GENETIC DIVERGENCE OF
*CHANNA* SPECIES
Chapter 2

*Genetic divergence in Channa species*

**Introduction**

Living organisms are incessantly undergoing micro-and macroevolutionary processes both at molecular and organismal levels. Genetic variation in fishes has proven valuable information in aquaculture and fisheries management, for identification of species and stocks, in selective breeding programmes and for estimating contributions to stock mixtures. Moreover, an efficient use of biological resources requires a thorough knowledge of the amount and distribution of genetic divergence within the species considered. Generally, species with greater genetic variability have higher growth rates, developmental stability, viability, fecundity, and resistance to environmental stress and diseases (Carvalho, 1993).
For the evolutionary biologist, the rapid accumulation of sequence data from whole genomes has been a major advance, because the very nature of DNA allows it to be used as a "document" of evolutionary history. Comparisons of the DNA sequences of various genes between different organisms can tell a lot about the relationships of organisms that cannot otherwise be inferred from morphology, or an organism's outer form and inner structure. Because genomes evolve by the gradual accumulation of mutations, the amount of nucleotide sequence difference between a pair of genomes from different organisms should indicate how recently those two genomes shared a common ancestor. Two genomes that diverged in the recent past should have fewer differences than two genomes whose common ancestor is more ancient. For many years, numerous investigators have been trying to better understand the evolution of the vertebrates, especially the phylogeny of the tetrapods or land vertebrates. Most of the information comes from paleontological or anatomic work. Nucleic acid and protein sequence analyses are very recent techniques and are providing new information on the subject (Bernardi et al., 1992)

Recent advances in molecular biology have changed this situation. The dramatic development of molecular genetics since the widespread use of allozyme in the 1970's, and currently exemplified by the Human Genome Project and other equally ambitious undertaking, has laid the ground work for genomics. Firstly, the blueprint of all organisms is written in the DNA and it is now possible to study the evolutionary relationships of organism by comparing their DNA sequences. Secondly, species identification based on molecular genetics is more reliable than
morphological characteristics which are often affected strongly by the environment. Keys to the emergence of genomics were advances in DNA marker technology. The application of DNA marker technology has allowed rapid progress in aquaculture investigations of genetic variability and inbreeding, parentage assignments, species and strain identification.

The development of DNA based genetic markers has had a revolutionary impact on animal genetics. With DNA markers, it is theoretically possible to observe and exploit genetic variation in the entire genome, (Liu and Cordes, 2004). Based on the transmission and evolutionary dynamics genetic markers can be categorized into viz., nuclear DNA and mitochondrial DNA (mtDNA) markers (Park and Moran, 1994) each of which has advantages and disadvantages. In 1960's initial studies involved proteins such as haemoglobin and transferrin, attention quickly turned to enzymatic protein (allozyme) variation in which most subsequent studies have been based on it (Ferguson et al., 1995). One of the main limitations of protein variants as genetic markers is the low level of polymorphism in some species and populations.

Nevertheless, DNA sequencing is the best way to directly estimate genetic variations of specific genes among taxa examined. Thus, we will focus on sequence-based molecular markers or gene regions related only to nuclear ribosomal DNA (rDNA) and mitochondrial DNA (mtDNA). Nuclear DNA genetic markers includes Allozymes, Randomly Amplified Polymorphic DNA (RAPDs), Amplified Fragment Length Polymorphism (AFLPs), and Variable
Number of Tandem Repeats loci (VNTRs: Minisatellites, Microsatellites) are biparentally inherited with potentially widespread utility in a variety of aquaculture endeavors (Liu and Cordes, 2004).

The mtDNA of multicellular animals consists of a closed circular DNA molecule except in some cnidarians, where it consists of one or two linear molecules (Warrior and Gall, 1985; Bridge et al., 1992). Metazoan mtDNAs ordinarily contain 36 or 37 genes, two for ribosomal RNAs (16S rRNA and 12S rRNA), 22 tRNAs encoded in both the heavy (H) and light (L) DNA strand and 12 or 13 subunits of multimeric proteins of the inner mitochondrial membrane (cytochrome oxidase I-III [COI-III], ATP synthase 6 and 8 [A6 and A8], NADH dehydrogenase 1-6 and 4L [ND1-6, ND4L], and cytochrome b apoenzyme [Cytb]) (reviewed by Meyer 1993; Chomyn et al., 1985; Rafael et al., 1995). In addition, there is usually at least one sequence of variable length that does not encode any gene (e.g. control region or A+T rich region). Within the mitochondrial genome, the D-loop (control region) has the highest nucleotide substitution rate (Brown et al. 1993). Even among populations of the same species, the length of this segment is highly variable due to the presence of tandemly repeated sequences and large insertions (Billington and Hebert 1991; Rand 1993; Faber and Stepien 1998; Nesbo et al. 1998).

Mitochondrial DNA is known to evolve much faster than the nuclear genome. The mitochondrial genome of vertebrates has been extensively used over nuclear DNA for resolving phylogenetic relationships at different evolutionary
Genetic divergence in Channa species

depth due to its unique properties, including presence of strictly orthologous genes, lack of recombination, and unique substitution rates (Allard et al., 1992). These unique features along with high copy number of mitochondrial DNA per cell compared to nuclear DNA makes it a more powerful tool than nuclear DNA for identification of unknown biological materials (Wilson et al., 2000; Branicki et al., 2003; Cheng et al., 2001). Although mtDNA generally exhibit faster evolution rate than nuclear genome, certain regions such as 12S rDNA, however, is highly conserved like the nuclear SSU rDNA, which has been employed to illustrate phylogeny of higher categorical levels such as in phyla or subphyla (Ballard et al., 1992). On the other hand, 16S rDNA is more variable than 12S rDNA. Thus, 16S rDNA is usually used for phylogenetic studies at mid-categorical levels such as in families or rare genera (Black and Piesman, 1994; Kambhampati et al., 1996).

Consequently, most of the mitochondrial protein coding genes such as Cytochrome b and COI have been used to examine phylogenetic relationships in relatively lower categorical levels such as in families, genera, species or populations (Hwang and Kim, 1999). Due to the high rate of substitution occurring in the third codon positions (wobble positions) of protein coding genes, the DNA sequences of protein coding genes have frequently been used for species level or population level phylogeny (Navajas et al., 1996). In contrast, because its amino acid sequences are much more conserved than nucleotide sequences, it has been applied to relatively higher levels such as in families or rarely among orders (Frati et al., 1997; Liu and Corder, 2004).
Genetic divergence in Channa species

The control region of mtDNA, the unassigned region, is hypervariable and there exist variations even between individuals. This region is generally variable in size and contains tandem repeats in many cases, like in the IGS of the nuclear rDNA. Thus, this region has been mainly used for phylogenetic studies among species, subspecies, or populations (Zhang and Hewitt, 1997). Uniparental inheritance also reduces the effective population size of mitochondrial genes, which means that variants are fixed more quickly between speciation events (Curole and Kocher, 1999).

It is almost ten years since the polymerase chain reaction (PCR) has been used for amplifying and sequencing of microbial, plant and animal nuclear and mitochondrial genes (Kocher, 1989). Mitochondrial genome study has revealed the tetrapod relationship of Actinopterygii as the out group to Sarcopterygii (Kocher, 1989). Furthermore, mtDNA retains a history of past isolations for a longer period relative to nuclear DNA, but the maternal inheritance of mtDNA limits its ability to provide information on the male component of population and hence detecting the natural hybrids of closely related species.

Presently DNA based techniques are extensively used in species identification (Blackett and Keim 1992; Grossam et al., 1994; Mansfield et al., 1998; Rehbein et al., 1997; Johnson and Kotowski 1996; Cipriano and Palumbi 1997; Palumbi et al., 1996). Most DNA methods reported involve sequence analysis of mitochondrial DNA (mtDNA) (Zehner et al., 1998; Cipriano and Palumbi 1999; Parson et al., 2000). The terminology DNA taxonomy (or DNA
barcoding) denotes two quite independent tasks have been merged (DeSalle et al., 2005): (1) identify and assign specimens to taxonomic groups (species, families) that have been previously described (Savolainen et al., 2005), and (2) predict and classify new taxa using DNA. The relationship between taxonomy and molecular divergence has been already studied by Avise and collaborators (1987). Avise and Johns (1999) demonstrated that comparable taxonomic ranks between animal phyla were not at all equivalent in molecular divergence (Johns and Avise 1998: Avise and Walker 1999).

