*Mitochondrial DNA*, June 2012; 23(3): 239–249

informa
healthcare

**FULL LENGTH RESEARCH PAPER**

# Morphological taxonomy and molecular divergence of four balitorid species (subfamily: *Nemachelinae*) from Central Himalaya, India

AJAY PANDEY[1], WAZIR S. LAKRA[2], RAJENDRA P. THAPLIYAL[1], MUKUNDA GOSWAMI[3], MAHENDER SINGH[3], & ABHISHEK K. MALAKAR[3]

[1]*Department of Zoology and Biotechnology, H.N.B. Garhwal Central University, Srinagar, Garhwal 246174, Uttarakhand, India,* [2]*Central Institute of Fisheries Education (Deemed University), Versova, Andheri (W), Mumbai 400061, Maharastra, India, and* [3]*National Bureau of Fish Genetic Resources, Canal Ring Road, P.O. Dilkusha, Lucknow 226002, Uttar Pradesh, India*

*(Received 2 December 2011; revised 6 March 2012; accepted 6 March 2012)*



**Abstract**

The fishes in this study belong to the difficult group. It is difficult to identify about more than 50 nemacheilid loach species in India, mainly due to the poor quality of the original descriptions, the lack of good reviews, and the similarity of a number of the banded species. In view of this, a morphometric and genetic study was conducted on four species that have been documented in Central Himalayan region of India. Using comparative methods, a contribution to the knowledge about the variability of four Himalayan hillstream loaches, a molted loach *Acanthocobitis botia* (Hamilton, 1822), *Nemacheilus corica* (Hamilton, 1822), a creek loach *Schistura beavani* (Gunther, 1868), and *Schistura montana* (McClelland, 1838), was accomplished. Morphometric data were analyzed by univariate (CV and analysis of variance) and multivariate analyses (discriminant function analysis and cluster analysis). Based on the results of these statistical analyses, generally all four species grouped into their own species cluster with minimal overlap between two species of genus *Schistura*. Mitochondrial DNA cytochrome b (Cyt b) gene analysis revealed 7.7–17.6% genetic distances among the four species. Molecular phylogenetic relationship among these species and other Balitoridae species was investigated using published mitochondrial Cyt b sequences. Dendrograms obtained by the maximum likelihood and maximum parsimony method analyses exhibited the same topology. According to this topology, all the four species represented distinct species group and *A. botia* form the most distinct species while *S. beavani* and *S. montana* are clustered together with *Nemacheilus* species to form a single group. This work may build the base for the revision of taxonomic identity of fishes of the family Balitoridae. The results may further help to enhance the knowledge of the ichthyologists in understanding the fish fauna of India and assist them in planning conservation and management strategies for the propagation of these less studied small indigenous species along their natural range of distribution.

**Keywords:** *Loaches, principal component analysis, discriminant function analysis, cytochrome b, cluster analysis*

## Introduction

The fish species belonging to the family Balitoridae, namely hillstream loaches, likely originated in the early Miocene probably in the SouthWest Chinese region and dispersed Westwards to Central Asiatic Highlands and subsequently to India along an Eastern Tibetan route (Hora 1937; Banarescu and Mirza 1972; Menon 1987). These species are quite varied and diverse, and form several recognizable groups or subgenera in the habitats ranging from warm tropical waters to cold water or mountainous streams and rivers (Kottelat et al. 1988).

Among balitorids, subfamily *Nemachelinae* is unique for the relatively high species diversity, including more than 200 fish species distributed in India, South China, Southeast Asia, Baluchistan, Western Iran, South Anatolia, and Laka Tsana in North East Africa

Correspondence: A. Pandey, Ornamental Fish Division, The Marine Products Export Development Authority, Ministry of Commerce and Industry, Government of India, Panampilly Nagar, P.O. Ernakulam 682036, Kerala, India. Tel: + 91-9418079824. Fax: + 91484 2313361. E-mail: ajayp_2@rediffmail.com

