>
<td>50°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R AAAGTCCGATGATGAAATA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K27</td>
<td>(CA)18</td>
<td>F CCATACGCTCTTGCTCTTCTTTC</td>
<td>AY173135</td>
<td>50°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R AAGGATCGAGCAGGAGGAAAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K37</td>
<td>(CA)16</td>
<td>F CCATTACGACATGCTTCTTCAG</td>
<td>AY173132</td>
<td>51°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R TGGAATATGTGTGCATCTCCCTC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Primers used for PCR and sequencing (Kumazawa et al., 1999).

<table>
<thead>
<tr>
<th>Mitochondrial gene</th>
<th>Primers</th>
<th>Primer sequences (5’-3’)</th>
<th>Annealing Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND2</td>
<td>L4296</td>
<td>F CGTAGGGATCACTTTGATAG</td>
<td>50°C</td>
</tr>
<tr>
<td></td>
<td>H5540</td>
<td>R CCGCTGAGGGCTTTGAAGGC</td>
<td></td>
</tr>
<tr>
<td>ND2-3</td>
<td></td>
<td>F TCMACCTGACARAAAACT</td>
<td>50°C</td>
</tr>
<tr>
<td>(Internal sequencing primer)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytb</td>
<td>Fyeb-3</td>
<td>F TMGGTCAATGGAATCTGAGG</td>
<td>51°C</td>
</tr>
<tr>
<td></td>
<td>H15990</td>
<td>R GTTTAATTTAGAATGCTGAGTTG</td>
<td></td>
</tr>
</tbody>
</table>

Statistical Analysis

Microsatellites

The observed \((H_o)\) and expected heterozygosities \((H_e)\) of the polymorphic loci of each strain were estimated using the ARLEQUIN version 2.000 [23] software. Hardy-Weinberg equilibrium (HWE) at each locus was assessed using the ARLEQUIN programme. A Markov-chain method with 1000 steps, 1000 dememorizations was used to calculate an unbiased estimate of the P value. The inbreeding coefficient, \(F_{IS}\) [31] was estimated to measure the HWE departures for each population using GENEPOP version 3.1c [22]. The software MICROCHECKER [29] was used to identify the presence of null alleles. MICROCHECKER was also used to test for another source of errors in mis-scoring due to stuttering and allelic dropout.

Population differentiation was measured by calculating pairwise weighted \(F_{ST}\) [31] values over all loci. The ARLEQUIN programme was used for computations. \(R_{ST}\) values were also calculated as sum of squared size differences based on number of repeats [26]. The probability associated with the \(F_{ST}\) and \(R_{ST}\) values was estimated through random permutation procedure (1000 permutation).

Effective population sizes for the two populations were estimated based on two methods: the linkage disequilibrium method [12] to estimate contemporary \(N_e\) and the unbiased expected heterozygosity under Stepwise Mutation Model (SMM) according to Nei[17] to calculate long-term \(N_e\). The contemporary estimates of \(N_e\) and 95% confidence intervals were estimated using the programme NeEstimator [20].