Therefore, mtDNA variations can be extremely useful for identifying and managing stocks of fish species (Grewe and Hebert, 1988; Billington et al., 1992). Complete mitochondrial DNA sequences have been published for 23 invertebrates; 3 echinoderms, 11 arthropods, 1 annelid, 4 molluscs, 3 nematodes, and 1 cnidarians (Hwang, 1998). In vertebrates, mitochondrial DNA (mtDNA) is present in multiple copies (usually 103–104 copies/cell) and is devoted to the expression of 13 polypeptide components of the respiratory-chain enzyme complexes located in the inner mitochondrial membrane (Gerald et al. 1997). Although they contribute only a small portion of the total set of proteins in these large complexes (comprised of at least 100 polypeptides), all 13 mtDNA-encoded components are presumed to be essential because they are necessary for oxidative phosphorylation in the mitochondrion and hence for the production of cellular ATP.
Generally, the mtDNA gene arrangements in animal groups are usually highly conserved within the phyla; thus, when the rearrangements occur, they are known to be powerful markers for inferring deep evolutionary history (Sankoff et al., 1992; Boore et al., 1995, 1998; Hwang et al., 1998). Compared to the nuclear rDNA, it is more difficult to design universal primers for amplifying specific regions in mtDNA due to a high variability. That is why only a few mitochondrial genes such as 12S rDNA, 16S rDNA, Cytb, COI and ND1 have been employed in phylogenetic studies. In general, 12S and 16S rDNAs are the most conserved regions among the mitochondrial genes. COI is the most conserved among the three cytochrome oxidase coding genes, and also ND1 among the seven NADH dehydrogenase coding genes. Traditionally coding genes like 16s and COI are more extensively used for phylogenetic studies (Ovenden et al. 1997; Sarver et al. 1998; Ptaceck et al. 2001) because they are conservative sequences, whereas the control region has been more used for micro-evolutionary processes at the population level because they are generally the most rapidly evolving regions (Shearer et al. 2002).

Mitochondrial DNA (mtDNA) has many advantages as a marker for phylogenetic analyses (Moore et al. 1999), it is one of the most frequently used markers in molecular systematics (Baker et al. 1997) and in studies for the determination of intraspecific population structure, helping to define conservation units in many endangered species (Goerlitz et al. 2003), or to identify fisheries stocks in commercial species (Cope 2004). The conventional way of identifying species of fishes is by morphological characteristics. In the case of *Channa*,
identification based on morphological descriptions seems to be misleading because these fishes resemble each other. Although Watanbe (2001) indicated that the variegated skin colour pattern, the fin and scale count are the best suited taxonomic key characters for the discrimination of species, these characteristics are not species-specific. In the case of *Channa*, it is difficult to differentiate the species because of the similarity in external morphology. Therefore, the reconstructed Phylogenetic trees based on morphology were controversial due to the complex evolutionary changes in either morphological or physiological characters (Xia et al., 2006).

Hebert et al. (2003) proposed that a single gene sequence would be sufficient to differentiate all, or at least the vast majority of, animal species, and proposed the use of the mitochondrial DNA gene cytochrome oxidase subunit I (Co1) as a global bioidentification system for animals popularly known as DNA Barcoding. Mitochondrial 16S rRNA gene and the protein-coding cytochrome c oxidase subunit I (COI) gene are reasonably well conserved proven to be a robust evolutionary marker for determining inter-specific relationship, and have been sequenced in various invertebrate taxa (Brown, 1985; Bermingham and Lessios, 1993; Brower, 1994, Clary and Wolstenholme, 1985; Baldwin et al., 1996; Hoeh et al., 1996; Boudry et al., 1998, Palumbi and Benzie, 1991; Levinton et al., 1996; Sturmbauer et al., 1996).

However sequence information of these genes of various channids of Indian waters that would be helpful to resolve taxonomic disputes, the present
study carried out the molecular characterization of Channa species using partial sequences of the 16S rDNA and Cytochrome oxidase subunit I genes of the mitochondrial DNA. DNA sequencing provides a tool for the detailed and relatively accurate identification of Channa species, particularly when there are difficulties in differentiating the species based on morphological characteristics. The present study is the first molecular characterization of the freshwater Channa in the India based on Cytochrome oxidase sub unit I and 16S rDNA gene sequences. Hence analysis of partial sequence of the COI and 16S rRNA genes was chosen for this study in Indian channids.
MATERIALS AND METHODS
Collection of samples

The tissue samples (fin clips) of seven species of *Channa* for DNA extraction were collected from the live fish using minimal-invasive methods immediately after capture (Table 3). The tissue samples were stored in sterile eppendorf tubes containing 95% ethanol and sealed with parafilm and kept at room temperature until further analysis. The details of specimen collection sites are as follows:

Table 3: Details of the species used for the study

<table>
<thead>
<tr>
<th>Name of the Species</th>
<th>No. of Fishes collected</th>
<th>Area of collection</th>
<th>Mode of preservation</th>
<th>Codes used</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Channa striatus</em></td>
<td>15</td>
<td>Thamirabharani river, Tirunelveli</td>
<td>95% ethanol</td>
<td><em>C. s</em></td>
</tr>
<tr>
<td><em>Channa punctatus</em></td>
<td>20</td>
<td>Thamirabharani river, Tirunelveli</td>
<td>95% ethanol</td>
<td><em>C. p</em></td>
</tr>
<tr>
<td><em>Channa marulius</em></td>
<td>15</td>
<td>Pampa river, Thakazhy, Kerala.</td>
<td>95% ethanol</td>
<td><em>C. m</em></td>
</tr>
<tr>
<td><em>Channa aurantimaculata</em></td>
<td>10</td>
<td>Brahmaputra river basin, Assam.</td>
<td>95% ethanol</td>
<td><em>C. a</em></td>
</tr>
<tr>
<td><em>Channa gachua</em></td>
<td>10</td>
<td>Brahmaputra river basin, Assam.</td>
<td>95% ethanol</td>
<td><em>C. g</em></td>
</tr>
<tr>
<td><em>Channa bleheri</em></td>
<td>10</td>
<td>Brahmaputra river basin, Assam.</td>
<td>95% ethanol</td>
<td><em>C. b</em></td>
</tr>
<tr>
<td><em>Channa diplogramme</em></td>
<td>15</td>
<td>Pampa river, Thakazhy, Kerala.</td>
<td>95% ethanol</td>
<td><em>C. mp</em></td>
</tr>
</tbody>
</table>
DNA Isolation

The total genomic DNA of the seven Channa species (all samples) was isolated using the salting out protocol (Miller et al., 1998). The concentrations of the stock solutions as well as working solutions and the protocols followed are given below:

Stock Solutions

0.5M Tris Cl (pH-8.0)
Tris base - 3.028g
Distilled water - 40ml
Adjust pH to 8.0 using HCl.
Make up the volume to 50ml, autoclave and store at 4°C.

0.5M Tris Cl (pH-8.3)
Take above composition, adjust pH to 8.3 using HCl, make up to 50ml, autoclave and to store at 4°C.

0.5M EDTA (pH-8.0)
EDTA - 9.31g
Distilled water - 40ml
Adjust pH to 8.0 using NaOH.
Make up the volume to 50ml, autoclave and store at 4°C.

10mM Tris Cl (pH-7.5)
Tris base - 0.03028g
Materials and Methods

Distilled water - 20ml

Adjust pH to 7.5 using HCl.

Make up the volume to 25ml, autoclave, and store at 4°C.

**RNAase buffer**

10mM Tris Cl (pH 7.5) - 10μl
15mM NaCl (0.8766mg/ml) - 30μl
Distilled water - 960μl

Autoclave and store at 4°C.

**TBE buffer 10X (pH-8.0)**

Tris Cl (0.9M) - 10.8g
Boric acid (0.9M) - 5.5g
EDTA (0.5M) - 0.75g

Make up the solution to 100 ml with double distilled water.

Autoclave and store at 4°C

**Bromophenol blue**

0.5% Bromophenol blue - 500mg

Make up the solution to 100 ml with double distilled water.

