MOLECULAR CHARACTERISATION AND PHYLOGENETICS OF MALAYSIAN GREEN AROWANA (Scleropages formosus) IN PENINSULAR MALAYSIA

M. Rizman-Idid¹, S. Rahman¹, H. A. Aziz, S. Muniandy² & M. Zakaria-Ismail¹

¹Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur
²Department of Molecular Medicine, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur
Asian Arowana (Scleropages formosus)

- Also known as dragon fish or “kelisa” (Malay)
- Highly prized ornamental fish (China, Japan, South East Asia)
- Value as much as USD$ 200 per fingerling and between USD$ 1,000 and 2,000 for 30cm length (Pouyaud et al., 2003)
- In Malaysia, its popularity began as freshwater game fish (Stuart, 1953; Locke, 1956).
- Popular as aquarium fish in 1960s (Yue et al., 2004).
ASIAN AROWANA
(Scleropages formosus)

- The body of the Asian arowana is compressed, dorsal and anal fin are placed on the rear half of the body.
- They have large scales with varieties of colours and two barbels protruding from the lower jaw.
SYSTEMATICS OF ASIAN AROWANA

- Ancient “bony-tongue” family (Osteoglossidae)
- Extant representatives restricted to freshwater (Nelson, 1994).
- Order Osteoglossiformes
  - Family Osteoglossidae (4 genera)
    - *Osteoglossum* (South America)
    - *Arapaima* (South America)
    - *Heterotis* (Africa)
    - *Scleropages* (Southeast Asia, Australia)
• Genus *Scleropages* (3 species):
  – Northern barramundi (*Scleropages jardinii*)
  – Spotted barramundi (*Scleropages leichardti*)
  – Asian arowana (*Scleropages formosus*).

• Distribution pattern:
  – *Scleropages jardinii* (Northern Australia, Southern New Guinea)
  – *Scleropages leichardti* (Eastern Australia)
  – *Scleropages formosus* (Southeast Asia).
COLOUR VARIETIES OF Scleropages formosus

- Several types of *S. formosus* with different colour patterns inhabit separate regions of Southeast Asia.

- Distributed into freshwater habitats during the Pleistocene glacial ages. (Goh & Chua, 1999)

- *S. formosus* have four natural colour varieties (Goh & Chua, 1999, Pouyaud et al., 2003):
  - **Cross Back Golden** - Peninsular Malaysia
  - **Super Red** - West Kalimantan (Indonesia)
  - **Red Tail Golden** - Pekan Baru (Indonesia)
  - **Green** – Peninsular Malaysia, Indonesia, Myanmar, Thailand
• Green and gold varieties only in the Peninsular Malaysia
• None has been reported from East Malaysia, Sabah and Sarawak (Suleiman, 2003).
• **Green variety** has been reported in some area in Peninsular Malaysia:-
  - Lake Tasek Bera (Pahang)
  - Endau River (Johor)
  - Terengganu drainage

• **Gold variety** is believed endemic to Bukit Merah Reservoir, in Perak, Malaysia. (Suleiman, 2003)
• Declining populations due to over-exploitations and destruction of their natural habitats.
• Placement of the species in the Red List of IUCN and Appendix I in CITES.
DISTRIBUTION OF S. formosus IN PENINSULAR MALAYSIA

Site: Bukit Merah Reservoir
Variety: Gold (extinct?)

Site: Lake Tasek Bera
Variety: Green (thriving)

Site: Endau River
Variety: Green (thriving)

Site: Trengganu
Variety: Green (extinct?)
MOLECULAR GENETIC STUDIES

• Molecular genetic techniques have been applied in captive breeding and conservation programs of this endangered species.
• Some of the techniques were based on RFLP, RAPD, AFLP, DNA fingerprinting, and microsatellite analysis.
• However, studies were more focused on captive stocks.
• Very few studies investigate the genetic variation of the natural and wild populations.
OBJECTIVES

• To describe the mitochondrial DNA sequence of natural population of green arowana from Lake Tasek Bera and Endau River (5’-end of partial cyt-\textit{b} and complete \textit{ND2} genes).

• To determine the accumulation of substitutions.

• To reconstruct the phylogenetic tree (\textit{cyt-b} and \textit{ND2} genes).