Mitochondrial DNA

All sequences were edited using the programme CHROMAS version 1.4 [16]. The ESEE programme [4] was used to align multiple sequences to identify polymorphic nucleotide sites and assign haplotypes. Sequence variation was assessed estimating nucleotide diversity [17] and haplotypic diversity [18] using the ARLEQUIN package version 2.0[23]. Analysis of molecular variance (AMOVA) [9] was performed on the mtDNA sequence data. The ARLEQUIN was also used to estimate thr \(\Phi_{ST}\) (an mtDNA analogue for \(F_{ST}\); Excoffier [9]. The \(\Phi_{ST}\) analysis was performed using a matrix Tamura and
Table 3: Number of alleles (N), heterozygosity observed (H₀), heterozygosity expected (Hₑ), and inbreeding coefficient (Fᵢₛ) at nine microsatellite loci for two populations of arowana.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Genetic variability indices</th>
<th>Tasik Bera</th>
<th>Endau River</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.918</td>
<td>0.737</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.639</td>
<td>0.512</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>-0.438</td>
<td>-0.448</td>
</tr>
<tr>
<td>D27</td>
<td>N</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.495</td>
<td>0.658</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.427</td>
<td>0.501</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>-0.161</td>
<td>-0.318</td>
</tr>
<tr>
<td>D85</td>
<td>N</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.327</td>
<td>0.419</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>D11</td>
<td>N</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.501</td>
<td>0.498</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>K10</td>
<td>N</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.464</td>
<td>0.447</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.366</td>
<td>0.372</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>-0.297</td>
<td>-0.276</td>
</tr>
<tr>
<td>K13</td>
<td>N</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.320</td>
<td>0.342</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.289</td>
<td>0.373</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>-0.140</td>
<td>0.084</td>
</tr>
<tr>
<td>K16</td>
<td>N</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.127</td>
<td>0.392</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>K27</td>
<td>N</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.515</td>
<td>0.553</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.420</td>
<td>0.448</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>-0.250</td>
<td>-0.239</td>
</tr>
<tr>
<td>K37</td>
<td>N</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.212</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Mean</td>
<td>N</td>
<td>2.56</td>
<td>2.33</td>
</tr>
<tr>
<td></td>
<td>H₀</td>
<td>0.301</td>
<td>0.304</td>
</tr>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.368</td>
<td>0.405</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>0.302</td>
<td>0.311</td>
</tr>
</tbody>
</table>

Table 4: Observed gene diversity (Hₑ, Hardy-Weinberg heterozygosity) and equilibrium gene diversity (Hₑₑ) of two populations of arowana under two-phased model.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Genetic variability indices</th>
<th>Tasik Bera</th>
<th>Endau River</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hₑ</td>
<td>0.639</td>
<td>0.477</td>
</tr>
<tr>
<td>D92</td>
<td>Hₑₑ</td>
<td>0.501</td>
<td>0.390</td>
</tr>
<tr>
<td></td>
<td>Fᵢₛ</td>
<td>0.186</td>
<td>0.269</td>
</tr>
<tr>
<td>D27</td>
<td>Hₑ</td>
<td>0.427</td>
<td>0.473</td>
</tr>
<tr>
<td></td>
<td>Hₑₑ</td>
<td>0.360</td>
<td>0.501</td>
</tr>
<tr>
<td>D85</td>
<td>Fᵢₛ</td>
<td>0.186</td>
<td>0.052</td>
</tr>
<tr>
<td>D11</td>
<td>Fᵢₛ</td>
<td>0.501</td>
<td>0.183</td>
</tr>
<tr>
<td>K10</td>
<td>Hₑ</td>
<td>0.501</td>
<td>0.183</td>
</tr>
<tr>
<td></td>
<td>Hₑₑ</td>
<td>0.360</td>
<td>0.501</td>
</tr>
<tr>
<td>K13</td>
<td>Fᵢₛ</td>
<td>0.501</td>
<td>0.183</td>
</tr>
<tr>
<td>K16</td>
<td>Hₑ</td>
<td>0.501</td>
<td>0.183</td>
</tr>
<tr>
<td></td>
<td>Hₑₑ</td>
<td>0.360</td>
<td>0.501</td>
</tr>
<tr>
<td>K27</td>
<td>Fᵢₛ</td>
<td>0.501</td>
<td>0.183</td>
</tr>
<tr>
<td>K37</td>
<td>Hₑ</td>
<td>0.501</td>
<td>0.183</td>
</tr>
<tr>
<td></td>
<td>Hₑₑ</td>
<td>0.360</td>
<td>0.501</td>
</tr>
</tbody>
</table>

Nei [27]. Fu’s test Fᵢₛ [11] was performed to indicate whether the mutations are neutral or under the influence to selection. Significance of Fᵢₛ was evaluated by using 1,000 random permutations in ARLEQUIN.

Sequences generated from the two-mitochondrial genes and tRNA-Thr were submitted to Genbank (ND2 gene sequences accession numbers DQ 864661-DQ864676, and tRNA- Thr gene sequences accession numbers DQ147903-DQ147905, DQ864687-DQ864699).