Autoclave and store at 4°C

**Working Solutions**

**Solution 1:**

50mM Tris-HCl (pH 8.0)
Materials and Methods

20mM EDTA (pH8.0)

2%SDS

Solution 2:

Saturated NaCl solution (6M)

Proteinase K (20mg/ml)

Proteinase K - 10mg

Autoclaved distilled water - 500μl

Dissolve proteinase K in distilled water and store at -20°C.

TE buffer

Stock 0.5M Tris Cl (pH-8.0) - 2.0ml (10mM)

Stock 0.5M EDTA (pH-8.0) - 0.2ml (1mM)

Make up the solution to 100 ml with distilled water, autoclave it, store at 4°C.

RNAase

RNAase - 10mg

RNAase buffer autoclaved - 1ml

Dissolve, heat at 100°C for 15mts in boiling water.

Cool at room temperature and store at 4°C.

Bromophenol blue

Stock 0.5% Bromophenol blue - 100μl

Autoclaved 1x TBE - 700μl

Glycerol (mol. grade) or Ficol - 200μl
Reagents for Agarose Gel Electrophoresis

1. Agarose
2. 1x TBE
3. Ethidium bromide

DNA Extraction Protocol (Salting out method)

Tissue kept in alcohol was washed with Tris buffer (pH8.0) by spinning.

1. Tissue samples were taken in 1.5ml tube & added 500μl Solution 1.
2. Homogenized the tissue with sterile homogenizer.
3. Added 5μl of Proteinase K (20mg/ml)
4. Incubated at 55°C in water bath for 2 hours (with occasional mixing).
5. Chilled on ice for 10 minutes.
6. Added 250μl Solution 2 and inverted several times to mix.
7. Chilled on ice for 5 minutes.
8. Centrifuged at 8000 rpm for 15 minutes.
9. Carefully collected about 500μl clear supernatant into a new-labeled 1.5ml tube.
Materials and Methods

[If supernatant was not clear, again centrifuged at 8000 rpm.]

10. Added twice the volume (i.e., 1ml) of 100% Molecular biology grade Ethanol to precipitate the DNA

11. Kept at -20°C for overnight.

12. Next day, again centrifuged at 11000 rpm for 15 minutes.

13. Removed the supernatant carefully.

14. Rinsed the DNA pellets in 500μl of ice-cold 70% ethanol.

15. Centrifuged at 11000 rpm for 5 minutes.

16. Carefully removed the supernatant and partially dried with lid off either at room temperature.

17. Resuspend partially dried DNA in 50-200μl of TE buffer depending on size of pellet (100μl average) by gently pipetting sample with wide-bore filter tip until dissolved.

18. The dissolved DNA was stored at -20°C.

**Determination of quality and quantity of DNA isolated**

The quality and concentration above isolated DNA samples were analysed on 0.7% Agarose Gels.

**Casting of 0.7% Agarose Gel:**
Setting of casting unit

Set unit according to manufacturer's instructions.

0.7% Agarose Gel

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agarose</td>
<td>210mg</td>
</tr>
<tr>
<td>10X TBE</td>
<td>3 ml</td>
</tr>
<tr>
<td>Distilled water</td>
<td>27 ml</td>
</tr>
</tbody>
</table>

**Protocol**

Agarose solution was boiled to dissolve agarose.

↓

It was cooled down to approximately 50°C.

↓

1.5μl (1mg/ml) of Ethidium bromide was added.

↓

It was then poured into gel casting plate with already adjusted gel comb.

↓

Allowed it to solidify at room temperature for 1/2hr for the gel to set, before loading the gel.

**Running agarose gel:**

Used cold 1X TBE as gel running buffer.

↓
Loaded 2μl of DNA solution in the wells along with known quantity of DNA.

Run at 90 V for 30 min.

Observed in ultraviolet light.

**PCR amplification and Sequencing**

Approximately 600 base fragments of the mitochondrial 16S rRNA and Cytochrome c Oxidase subunit 1 (COI) genes were amplified from five samples each of the seven species using 1μl of the DNA extract as a template, using the following primers L2510 (5’CGC CTG TTT ATC AAA AAC AT 3’) and H3080 (5’ CCG GTC TGA ACT CAG ATC ACG T 3’) (Palumbi et al., 1991) for 16S rRNA gene and the Cytochrome c Oxidase subunit 1 (COI) gene was amplified using two sets of COI primers (Ward et al., 2005), FishR2-(5’ TCA ACC AAC CAC AAA GAC ATT GGC AC 3’), FishR1- (5’ TAG ACT TCT GGG TGG CCA AAG AAT CA 3’), FishF2-(5’ TCG ACT AAT CAT AAA GAT ATC GGC AC 3’), FishF1- (5’ ACT TCA GGG TGA CCG AAG AAT CAG AA 3’). PCR reactions were carried out in PTC 200 gradient thermal cycler (M.J. research, Inc., Watertown, Massachusetts, USA). The amplifications were performed in 25μl reactions containing 1x assay buffer (100mM Tris, 500mM KCl, 0.1% gelatin, pH 9.0) with 1.5mM MgCl₂ (Genei, Bangalore, India), 10 p moles/μL of primer mix, 10 mM dNTPs (Genei, Bangalore, India), 1.5 U Taq DNA polymerase and 20 ng of template DNA. To check DNA contamination, a negative control was set up.
omitting template DNA from the reaction mixture. The reaction mixture was initially denatured at 95°C for 5 minutes followed by 29 cycles (denaturation at 94°C for 45 seconds, annealing at 50°C (for 16S) or 54°C (for COI) for 30 seconds and 72°C for 45 seconds). The reaction was then subjected to a final extension at 72°C for 5 minutes.

The PCR products and 100bp DNA ladder (Genei, Bangalore, India) were electrophoresed on a 1.5% agarose gel (in 1 X TBE) for 30 minutes. The gel was then visualized under a UV transilluminator and documented using Image Master VDS (Pharmacia Biotech, USA). The remaining PCR product was cleaned using GeNei™ Quick PCR purification kit (Genei, Bangalore, India). The cleaned up PCR products were sent to the sequencing facility for sequencing in AB 3130 x 1 genetic analyser (Applied Bio systems).

Data analysis

The DNA sequences of *Channa aurentimaculata*, *C. bleheri*, *C. diplogramme*, *C. gachua*, *C. marulius*, *C. punctatus* and *C. striatus* were edited using BIOEDIT (Hall, 1999) and aligned using CLUSTAL X as implemented in BIOEDIT sequence alignment editor version 7.0.5.2. The extent of sequence difference between individuals was calculated by averaging pairwise comparisons of sequence difference along all individuals. Pairwise evolutionary distances among the haplotypes were determined by the Kimura-2-Parameter (K2P) method (Kimura, 1980) using the software program MEGA version 3.1 (Kumar *et al.*, 2004). Phylogenetic relationships among the haplotypes were
Materials and Methods

derived using neighbor-joining (NJ) and maximum-parsimony (MP) using MEGA. The data set was bootstrapped 1000 times to estimate the internal stability of the tree nodes. A representative species from the Badidae, Osphronemidae, Carangidae and Cyprinidae families was chosen as outgroups. The Percoidei and Anabantoidei have been postulated to have a close phylogenetic relationship to the Channoidei, which contains the Channidae as the only family member (Gosline, 1971, Rosen and Patterson, 1990, Britz, 1997, Ruber et al., 2004, Xia Li et al., 2006).
RESSURECTION OF CHANNA
DIPLOGRAMME
Resurrection of Channa diplogramme

Day (1865a, b) described Ophiocephalus diplogramme based on one specimen, about 42 mm in length, collected in October 1863 near the mouth of the Cochin river in the port city of Cochin in Southwestern India and named it as Malabar snakehead. The colouration of this juvenile specimen matched with that of the juveniles of another species Channa micropeltes originally described by Cuvier and Valenciennes, 1831 from Java, Indonesia. These two species were easily confused due to the similarity in the colour pattern of juveniles (external morphology, including number of fin rays, scales and colouration of the live juvenile specimen) and Walter and James (2004) was of the opinion that C. micropeltes was introduced to Kerala during mid 1800's by traders, because moving an air breathing fish to a new location even on sailing ships is an easy task and Cochin has been a major trading port for several centuries. However, on close examination of the Malabar snakehead, differences were observed in many characters from that of the original Southeast Asian C. micropeltes. Hence, the
study was taken up to accurately identify and validate the status of Malabar snakehead, which is endemic to Kerala.