ISSN 1940-1736 print/ISSN 1940-1744 online © 2012 Informa UK, Ltd.
DOI: 10.3109/19401736.2012.674117

and about 80 species alone are reported in Indian region. These loaches are preferred locally for food besides their role in lucrative aquarium trade (Menon 1987). Apart from this, the caves dwelling species are of scientific value from an evolutionary viewpoint (Romero and Green 2005; Kottelat et al. 2007; Juan et al. 2010). Furthermore, from the biodiversity perspective too, this group has immense potential as most of the ichthyofaunal discoveries during the recent times have been from this subfamily (Vishwanath and Nebeshwar 2004, 2005; Sreekantha Gururaja et al. 2006; Vishwanath and Shanta 2006; Kottelat et al. 2007) and many more should be expected in India. Although of much significance no study so far focused on the molecular characteristics, systematics, and phylogenetics of these small indigenous fishes in India (except the partial taxonomic revision of the few genera from NorthEast Himalayan regions; Hora 1932). Consequently, the taxonomy of these fishes has been in great confusion (Menon 1987). These fishes also show high levels of morphological and ecological conservatism (Kottelat 1988). In the past, this conservatism allowed several authors to ascribe most loaches to the genus *Nemacheilus* based on body coloration (i.e. number and type of bands), number of branched rays in dorsal fin, and the nature of caudal fin (Menon 1987; Kottelat 1988; Talwar and Jhingran 1991). However, recent cladistic analyses recognized the generic status to the numerous subgenera within the subfamily (Jayaram 2010; Eschmeyer 2011; Froese and Pauly 2011).

The Central Himalayan region (latitudes 29°05′–31°25′N and longitudes 77°45′–81°E; area 51,124 $km^2$) of India is endowed with vast freshwater resources (including upper River Ganga basin), and thus known to harbor a rich diversity of freshwater fishes (Hora and Mukerji 1936; Hora 1937; Menon 1949, 1974; Pant 1970; Singh 1990; Khan 2000). About seven *Nemachelinae* species reported so far in this region of India. These species formally assigned to the genera *Acanthocobitis*, *Nemacheilus*, and *Schistura* and recently *Paraschistura* based on the external anatomical characters (Prokofiev 2009; Jayaram 2010; Eschmeyer 2011; Froese and Pauly 2011).

Besides taxonomic and systematic questions, high priority is also given to address conservation programs of some threatened loach species (CAMP 1998), since several anthropogenic factors such as sewage and industrial pollution, extensive habitat degradation and alteration, and the recent proliferation of exotic fish species lead to significant population decline (IUCN 2011). Any further impact on the habitat quality would immediately move them to Critically Endangered (Sharma 2004; Atkore et al. 2011; Sarkar et al. 2011). Therefore, for monitoring species-specific conservation program, correct identification of the unit to be conserved is vital. Incorrect taxonomy leads to the lack of identification of threats to the survival of important components of biodiversity. Conversely, conservation resources can be a waste by focusing on supposedly rare species that are merely nongenetic phenotypic variants of a common species (Avise 1994). Nevertheless, the similarity of morphological and ecological features of these species may be an obstacle to positive species identification and reliable population estimates.

Genetic markers are powerful tools for solving questions concerning ecology, evolution, and conservation of several marine and freshwater taxa (Agnese et al. 1997; Yoshizaki et al. 1997; Chapman et al. 1999; Wilson et al. 2000; Thangaraj and Lipton 2010). Presently, among the DNA markers used for species discrimination, the mitochondrial markers are more popular (Teletchea 2009). These markers are already proven as important tool to address the molecular systematics status of the various loach species in several other countries (Borowsky and Mertz 2001; Janko et al. 2005; Mihara et al. 2005; Shedko et al. 2008) and other commercially important species in India (Jayasankar et al. 2004; Mohindra et al. 2005; Lakra et al. 2007, 2010).

In this study, taxonomical characterization and genetic divergence were investigated among the four balitorid species from the Central Himalayan region of India based on the results of morphology and mitochondrial gene cytochrome b (Cyt b) sequence analysis. The study also attempted to clarify the systematic status of these species using the corresponding Cyt b sequences of other balitorids retrieved from GenBank. The results of this study could form taxonomic base for the revision of identity of these species in India subcontinent.

## Methods

### *Study area and sample collection*

In total, 115 specimens of four species of *Nemachelinae* subfamily, namely *Acanthocobitis botia*, *Nemacheilus corica*, *Schistura beavani*, and *Schistura montana* were collected from streams and rivulets belonging to upper Ganga river basin during the year 2007–2009 (Figure 1) using cast nets (mesh size: 9′, 1″, 9′, 1/2″) and other indigenous methods. In order to increase the sample size of each species, the experimental fishing was performed multiple times during the study period at the same geographical locality (Table I). All the four species were finally recognized as per the available taxonomic keys (Menon 1987; Jayaram 2010).