Results

Genetic variability

All nine microsatellite loci were polymorphic in both the Tasik Bera and Endau River populations.
The frequencies of null alleles were 0.33 for K37, 0.17 for K16 and K37 in both populations. MICROCHECKER indicated that departure from HWE in these two populations was not due to stuttering or large allele drop out. Null alleles were detected at loci D11 and K37 in the Tasik Bera and Endau River populations. The frequencies of null alleles were 0.33 for D11, 0.17 for K37 in the Tasik Bera population and 0.32 for D11, 0.09 for K37 in the Endau River population. After adjusting allele frequencies for the null alleles, these two loci were still deviated from HWE.

The Endau River has experienced significant heterozygosity excess (p<0.05) (Table 4) but not the Tasik Bera population. The allelic distribution shows a significant ‘mode-shift’ (lack of low frequency alleles) in the former population while the latter showed normal L-shaped distribution (Fig. 4). Using allele frequencies adjusted for null alleles, both of the populations exhibited significant heterozygosity excess (p<0.05) and L-shape in the distribution of allele frequencies (results not presented).

Among 1041 bp of ND2 gene sequenced, 9 sites were variable. Five unique haplotypes were detected (Fig. 5). Of these, 2 unique haplotypes were found in Tasik Bera whereas 3 were found in Endau River. For cytb, 4 variable sites were detected out of 628 aligned nucleotides (Fig. 6). Of 72 aligned tRNA-Thr nucleotides, 2 were variables (Fig. 7). Haplotype diversity was 1.000 for both populations and nucleotide diversity was low, 0.004 for Tasik Bera and 0.006 for Endau River (Table 5).

**Population Differentiation**

$F_{ST}$ of Tasik Bera and Endau River (0.020) was significantly different from zero (p<0.05). This showed that there was pronounced genetic differentiation between these two populations. Significant (p<0.05) pairwise $R_{ST}$ value of 0.021 were observed between these two populations. After adjusting the allele frequencies for null alleles, the $F_{ST}$ and $R_{ST}$ values were still significantly (p<0.05) differentiated between the two populations. Both of the $F_{ST}$ and $R_{ST}$ values were 0.022. The $\phi_{ST}$ of -0.037 did not reveal significant difference between the two populations (Table 6).

**Population Expansion and Effective Population Size**

Fu’s $F_{S}$ of Tasik Bera and Endau River were -4.263 and -0.286 respectively. However only the Fu’s $F_{S}$ of Tasik Bera was significantly (p<0.05) different from zero which indicated population expansion. For microsatellite data, contemporary $N_e$ estimates for Tasik Bera and Endau River were 28.0 (95% confidence interval [CI] = 20.1-39.9) and 14.6 (95% CI = 10.9-19.5) respectively. Historical $N_e$ showed that the Tasik Bera population had an $N_e$ of 1514 which was higher than that of the Endau River population (995). The estimates were slightly changed after correcting the allele frequencies for null alleles. The contemporary $N_e$ for Tasik Bera was 26.7 (95% CI = 19.8-36.6) while $N_e$ for Endau River was 21.4 (95% CI = 16.2-28.5).

**Discussion**

Our microsatellite data suggest the genetic variation within and between the two populations was low. Previous genetic work on the arowana indicated higher genetic variability [34-36]. The discrepancy between our results and these earlier studies may be explained by the unequal representation of populations in different geographical regions and time frame. The stock in previous studies was either pooled from numerous farms or established from wild stock collected in the early 1980s [34,35]. This might not represent the recent natural populations.

The deviation from Hardy Weinberg observed in the present study might result from the territorial behavior of arowana. The entire population may be divided into several subpopulations and only matings between related individuals is likely to happen. The presence of positive $F_{IS}$ values showed further evidence of inbreeding. Although null allele frequencies were taken into consideration, the results did not change. Thus, null alleles seemed unlikely to cause the positive $F_{IS}$ values. The observation of deviation from HWE might result from sampling error. Adult male arowana are territorial mouth brooder and fry can be found near the adults. So, collections in the same location might be composed of siblings produced by a relatively few adults. This could prevent the detection of high genetic variation. Both Tasik Bera and Endau River exhibited heterozygote excess, suggesting that these two populations have experienced a reduction in effective population size ($N_e$). This is because allelic richness is lost faster than heterozygosity after a population reduction. Both populations showed L-shaped distribution of allele frequencies. Thus these two populations have not experienced bottleneck yet because the low frequency alleles are still conserved.
Fig 4: Allele Frequency Distributions from Two Populations of Arowana (Tb; Tasik Bera and Er; Endau River).