Channids are well known for the fact that the colour patterns of their young are very different from the adults (Day, 1878-1888; Weber and de Beaufort, 1922; Smith, 1945). Many of the synonyms reflected a lack of knowledge about such changes. The reasons for these drastic changes in colour and pattern remain unknown, but are possibly associated with their schooling behaviour as young fishes. All the young have longitudinal stripes which are longer in Channa micropeltes and shorter in C. gachua (Smith, 1945). Morphology is a complex and non-neutral marker for taxonomical purposes. Consequently, morphological taxonomy could lead to under- or over-estimation of biodiversity.

With today’s technology for production of molecular sequences, cryptic speciations have been reported from all phyla and seem to be a frequent bias associated with morphological taxonomy. Some authors (Hebert et al., 2003; Tautz et al., 2003; Blaxter, 2004) suggested the use of DNA in taxonomy to overcome those “impediments” and to face with the enormous quantity of living forms. DNA, in contrast, being heritable with a clear mode of transmission, can potentially provide more accurate resolution of taxonomic ambiguity and evolutionary relationships among such taxa (Avise, 1994). Thus, phylogenetic relationships among the specimens (colour morphs of C. diplogramme) with C.
Resurrection of Channa diplogramme

*micropeltes* were constructed based on the partial sequence data of the mtDNA 16S rRNA (Fig. 4) and barcoded using COI gene.

*Channa diplogramme* showed similarity to juvenile specimens of *C. micropeltes* in colouration that is known to change with growth in Channids (Fig. 2). The specimens caught from Thakazhy, Kerala in juvenile stage were reared in earthen ponds at Centre for Aquaculture Research and Extension (CARE) were periodically examined and compared with the individuals of the same length and weight caught from the wild. *C. diplogramme* exhibited profound colour changes during its life cycle, a broad black band passed through the eye direct to the upper half of caudal fin and a second black line commenced at the angle of the mouth, and proceeded to lower half of caudal; and as the fish grew, the bands gradually broke up into several (4–6) white blotches and in adults the upper half of the body exhibiting dark colouration (Fig. 1). Somewhat, changes in body colour and patterns between young fishes and adults of *C. micropeltes* had been documented also by Day (1878-1888) and Tweedie (1949) (Fig. 1).

Of the 28 Channid species that are presently considered as valid (Bonou and Teugels, 1985; Ng et al., 1999; Musikasinthorn, 2000), 7 species possess a patch of scales on the gular portion (Fig. 3); these are 4 species of the Asian genus...
Fig. 1: Different stages of *C. diplogramme*

1A - Juvenile stage 1 of *C. diplogramme*.
1B - Juvenile stage 2 of *C. diplogramme*.
1C - Intermediate stage of *C. diplogramme*.
1D - Intermediate stage of *C. diplogramme*.
1E - Adult stage of *C. diplogramme*.
1F - Adult stage of *C. diplogramme*.

*C. micropeltes* figured by Day (1878-1888) and Tweedie (1949)

1G: Schematic drawing of *C. micropeltes* juvenile.
1H: Schematic drawing of *C. micropeltes* intermedia.
1I: Schematic drawing of *C. micropeltes* adult.

Photograph of *C. micropeltes*
Channa, i.e., C. bankanensis, C. lucius, C. micropeltes, and C. pleurophthalmus, and all species of the African genus Parachanna, i.e., P. africana, P. insignis, and P. obscura (Musikasinthorn and Taki, 2001; Zhang et al., 2002). It is interesting to note that none of the Indian Channids (Channa punctatus, C. marulius, C. diplogramme, C. bleheri, C. striatus, C. gachua, and C. aurentimaculata) possess the gular scales.

The morphomerestic comparisons of C. diplogramme with the absence of gular scales indicated that they only exhibit similarity with C. micropeltes in colouration during their juvenile stage. All the examined meristic and morphometric characters of the C. diplogramme fall out of the ranges for C. micropeltes, excepting slight changes in the scale count (Table: 4).

However, Malabar snakehead exhibited marked differences in the following characters from that of the Southeast Asian C. micropeltes: Absence of gular scales:- In all the specimens examined gular scales were absent in Malabar snakehead, while the original C. micropeltes from Southeast Asia was reported to have a patch of 15 gular scales (Musikasinthorn et al., 2001). In addition, none of
the channids reported from the Indian sub-continent had gular scales. Differences in lateral line scale count:- The lateral line scale count in all the colour morphs of Malabar snakehead from Kerala was in the range of 103 to 105, whereas the same in *C. micropeltes* from Southeast Asia exhibited 83 to 94 lateral line scales (Kottelat 2001). Partial sequence information of 16S rRNA gene between *C. diplogramme* and *C. micropeltes* showed 3.2 per cent divergence.

The partial 16S rRNA sequence alignment (488 bp) among the specimens of the malabar snakehead (*C. diplogramme*) (CD1, CD2, CD3, CD4, CD5 and CD6) was highly conserved and did not reveal any significant variation and no genetic distance was observed. When compared with sequence of *C. micropletes* from National Centre for Biotechnology Information (NCBI) a genetic divergence (K-2 Parameter) value of 3.2 per cent was observed. Partial sequence information of COI gene of *C. diplogramme* was generated, however, genetic divergence value could not be estimated due to the absence of sequence data of *C. micropeltes* in NCBI genbank.

The amount of pairwise sequence divergence within each of *C. diplogramme* was found to be zero. The phylogenetic affinities of *C. diplogramme* with the other Indian channid species are presented in the figures 2.1 and 2.2. Based on the above observations, the status of Malabar snakehead *Channa diplogramme* is revalidated. The mitochondrial data unanimously support genetic similarity/relatedness. Genetic data strongly support that the species status of colour morphs of *C. diplogramme*, which was until recently been confused as
Resurrection of *Channa diplogramme*

*C. micropeltes,* was found to be genetically dissimilar to *C. micropeltes.* Thus the mitochondrial sequence based approach was useful for the accurate identification of the species even from the early stage.

Our results demonstrated that the current taxonomy is in global agreement with molecular differentiation (Xia *et al,* 2006). It was found that mitochondrial DNA sequence variations are suitable to delimit what Channid taxonomists have called species. A molecular threshold was thus developed to help draw up the boundaries of *Channa diplogramme* from *C. micropeltes.* The present records (morphological and molecular evidence) thus confirming the presence of *C. diplogramme* in Kerala; 142 years after Day’s report. The specimen earlier suspected as *C. micropeltes* in India is being resurrected as *C. diplogramme* on the basis of the above cited molecular and morphomerestic data. It is also noteworthy that site of collection of specimens of *C. diplogramme* in the present was in close proximity to the type locality (Cochin) as described by Day (1865). Five of the specimens collected were deposited in the Zoological Survey of India (ZSI) (Topotype accession number: Topotype ZSI SRS F.8109 and Topotype ZSI SRS F.8110) and the others are kept in live condition for further studies at Centre for Aquaculture Research and Extension (CARE).
Table 4: Comparative morphomeric characters of *Channa diplogramme* from Kerala and *C. micropeltes* from Southeast Asia

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Lateral line scales</td>
<td>95 - 110</td>
<td>83-94</td>
<td>104</td>
</tr>
<tr>
<td>Pre pelvic scales</td>
<td>22</td>
<td>----</td>
<td>22</td>
</tr>
<tr>
<td>Dorsal fin rays</td>
<td>43 – 46</td>
<td>----</td>
<td>43</td>
</tr>
<tr>
<td>Pectoral fin rays</td>
<td>15</td>
<td>----</td>
<td>16</td>
</tr>
<tr>
<td>Ventral fin rays</td>
<td>6</td>
<td>----</td>
<td>6</td>
</tr>
<tr>
<td>Anal fin rays</td>
<td>27 - 30</td>
<td>----</td>
<td>27</td>
</tr>
<tr>
<td>Caudal fin rays</td>
<td>15</td>
<td>----</td>
<td>16</td>
</tr>
</tbody>
</table>

Fig 4: Neighbor joining tree of 16S rRNA partial sequence of *C. diplogramme* compared with a sequence of *C. micropeltes* from NCBI.
Results

Isolation of DNA

The DNA was isolated from samples of *Channa aurentimaculata*, *C. punctatus*, *C. gachua*, *C. diplogramme*, *C. bleheri*, *C. striatus*, and *C. marulius*, by salting out method. The extracted DNA was electrophoresed through 0.7% agarose gel containing ethidium bromide (Fig. 5).