• To determine whether Tasek Bera population is different from Endau River population.
MATERIALS AND METHODS

• Collection of samples (16 specimens) from:
  - Lake Tasek Bera (Pahang)
  - Endau river (Johor)

• Total of 23 sequences:
  - 15 green (10 Tasik Bera and 5 Endau Rompin)
  - 1 gold (Bukit Merah)
  - 5 outgroup sequences (*Arapaima gigas*, *Osteoglossum bicirrhosum*, *Osteoglossum ferreirai*, *Scleropages jardinii* and *Scleropages leichardti*).
• Peat swamp (Black water)
• Water is dark, tea coloured, and is black as seen by reflected light
• Acidic (pH4.3-6.45)
• Less water flow
• Vegetation:
  - *Pandanus heliocoptus*
  - *Lepironia articulata*
• Riverine system with pools
• Variable flow of water current
• Acidic (pH 5.47-6.97)
• Typical rainforest vegetation
Preserved Scales (90% EtOH)

DNA Extraction
-lysed with TNES-Urea buffer (Asahida et al., 1996)
-modified phenol-chloroform (Tang et al., 2004)

PCR
-ND2 Primers (Kumazawa et al, 1999):
  L4296 (5’-ACGTAGGGATCACTTTGATAG-3’)
  H5635 (5’-AGGTCTTAGCTTAATTAAAG-3’) or
  fNd2-3 (5’-TCMACCTGACARAAACT-3’)
-Cyt B Primers (Kumazawa et al, 1999):
  fCytb-3 (5’-TMGTMCAATGAATCTGAGG-3’)
  H15990 (5’-AGTTTAATTAGAATCYTGCGTTTG-3’)

Purification of PCR Product
Wizard PCR Preps DNA Purification System (Promega)

Direct Sequencing

Sequence & Phylogenetic Analysis
<table>
<thead>
<tr>
<th>Species name</th>
<th>Location (Species ID)</th>
<th>No of specimen</th>
<th>Accession No. (CytB)</th>
<th>Accession No. (ND2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. formosus</em></td>
<td>Endau Rompin</td>
<td>5</td>
<td>DQ864671-5</td>
<td>DQ864684-5</td>
</tr>
<tr>
<td><em>S. formosus</em></td>
<td>Bukit Merah</td>
<td>1</td>
<td>DQ864676</td>
<td>DQ864677-83</td>
</tr>
<tr>
<td><em>S. formosus</em></td>
<td>Tasik Bera</td>
<td>10</td>
<td>DQ864661-70</td>
<td>DQ864686</td>
</tr>
<tr>
<td><em>S. formosus</em></td>
<td>Indonesia</td>
<td>2</td>
<td>AB035234-5</td>
<td>AB035221</td>
</tr>
<tr>
<td><em>S. jardinii</em></td>
<td>Northern Australia</td>
<td>1</td>
<td>AB035236</td>
<td>AB035223</td>
</tr>
<tr>
<td><em>S. leichardti</em></td>
<td>Eastern Australia</td>
<td>1</td>
<td>AB035237</td>
<td>AB035224</td>
</tr>
<tr>
<td><em>O. bicirrhosum</em></td>
<td>South America</td>
<td>1</td>
<td>AB035238</td>
<td>AB035225</td>
</tr>
<tr>
<td><em>O. ferreirai</em></td>
<td>South America</td>
<td>1</td>
<td>AB035239</td>
<td>AB032526</td>
</tr>
<tr>
<td><em>Arapaima gigas</em></td>
<td>South America</td>
<td>1</td>
<td>AB035241</td>
<td>AB032528</td>
</tr>
</tbody>
</table>
### BASE COMPOSITION