Fig 5: Variable sites of the ND2 gene. Dots indicate identity to the sequences of TB19. TB03-10, TB19 and TB79 indicate Tasek Bera Individuals green variety); ER04-05, 06, 16 and 17 indicates Endau River Individuals (green variety).

Table 5: MtDNA sequence variable for green arowana from Tasik Bera and Endau River: number of individuals (n), number of haplotypes (nh), haplotype diversity (h) and nucleotide diversity (π).

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Number of individuals (n)</th>
<th>Number of haplotypes (nh)</th>
<th>Haplotype diversity (h)</th>
<th>Nucleotide diversity (π)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tasik Bera</td>
<td>10</td>
<td>10</td>
<td>1.000</td>
<td>0.004</td>
</tr>
<tr>
<td>Endau River</td>
<td>5</td>
<td>5</td>
<td>1.000</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Table 6: Amova table of genetic variation of green arowana form Tasik Bera and Endau River.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Sum of square</th>
<th>Variance components</th>
<th>% of total variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among populations</td>
<td>1</td>
<td>3.267</td>
<td>-0.154</td>
<td>-3.72</td>
</tr>
<tr>
<td>Within populations</td>
<td>13</td>
<td>55.800</td>
<td>4.292</td>
<td>103.72</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>59.067</td>
<td>4.138</td>
<td></td>
</tr>
</tbody>
</table>

Fixation index \( \Phi_{ST} = -0.037 \)
Fig. 6: Variable sites for cyt-b gene. Dots indicate identity to the sequences of TB19, TB03-10, TB19 and TB79 indicate Tasek Bera Individuals (green variety); ER04-05, 06, 16 and 17 indicates Endau River Individuals; (green variety).

Fig. 7: Variable sites for tRNA-Thr gene. Dots indicate identity to the sequences of TB19, TB03-10, TB19 and TB79 indicate Tasek Bera Individuals (green variety); ER04-05, 06, 16 and 17 indicates Endau River Individuals; (green variety).
Mitochondrial markers which have a 4-fold smaller effective population size than nuclear markers are expected to lose variation faster due to genetic drift [32]. Surprisingly, in the present study, mitochondrial gene diversity ($h$) was high. The significantly negative Fu’s $F_{S}$ in Tasik Bera indicated this population has undergone historical expansion and this might be the cause of the high mitochondrial gene diversity. However, this expansion was followed by a more recent population decline. The historical and contemporary $N_{e}$ of both Tasik Bera and Endau River based on microsatellite data suggest that the size of the arowana populations reduced drastically. Genetic differentiation based on microsatellite and mtDNA data between the arowana from Tasik Bera and Endau River was low. This observation may be result from changes of land mass configurations due to glaciation and deglaciation and the accompanying lowering and rising of sea levels during the late Pleistocene. During glacial maxima (17,000 years before present), extensive land bridges connected Peninsular Malaysia, the Indonesian island and mainland Asia to form a huge continental shelf called Sundaland. The Endau, Pahang, Trenggano and Kelantan Rivers of Peninsular Malaysia joined Thailand’s Chao Phraya River to form the Siam River System. The Pahang and Endau River disconnected 13,000 yr BP when sea level rose to 75 m below present level.

Acknowledgments

The research was partly funded by a research grant from the Ministry of Science, Technology and Innovation (05-01-03-SF0172) and University’s short-term research grants 0139/2002A, 0142/2003A and 0131/2004A. The authors wish to express sincere thank to the Government of Malaysia for granting the sponsorship through the Malaysian Technical Co-operation Programme (MTCP) for the study.

References