![Representative gel showing Genomic DNA of different Channa species.](image)
Quantification and Purity of DNA

The quantity and quality of the extracted DNA was checked by taking the optical density (OD) using spectrophotometer at wavelength 260nm and 280nm. Most of the extracted DNA had very high concentration; therefore, the samples were diluted with sterile double distilled water to get appropriate concentration for PCR reactions. The OD ratio (260nm to 280nm) of the each sample was in between 1.7 and 1.9. Therefore, the samples were in pure condition without contamination of protein and RNA.

Mt DNA Analysis

Sequence characteristics and genetic divergence based on 16S rRNA (Fig: 6, 8.) and COI gene (Fig: 7, 8.). A total of 513 base pairs of aligned sequences of 16S rRNA gene and 560 nucleotides of COI were studied (Appendix 1) in seven Indian *Channa* species, it was attempted to include five individuals per species. When multiple individuals were available for the same taxon partial sequence genes of the second individual were done to show good correspondence of the nucleotide sequence within each taxonomic unit identified under the current classification.

Sequences representing each species were submitted to the nucleotide database of GenBank (www.ncbi.nlm.nih.gov) with the following GenBank accession numbers: EU342175, EU342176, EU342177, EU342178, EU342179, EU342180, EU342181, EU342182, EU342183, EU342184, EU342185, EU342186, EU342187, EU342188, EU342189, EU342190, EU342191, EU342192, EU342193,
Fig. 6.: mtDNA profile of different *Channa* species with 16S rRNA primer (lane 1-C. striatus, lane 2-C. punctatus, lane 3-C. aurentimaculata, lane 4-C. marulius, lane 5-C. bleheri, lane 6-C. gachua, M – marker).

Fig. 7.: mtDNA profile of 6 species of *Channa* with COI primer. (lane 1-C. striatus, lane 2-C. punctatus, lane 3-C. aurentimaculata, lane 4-C. marulius, lane 5-C. bleheri, lane 6-C. gachua. Lane 7-100bp ladder, lane 8-C. striatus, lane 9-C. punctatus, lane 10-C. aurentimaculata, lane 11-C. marulius, lane 12-C. bleheri, lane 13-C. gachua).

Fig. 8.: mtDNA profile of 2 individuals of *C. diplogramme* with 16S rRNA primer (lanes 1 & 2) and COI primer (Lanes 4 & 5) (lanes 1 - CD1 & 2 - CD2, lanes 4 - CD1 & 5 - CD2, M – marker 100 bp DNA ladder, C – negative control [reaction mixture without template DNA]).
Results

EU342194, EU342195, EU342196, EU342197, EU342198, EU342199, EU342200, EU342201, EU342202, EU342203, EU342204, EU342205, EU342206, EU342207, EU342208, EU342209, EU342210.

In general the nucleotide sequences were more variable in protein coding regions than in rRNA gene regions. Alignment of sequence was simple and unambiguous for both the genes. Sequence were aligned using BIOEDIT sequence alignment editor version 7.0.5.2 (Hall, 1999) and were used to estimate genetic divergence values and for constructing phylogenetic trees (Neighbor Joining ‘NJ’ and Maximum Parsimony ‘MP’) using MEGA 3.1 (Kumar et al., 2004). These two analytical methods (NJ and MP) provided nearly identical trees, and many clades were established with strong tree-support values (Fig. 9 - 12).

Nucleotide-sequence divergences were calculated using the Kimura-two-parameter (K2P) model, the best metric when distances are low (Nei and Kumar 2000) as in this study. ‘Indels’ (insertions/deletions) or ‘gaps’ during aligning the sequences were treated as “fifth character”. Kimura-2-parameter method was identified as the mutational model in the present study to accommodate higher rate of transitions that is generally observed in teleosts.

The sequence data of Badis badis (Family: Badidae; Accession No. AY662746 for 16S and AY662699 for COI), Trichopsis vittata (Family: Osphronemidae; Accession No. AF519658) for 16S, Polymixia japonica (Family: Polymixiidae; Accession No. NC 002648) for 16S and COI and Gobicypris rarus (Family: Cyprinidae; Accession No. AY899292) for COI from NCBI was used as
Fig. 9.: Maximum Parsimony phylogram showing the relationship between 16S rRNA haplotypes from Channa species.

Fig. 10.: Neighbour-joining phylogram showing the relationship between 16S rRNA haplotypes from Channa species.
Fig. 11.: Maximum Parsimony phylogram showing the relationship between COI haplotypes from Channa species.

Fig. 12.: Neighbour-joining phylogram showing the relationship between COI haplotypes from Channa species.
Results

outgroups. The phylogenetic trees supported monophyly of the channid species in relation to the outgroup species used in this study.

Sequencing of the 16SrRNA gene produced an average of 513 (range 497-515) nucleotide base pairs per taxon. Multiple alignments resulted in a consensus length of 513 sites including base pairs and gaps. This alignment consisted of 513 positions, of which 160 were variable and 124 were parsimony informative.

The average nucleotide composition (Table 5.) for 16S was 30.3% A, 21.8% T, 25.3% C, and 22.7% G.

Sequence divergence (Table 6.) for 16S gene within Channa ranged from 3.09% (C. aurenticulata and C. gachua) to 13.66% (C. punctatus and C. striatus). The outgroup sequence differed by 9.0%–17.8% from the Channa sequences. All species differed by more than 3% from the averages. Observed transitions and transversions were plotted for all comparisons Appendix 2. The plotted

| Table 5: Nucleotide Frequencies 16S |
|---------|---------|---------|---------|---------|
| C. striatus 1 | 21.3 | 25.2 | 30.4 | 23.1 |
| C. striatus 2 | 21.1 | 25.2 | 30.8 | 22.9 |
| C. punctatus 2 | 21.2 | 25.9 | 30.9 | 22.0 |
| C. punctatus 1 | 21.2 | 25.9 | 30.9 | 22.0 |
| C. diplogramme 1 | 21.6 | 25.2 | 30.8 | 22.4 |
| C. diplogramme 2 | 21.6 | 25.2 | 30.8 | 22.4 |
| C. marulius 2 | 21.4 | 25.4 | 30.6 | 22.6 |
| C. marulius 1 | 21.4 | 25.4 | 30.8 | 22.4 |
| C. bleheri 2 | 21.8 | 26.2 | 29.0 | 23.0 |
| C. bleheri 1 | 21.8 | 26.4 | 29.0 | 22.8 |
| C. aurenticulata 1 | 22.1 | 26.0 | 29.6 | 22.3 |
| C. aurenticulata 2 | 22.1 | 25.8 | 29.8 | 22.3 |
| C. gachua 2 | 22.6 | 25.0 | 30.6 | 21.8 |
| C. gachua 1 | 22.4 | 25.2 | 30.6 | 21.8 |
| P. japonicus | 22.5 | 24.1 | 30.6 | 22.9 |
| B. badis | 20.7 | 23.8 | 32.5 | 23.0 |
| T. vittata | 23.3 | 24.1 | 29.7 | 22.9 |
| Avg. | 21.8 | 25.3 | 30.3 | 22.7 |