Base composition of cyt-b sequences

<table>
<thead>
<tr>
<th>Species</th>
<th>T(U)</th>
<th>C</th>
<th>A</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. formosus</td>
<td>24.3*</td>
<td>35.8*</td>
<td>28.9*</td>
<td>11.0*</td>
</tr>
<tr>
<td>S. jardiniii</td>
<td>26.5</td>
<td>34.9</td>
<td>25.7</td>
<td>12.9</td>
</tr>
<tr>
<td>S. leichardti</td>
<td>26.6</td>
<td>34.9</td>
<td>26.3</td>
<td>12.1</td>
</tr>
<tr>
<td>O. bicirrhosum</td>
<td>29.3</td>
<td>29.8</td>
<td>29.7</td>
<td>11.2</td>
</tr>
<tr>
<td>O. ferreirai</td>
<td>28.7</td>
<td>30.5</td>
<td>30.3</td>
<td>10.5</td>
</tr>
<tr>
<td>A. gigas</td>
<td>30.5</td>
<td>30.6</td>
<td>28.1</td>
<td>10.8</td>
</tr>
</tbody>
</table>

*average of 15 S. formosus (green arowana) sequences
<table>
<thead>
<tr>
<th>Species</th>
<th>T(U)</th>
<th>C</th>
<th>A</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. formosus</td>
<td>20.9*</td>
<td>35.8*</td>
<td>31.7*</td>
<td>11.6*</td>
</tr>
<tr>
<td>S. jardinii</td>
<td>21.8</td>
<td>34.1</td>
<td>29.8</td>
<td>14.3</td>
</tr>
<tr>
<td>S. leichardti</td>
<td>21.8</td>
<td>34.6</td>
<td>32.5</td>
<td>11.1</td>
</tr>
<tr>
<td>O. bicirrhosum</td>
<td>23.2</td>
<td>32.2</td>
<td>34.5</td>
<td>10.1</td>
</tr>
<tr>
<td>O. ferreirai</td>
<td>23.3</td>
<td>32.0</td>
<td>34.8</td>
<td>9.9</td>
</tr>
<tr>
<td>A. gigas</td>
<td>27.2</td>
<td>29.9</td>
<td>32.2</td>
<td>9.7</td>
</tr>
</tbody>
</table>

*average of 15 S. formosus (green arowana) sequences
### SEQUENCE VARIATION

**cyt-b**

<table>
<thead>
<tr>
<th>Type of site</th>
<th>Base frequency</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conserved</td>
<td>599</td>
<td>95.5</td>
</tr>
<tr>
<td>Variable</td>
<td>28</td>
<td>4.5</td>
</tr>
<tr>
<td>Singleton</td>
<td>13</td>
<td>2.1</td>
</tr>
<tr>
<td>Parsimony info</td>
<td>15</td>
<td>2.4</td>
</tr>
</tbody>
</table>

**ND2**

<table>
<thead>
<tr>
<th>Type of site</th>
<th>Base frequency</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conserved</td>
<td>1012</td>
<td>97.2</td>
</tr>
<tr>
<td>Variable</td>
<td>29</td>
<td>2.8</td>
</tr>
<tr>
<td>Singleton</td>
<td>9</td>
<td>0.9</td>
</tr>
<tr>
<td>Parsimony info</td>
<td>20</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Nucleotide frequency of variable sites of Cyt B (627bp) and ND2 (1041bp) from the alignment of 18 *S. formosus* sequences
DISTRIBUTION OF SEQUENCE VARIATION

**cytb**

Most variation found at position 50-250 and 400-600.

Average Pi nucleotide diversity = 0.088

**ND2**

Most variation found at position 80-90, 250-300, 470-510, 625-700 and 1000.

Average Pi nucleotide diversity = 0.09
### AMINO ACID SEQUENCE VARIATION

#### Cyt-b = 209 amino acids

<table>
<thead>
<tr>
<th>Type of site</th>
<th>Amino acid frequency</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conserved</td>
<td>202</td>
<td>96.7</td>
</tr>
<tr>
<td>Variable</td>
<td>7</td>
<td>3.4</td>
</tr>
<tr>
<td>Singleton</td>
<td>6</td>
<td>2.9</td>
</tr>
<tr>
<td>Parsimony info</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

#### ND2 = 347 amino acids

<table>
<thead>
<tr>
<th>Type of site</th>
<th>Amino acid frequency</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conserved</td>
<td>339</td>
<td>97.7</td>
</tr>
<tr>
<td>Variable</td>
<td>8</td>
<td>2.3</td>
</tr>
<tr>
<td>Singleton</td>
<td>5</td>
<td>1.44</td>
</tr>
<tr>
<td>Parsimony info</td>
<td>3</td>
<td>0.9</td>
</tr>
</tbody>
</table>
Amino acid composition of partial cyt-b sequences showed high percentage of leucine (19.56%) and phenylalanine (10.1%) residues.