73
Table- 6. Genetic distance values of Channa species using 16S gene.

<table>
<thead>
<tr>
<th></th>
<th>C.s1</th>
<th>C.s2</th>
<th>C.p2</th>
<th>C.p1</th>
<th>C.dip1</th>
<th>C.dip2</th>
<th>C.m1</th>
<th>C.m2</th>
<th>C.b1</th>
<th>C.b2</th>
<th>C.a1</th>
<th>C.a2</th>
<th>C.g1</th>
<th>C.g2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.s1</td>
<td>*****</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.s2</td>
<td>0.0040</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.p2</td>
<td>0.1366</td>
<td>0.1315</td>
<td>*****</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.p1</td>
<td>0.1366</td>
<td>0.1315</td>
<td>0.0000</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.dip1</td>
<td>0.0659</td>
<td>0.0660</td>
<td>0.1320</td>
<td>0.1320</td>
<td>0.0000</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.dip2</td>
<td>0.0659</td>
<td>0.0660</td>
<td>0.1320</td>
<td>0.1320</td>
<td>0.0000</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.m1</td>
<td>0.0635</td>
<td>0.0635</td>
<td>0.1308</td>
<td>0.1308</td>
<td>0.0679</td>
<td>0.0679</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.m2</td>
<td>0.0635</td>
<td>0.0635</td>
<td>0.1284</td>
<td>0.1284</td>
<td>0.0657</td>
<td>0.0657</td>
<td>0.0020</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.b1</td>
<td>0.0983</td>
<td>0.0985</td>
<td>0.1243</td>
<td>0.1243</td>
<td>0.1121</td>
<td>0.1121</td>
<td>0.1030</td>
<td>0.1030</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.b2</td>
<td>0.1008</td>
<td>0.1009</td>
<td>0.1269</td>
<td>0.1269</td>
<td>0.1147</td>
<td>0.1147</td>
<td>0.1055</td>
<td>0.1055</td>
<td>0.0020</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.a1</td>
<td>0.1075</td>
<td>0.1076</td>
<td>0.1171</td>
<td>0.1171</td>
<td>0.1167</td>
<td>0.1167</td>
<td>0.1030</td>
<td>0.0980</td>
<td>0.0395</td>
<td>0.0417</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.a2</td>
<td>0.1050</td>
<td>0.1051</td>
<td>0.1146</td>
<td>0.1146</td>
<td>0.1141</td>
<td>0.1141</td>
<td>0.0999</td>
<td>0.0955</td>
<td>0.0417</td>
<td>0.0438</td>
<td>0.0020</td>
<td>*****</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.g1</td>
<td>0.1026</td>
<td>0.1028</td>
<td>0.1074</td>
<td>0.1074</td>
<td>0.1122</td>
<td>0.1122</td>
<td>0.1002</td>
<td>0.1002</td>
<td>0.0372</td>
<td>0.0394</td>
<td>0.0309</td>
<td>0.0330</td>
<td>*****</td>
<td></td>
</tr>
</tbody>
</table>
| C.g2 | 0.1049 | 0.1050 | 0.1051 | 0.1051 | 0.1145 | 0.1145 | 0.1025 | 0.1025 | 0.0351 | 0.0373 | 0.0329 | 0.0351 | 0.0020 | ***

C. s 1- C. striatus 1, C. s 2- C. striatus 2, C. p 2- C. punctatus 2, C. p 1-C. Punctatus 1, C. dip 1- C. diplogramme 1, C. dip 2- C. diplogramme 2, C. m 1- C. marulius 1, C. m 2- C. marulius 2, C. b 1- C. breheri 1, C. b 2- C. breheri 2, C. a 1- C. aurentimaculata1, C. a 2- C. aurentimaculata 2, C. g 1- C. gachua 1, C. g 2- C. gachua 2.
Results

Observations indicated that transitions continue to outnumber transversions, even in the comparisons between the ingroup and the outgroups. All two sequences from each species were included in the phylogenetic analysis. The MP (Fig. 9) and NJ (Fig. 10) trees revealed identical phylogenetic relationship among the species. Two major clusters were obtained with the first cluster formed by the *C. bleheri*, *C. gachua* and *C. aurentimaculata* with *C. punctatus* as a subcluster. The second cluster was formed by *Channa diplogramme*, *C. marulius* and *C. striatus*.

In both the trees, these clusters were supported by high bootstrap values (NJ 88-100%; MP 84-99%).

Sequencing the COI gene produced an average of 560 nucleotide base pairs per taxon. No insertions, deletions or stop codons were observed in any sequence and multiple alignments resulted in a consensus length of 560 sites including base pairs and gaps. The COI alignment consisted of 560 positions, of which 223 were variable and 204 were parsimony informative. The DNA sequence information of COI gene was translated to amino

| Table 7. Nucleotide Frequencies COI |
|---|---|---|---|---|---|
|   | T  | C  | A  | G  | Total |
| *C. aurentimaculata* 1 | 29.2 | 30.3 | 25.3 | 15.2 | 561 |
| *C. aurentimaculata* 2 | 29.2 | 30.1 | 25.5 | 15.2 | 561 |
| *C. bleheri* 1 | 28.0 | 31.0 | 23.5 | 17.5 | 561 |
| *C. bleheri* 2 | 28.0 | 31.0 | 23.5 | 17.5 | 561 |
| *C. gachua* 1 | 30.5 | 28.2 | 25.8 | 15.5 | 561 |
| *C. gachua* 2 | 30.5 | 28.2 | 25.8 | 15.5 | 561 |
| *C. marulius* 1 | 28.3 | 31.0 | 23.7 | 16.9 | 561 |
| *C. marulius* 2 | 28.3 | 31.0 | 23.7 | 16.9 | 561 |
| *C. diplogramme* 1 | 25.1 | 32.8 | 24.8 | 17.3 | 561 |
| *C. diplogramme* 2 | 25.1 | 32.6 | 25.0 | 17.3 | 561 |
| *C. punctatus* 1 | 28.9 | 30.7 | 23.7 | 16.8 | 561 |
| *C. punctatus* 2 | 28.9 | 30.7 | 23.7 | 16.8 | 561 |
| *C. striatus* 1 | 30.7 | 28.5 | 24.6 | 16.2 | 561 |
| *C. striatus* 2 | 30.5 | 28.7 | 24.6 | 16.2 | 561 |
| *B. badis* | 29.2 | 30.7 | 22.8 | 17.3 | 561 |
| *P. japonicus* | 31.0 | 27.3 | 24.2 | 17.5 | 561 |
| *G. rarus* | 30.3 | 27.5 | 23.5 | 18.7 | 561 |
| Avg. | 28.9 | 30.0 | 24.3 | 16.7 | 561 |
Results

acid sequences using MEGA 3.1 (Kumar et al., 2004) and all codon positions were identified. The average nucleotide composition (Table 7) for COI was 24.3% A, 28.9% T, 30% C, and 16.7% G.

Sequence divergence (Table 8) for COI gene within Channa ranged from 14.0% (C. aurantimaculata and C. bleheri) to 25.8% (C. striata and C. bleheri). The outgroup sequence differed by 25.3%-27.3% from the Channa sequences. All species differed by more than 14% from the averages. The average pair wise ratio of transitions (Si) v. transversions (Sv) between different Channa species was calculated and the results are given in Appendix 3. Most of the variability among sequences was observed in the second and third codon position. First codon position showed little variation. In the present study, all the codon sites were used for the analysis. Nucleotide frequencies differ greatly among codon positions.

The average observed frequency (Appendix 3) of nucleotides in second position among the Channa species was observed to be; G 10.0%, A 25.8%, T 30.4%, C 33.8%; and third codon was G-22.6%, A-29.7%, T-21.5%, C-26.2%. All two COI sequences for each species were included in the phylogenetic analysis. The MP (Fig: 11) and NJ (Fig: 12) tree revealed identical phylogenetic relationship among the species. As with 16S rRNA, two major clusters were obtained where the first cluster formed with C. bleheri, C. aurantimaculata and C. gachua with C. punctatus as a subcluster. The second cluster was formed by Channa diplogramme, C. marulius and C. striatus. As with 16S rRNA, the clusters were supported by high bootstrap values (NJ 98-100 % and MP 95-99%).
Table- 8. Genetic distance values of Channa species using CO1 gene.

<table>
<thead>
<tr>
<th></th>
<th>Ca1</th>
<th>Ca2</th>
<th>Cb1</th>
<th>Cb2</th>
<th>Cg1</th>
<th>Cg2</th>
<th>Cm1</th>
<th>Cm2</th>
<th>Cdip1</th>
<th>Cdip2</th>
<th>Cp1</th>
<th>Cp2</th>
<th>Cs1</th>
<th>Cs2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ca1</td>
<td>*****</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ca2</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cb1</td>
<td>0.140</td>
<td>0.143</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cb2</td>
<td>0.140</td>
<td>0.143</td>
<td>0.000</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cg1</td>
<td>0.153</td>
<td>0.155</td>
<td>0.162</td>
<td>0.162</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cg2</td>
<td>0.153</td>
<td>0.155</td>
<td>0.162</td>
<td>0.162</td>
<td>0.000</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cm1</td>
<td>0.221</td>
<td>0.223</td>
<td>0.227</td>
<td>0.227</td>
<td>0.230</td>
<td>0.230</td>
<td>0.000</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cm2</td>
<td>0.221</td>
<td>0.223</td>
<td>0.227</td>
<td>0.227</td>
<td>0.230</td>
<td>0.230</td>
<td>0.000</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cdip1</td>
<td>0.212</td>
<td>0.214</td>
<td>0.239</td>
<td>0.239</td>
<td>0.223</td>
<td>0.223</td>
<td>0.194</td>
<td>0.194</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cdip2</td>
<td>0.218</td>
<td>0.220</td>
<td>0.245</td>
<td>0.245</td>
<td>0.229</td>
<td>0.229</td>
<td>0.200</td>
<td>0.200</td>
<td>0.005</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cp1</td>
<td>0.197</td>
<td>0.200</td>
<td>0.196</td>
<td>0.196</td>
<td>0.207</td>
<td>0.207</td>
<td>0.190</td>
<td>0.190</td>
<td>0.199</td>
<td>0.202</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cp2</td>
<td>0.193</td>
<td>0.196</td>
<td>0.196</td>
<td>0.196</td>
<td>0.207</td>
<td>0.207</td>
<td>0.190</td>
<td>0.190</td>
<td>0.199</td>
<td>0.202</td>
<td>0.004</td>
<td>*****</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cs1</td>
<td>0.243</td>
<td>0.246</td>
<td>0.255</td>
<td>0.255</td>
<td>0.233</td>
<td>0.233</td>
<td>0.194</td>
<td>0.194</td>
<td>0.224</td>
<td>0.225</td>
<td>0.192</td>
<td>0.197</td>
<td>*****</td>
<td></td>
</tr>
<tr>
<td>Cs2</td>
<td>0.246</td>
<td>0.248</td>
<td>0.258</td>
<td>0.258</td>
<td>0.236</td>
<td>0.236</td>
<td>0.192</td>
<td>0.192</td>
<td>0.222</td>
<td>0.222</td>
<td>0.195</td>
<td>0.199</td>
<td>0.002</td>
<td>***</td>
</tr>
</tbody>
</table>

C. a 1- C. aurentimaculata 1, C. a 2- C. aurentimaculata 2, C. b 1- C. bleacheri 1, C. b 2- C. bleacheri 2, C. g 1- C. gachua 1, C. g 2- C. gachua 2, C. m 1- C. marlius 1, C. m 2- C. marlius 2, C. dip 1- C. diplogramme 1, C. dip 2- C. diplogramme 2, C. p 2- C. punctatus 2, C. p 1- C. Punctatus 1, C. s 1- C. striatus 1, C. s 2- C. striatus 2.
Results

Pair wise comparisons, sequences alignment and DNA analysis of the 16S rDNA gene and Cytochrome oxidase sub unit I gene partial sequence show low values of nucleotide differences. Billington and Herbert (1991) and Sang et al., (1994) argued that selective pressures on the fishery stocks may result in genetic divergence. Hence, the low values of the nucleotide differences indicate that the pressures (e.g., fishing pressures, exploitation and natural barriers) influencing the genetic divergence of Indian Channa. The seven species studied showed large differences in COI sequences that corresponded to interspecies levels, but showed only slight differences in the 16SrDNA gene sequence. Similar findings were reported by Xia et al., 2006. The existence of a large number of species-specific nucleotide positions makes COI a more suitable marker for genetic identification in Channids.
Species identification based on the colour pattern and morphometric was not reliable for discriminating similar species of *Channa* for systematic analysis. Environmental factors can influence the morphological characters of the *Channa*, thereby affecting the reliability of the identification (Musikasinthorn, 1998; 2000). The goal of this study was to obtain a general view of the genetic divergence, taxonomic status and to develop the species specific sequence to resolve the taxonomic ambiguities between the species of the *Channids* from Indian waters, with the intent to facilitate further developments of the captive breeding and restocking programs for aquaculture and conservation. *Channids* enjoy a good market value as culture and ornamental species outside India (Gopalakrishnan and Ponniah 2000, Haniffa *et al.*, 2004; 2006) and species such as *C. bleheri* *C. barca* *C. aurantimaculata* and *C. orientalis* are exported in large numbers. Unfortunately due to overexploitation and habitat destruction species such as *C. diplogramme*, *C. gachua*, *C. punctatus* and *C. bleheri* started exhibiting a declining trend in natural bodies (IUCN 2003).
Attempts at captive breeding of these species started more than 10 years back and were only successful very recently (Haniffa et al., 2000). This success has opened an opportunity for mass production of these valued species potentially capable of serving both aquaculture for commercial production and restocking for conservation purposes. However, no attempts have been undertaken to assess the genetic status of these species, both in captivity and in the wild. The present study is unusual in that it presents one of a few instances whereby a genetic survey is undertaken prior to commencement of commercial aquaculture production and restocking.

The natural extension of the present study is to examine the finer scale genetics diversity and phylogeny of these species. This approach will help to explain the historical processes that are responsible for the distribution and population structure of these species. Furthermore, the information will also be useful to determine evolutionary significant units (ESUs) and management units (MU) for conservation purposes. Correct taxonomic definition and genetic variation of species must be established before management and conservation strategies to be adopted.

The taxonomic value of molecular genetic data is now well accepted for the identification and delineation of species both in general (Avise 1974, Thorpe 1982, Richardson et al. 1986, Baverstock and Moritz 1996) and specifically in relation to commercial fish species (Shaklee 1983, Ward and Grewe 1994). The
approach to the delineation of species using molecular genetic data is largely the same as for any other type of data.

The present study investigates phylogenetic relationships among seven species of the family *Channidae* based on partial sequencing of the mitochondrial 16S rRNA, and CO I genes. Neighbour-joining (NJ) analysis, implemented in MEGA2.1 (Kumar *et al.*, 2001), was employed to both examine relationships among taxa in the profiles and for the subsequent classification of 'test' taxa because of its strong track record in the analysis of large species assemblages. This approach has the additional advantage of generating results much more quickly than alternatives. COI gene (mitochondrial Cytochrome c oxidase subunit I) was used because of its faster evolving nature it was applied to evaluate not only the genetic differentiations obtained but also the status of species that shared identical haplotypes in 16S. Phylogenetic relationships based on morphological characters and molecules are mostly concordant (Ward *et al.*, 2005). The seven species of *Channids* from Indian were genetically distinct from each other without any haplotypes sharing or overlapping, based on the partial sequence information of both 16S rRNA and COI genes. The sequences of both regions of mtDNA demonstrated simplicity and unambiguity.

Mitochondrial DNA (mt DNA) has been extensively used to understand, in a large variety of organisms, the molecular relationships among individuals, populations and species (Manchado *et al.*, 2004). The gene content and organization of vertebrate complete mitochondrial genomes (mitogenomes) is
Discussion

quite conserved, although gene rearrangements have been described in some species (Manchando et al., 2004). In fact, mitogenomes have demonstrated adequacy in resolving persistent controversies over high level relationships in teleosts.

Within the mitochondrial genome, the 16S ribosomal DNA (16S rDNA) gene is widely studied (Olsen et al., 1991; Lindeque et al., 1999). Owing to its slowest mutation and lower substitution rates compared with other mt DNA genes, it has been reported to be useful when analyzing families, species and population (Meyer, 1994; Garland and Zimmer, 2002). COI has proved to be a robust evolutionary marker for determining intraspecific and interspecific relationships (Clary and Wolstenholme, 1985; Baldwin et al., 1996; Hoeh et al., 1996; Boudry et al., 1998).

Numerous studies have used 16S DNA sequences to resolve evolutionary relationships at different taxonomic levels (Palumbi and Benzie, 1991; Levinton et al., 1996; Sturmbauer et al., 1996). Thus the COI and 16S rRNA was chosen for this study because of their potential to be phylogenetically informative at lower taxonomic levels. The wide range in utility of these mitochondrial genes at various taxonomic levels suggests that the different rates of molecular evolution within the genes, due to varying functional constraints, greatly affect its phylogenetic utility.