ND2 sequences also had high percentage of leucine (19.88%) and threonine (15.55%) residues.
Average of Cytochrome b Amino acid composition of 18 S. formosus
AMINO ACID SEQUENCE VARIATION

Average of ND2 Amino acid composition of *S. formosus*
• Transition and transversion rate against the genetic distance (Kimura-2-Parameter) in the cyt-b sequence of 18 S. formosus and 5 outgroup species.

• Within S. formosus - more transitions than transversions

• Between genus/species - high level of transitions and transversions. Members of Osteoglossidae are highly divergent.
Transition and transversion rate against the genetic distance (Kimura-2-parameter) in the ND2 sequence of 18 S. formosus and 5 outgroup species.

- Similar substitution pattern as in cyt-b, but with higher substitution frequencies.
- For closely related species, the transition/transversion ratio approximately 10:1. More divergent species have transitions/transversion ratio of 2:1.
Neighbor Joining tree of 23 cytochrome b sequences (627 bp) using Kimura 2 parameter model (1000 bootstraps) and rooted with A. gigas.

PHYLOGENETIC TREE

S. formosus (green)

S. formosus (red / gold)

S. jardinii
S. leichardti
O. bicirrhosum
O. ferreirai
A. gigas

S.f 2
BM01
S.f 1

0.005
NeighborJoining tree of 23 ND2 sequences (1041 bp) using Kimura 2 parameter model (1000 bootstraps) and rooted with *A. gigas*.
Neighbor Joining tree of combined cytochrome b (627 bp) and ND2 (1041 bp) sequences using Kimura 2 parameter model (1000 bootstraps) and rooted with A. gigas

- Genus Scleropages is monophyletic
- S. formosus is monophyletic
- Green arowana is monophyletic
- No genetic differentiation between Tasek Bera and Endau Rompin populations
- Gold arowana more related to red arowana
DISCUSSION

• Both cyt-b and ND2 genes were highly conserved with low sequence variation among *S. formosus*, especially within the green arowana population.

• Insufficient to distinguish the wild populations of green arowana between Tasek Bera and Endau Rompin.

• However, green arowana were genetically different from red/gold arowana.
Pouyaud et al., 2003 have suggested that colour varieties of 
*S. formosus* may represent different species.

The species are
- *Scleropages formosus* (Green varieties)
- *Scleropages aureus* (Silver)
- *Scleropages macr ocephalus* (Red-tailed Golden)
- *Scleropages legendrei* (Super Red)

His analysis of partial cyt-b sequences (3’end) suggests a close 
relationship between the silver and the green varieties.

Our study suggests that the Malaysian gold variety appears to 
be more related to the red varieties.
• Strong evidence that the colour varieties may have diverged from at least 2 major lineages of Scleropages which are silver-green and red-gold.

• The further splitting of these 4 varieties into species as suggested by Pouyaud et al., 2003 may be premature as most of his measurements were made on cultivated specimens.

• More wild specimens are needed to resolve the splitting of the 4 colour varieties.
• Genetic differentiation was not detected between the wild populations due to the slow mutation rate of the cyt-b and ND2 genes.
• Very few transversions and almost saturated transitions.
• Faster evolving genetic marker may potentially reveal genetic differentiation (eg- D-loop).
• Current study on morphometric and meristic of *S. formosus* revealed no significant difference between wild populations of Tasek Bera and Endau River.
• Variability for cyt-b and ND2 genes were low.
• Selective constraints maintained the function of the genes (selection for hydrophobic amino acid residues).
• Genus *Scleropages* is monophyletic.
• At least 2 major lineages of *Scleropages* were identified, the 1)silver-green and 2)red-gold lineages.
• Gold arowana from Bukit Merah Reservoir was more related to the red arowana from Indonesia than the Malaysian green arowana.
• Present study was unable to detect genetic differentiation between the wild populations of Lake Tasek Bera and Endau River.
• More representative samples are required for future studies.