Sequence analysis revealed that the average nucleotide composition for 16S was 30.3% A, 21.8% T, 25.3% C, and 22.7% G and the average nucleotide
Discussion

composition for COI was 24.3% A, 28.9% T, 30% C, and 16.7% G. The genetic divergent values obtained in the present study are comparable with the same obtained in other fish species using the 16S rRNA and COI. However, the individuals within the species showed very low levels in intra-specific variation. This was reflected in the lowest number of haplotypes (one each for all the species) and may be due to the high proportion of the haplotypes detected in the present study or the limited number of individuals (2 each) collected for the present study. The presence of fixed haplotype differences between the species showed high genetic divergence values. Similar results are reported in a number of teleost species (Nguyen et al., 2005; Rossi et al., 2004; Murphy et al., 1997; Meyer et al., 1993). The genetic divergent values obtained in the present study are comparable with the same obtained in other fish species using 16S rRNA sequence data (Chakraborty et al., 2006; Garcia et al., 2000).

The natural extension of the present study is to examine the finer scale of genetics divergence of these seven species with more intensive sampling using mitochondrial 16S rRNA and COI genes. This approach will help to explain the historical processes that are responsible for distribution and genetic divergence of these species. In addition, the information will also be useful to determine evolutionary significant units (ESUs) and management units (MU) for conservation purposes. The results of 16S rRNA sequences are entirely consistent with the results from analyses of Xia Li et al., 2006, Ambok et al., 2007 in Channids species. The fact that 16S rRNA and COI display very similar patterns of heterogeneity and species structuring pattern provides the evidence that these
species show high level of genetical divergence. Different classes of molecular markers have been employed simultaneously to resolve taxonomic ambiguity and to estimate intra-specific variations in teleost species. Similar concordant results are reported in flying fish using mt DNA (Gomes et al., 1998); in needle fishes using sequence information of nuclear (creatine kinase B) gene and mt DNA genes (Banford et al., 2004) and in bony-tongue fishes using sequence analysis of five mitochondrial genes (Lavoue and Sullivan, 2004).

In the present study there was no nucleotide compositional bias in 16S rRNA gene. There was significant bias, however, in the second and third codon position of the COI gene. Compositional bias could have two origins, selection or mutation pressure (Sueoka, 1988). A consequence independent of the origin of the bias is that the transition/ transversion ratio will vary at different positions due to differences in the selection pressure (Irwin et al., 1991). It is generally believed that the ratio of transitional mutations to transversions in the mtDNA molecule decreases toward higher taxonomic levels because undetectable multiple transitional events at more variable sites accumulate as a function of time since divergence (Brown et al., 1985).

The generally low ratio of transitions/ tranversions found for COI (especially at the third codon position) might suggest a relative saturation of the COI sites, as was also shown in the scatter plot of transitions and pairwise genetic distance. This situation could affect the accuracy of the analysis used. In order to examine this possibility the additional phylogenetic analysis based on the COI
Discussion

segment by excluding those sites was done but the results obtained were similar to those described below.

The levels of divergence estimated among the five species using the combined data sets ranged from 3.09% to 13.66% in 16S and 14.0% to 25.8% in COI. These values are in general agreement with those reported by Billington and Hebert (1991), as well as with those proposed by Powers (1993) for marine species. Moreover, the level of nucleotide divergence observed among the three *Liza* species (4.9%) is in congruence with that proposed by Avise *et al.*, (1987) and Moritz *et al.* (1987) among congeneric species.

The highest degree of genetic divergence has been estimated among *Channa striatus* and *Channa bleheri* this could be the result of the substitution rate observed in this species, and it could be explained as a combined effect of nucleotide bias (especially on the third codon position of COI) and saturation of signal (Martin, 1995). This observation is in agreement with previous studies by Papasotiropoulos *et al.*, (2001, 2002) using allozyme and PCR-RFLP analysis, as well as with the studies performed by Murgia *et al.*, (2002); Rossi *et al.*, (2004); and Turan *et al.*, (2005).

The observed transition vs transversion ratios in Channids are also comparable to the same in many teleosts (Ward *et al.*, 2005, Chakraborty *et al.*, 2006a). Transitions outnumbered transversions in the present study in accordance with the previous reports on mtDNA in fish (Vinson *et al.*, 2004). Within the species, levels of intra-specific variation were low and this was reflected in the
lowest number of haplotypes (one to two for these species). This may be due to
the high proportion of the identified haplotype in the samples or the limited
number of individuals (2, each) collected for the present study. The observed
genetic divergence values and transition vs. transversions ratios are comparable to
the mean 2.3% in sharks (Martin, 1995) and in many perciformes (Garcia et al.,
2000) and in ribbon fishes (Chakraborty et al., 2006). Transitions outnumbered
transversions in the present study in accordance with the previous studies of
mitochondrial DNA in fish (Garcia et al., 2000).

Generally for mtDNA, a much larger excess of transitions related to
transversions is typically observed (Page and Holmes, 1998); hence further
analysis and tree construction, “Kimura-two-Parameter mutational model” is
generally adopted (Hall, 2004). Kimura-two-Parameter method accommodates
higher rate of transitions i.e. this method allows differential weighing of transition
and transversion probabilities (Hall, 2004). In the present study also, for
estimating genetic divergence Kimura-two-Parameter method was adopted.

Phylogenesis from molecular data are often computed by pair-wise genetic
distance based (numerical) methods like Neighbor Joining (NJ) and Maximum
Parsimony (MP) tree, with branch lengths that are proportional to the amount of
divergence, making it possible to estimate the relative times of separation. NJ tree
making method is a widely use distance-clustering algorithm that allows unequal
rates of divergence among lineages. Phylogenetic trees are also made based on
“discrete methods” that operate directly on sequences like the Maximum
Parsimony (MP) tree. MP chooses the tree (or trees) that require the fewest evolutionary changes \((i.e. \) it makes trees from sequences exhibiting smallest evolutionary changes). In variant characters (bases) those that have the same state in all taxon are obviously useless (phylogenetically uninformative) and are ignored by the MP method. Both the numerical (NJ) and discrete (MP) tree making methods are used in the analysis in most of the species (Hall, 2004) as in the present study.

Many members of the Family Channidae have been difficult to analyse with a rigorous phylogenetic content due to overlapping of useful morphological characters and size of the species. Generally speaking, independent analyses of molecular data unequivocally support the genetic divergence of \textit{Channa} species. The relatively large mt DNA distances between most of the Channids begs the question of whether the species are relatively old or characterized by an accelerate rate of molecular evolution. This is similar to the observed divergence values of \textit{Scomberomorous} \textit{sp.} (Banford et al., 1993), belonid species (Banford et al., 2004) and Channidae (Xia Li et al., 2006, Ambok et al., 2007).

Sequence analysis of mt DNA genes alone will prove if molecular evolution is accelerated in \textit{Channa} and the rate of increase affect in mitochondrial genes. Our phylogenetic results complement the tree presented by Xi Li et al., 2006. Our results confirm the monophyly of Asian \textit{Channa} species as found out in the work of Xi Li et al., 2006. The molecular phylogeny presented here is largely congruent with the subgeneric apportionment of Xia Li et al., (1960),
supporting monophyly of the subgenera *Channa*. The Neighbor Joining and Maximum Parsimony analysis also produced similar tree topology. The Maximum parsimony (MP) and the Neighbour joining (NJ) analyses using Kimura 2 parameter yielded trees with identical topology with high bootstrap support values. Seven well-supported highly divergent lineages were identified corresponding to *Channa striatus*, *C. punctatus*, *C. marulius*, *C. gachua*, *C. diplogramme*, *C. bleheri* and *C. aurantimaculata*.

The phylogenetic relationship among the major Indian channid species revealed that the monophyly of *channids* relative to representative species of some supposedly related families seems reasonable because all species of the Channidae share several distinct morphological features. The monophyly of Indian channids is also in agreement with the present classification which is widely accepted (Teugels and Daget, 1984; Bonou and Teugels, 1985; Nelson, 1994; Musikasinthorn, 1998, 2000; Zhang *et al.*, 2002). The genetic data presented in this study strongly suggests that two species of *Channa* was confused under the name *Channa micropletes*, were as *C. micropletes* originally described from specimens originating from Java and *C. diplogramme* described from Kerala. This study demonstrates that species specific sequence data from the mitochondrial DNA has considerable potential for addressing systematic questions and taxonomic problems in channid fishes. The 16S and COI region of the channid mitochondrial genome contained sufficient genetic variation and phylogenetic signal to be useful in species identification.