### *Morphometric and meristic marker analyses*

The following morphometric and meristic characters were analyzed for all fish species: standard length (SL), head length (HL), body depth (BD), snout

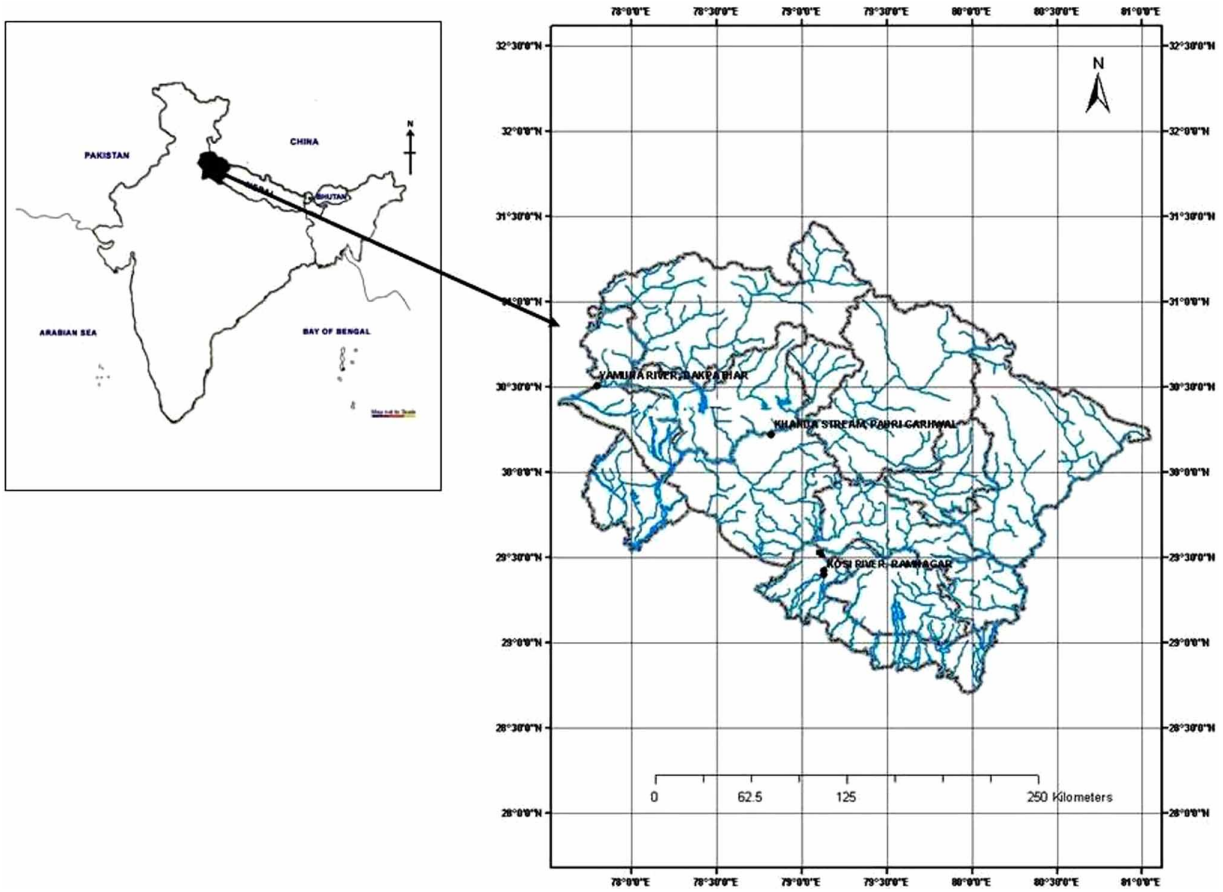


Figure 1. Drainage map of upper Ganges basin showing the sampling sites.

length (SntL), post-orbital distance of head, eye diameter, interorbital width, dorsal fin length, dorsal fin base length, pectoral fin length, pelvic fin length, anal fin length (AFL), caudal fin length, pre-dorsal length, post-dorsal length, pre-pelvic length (PrPlvL), pre-anal length (PrAL), distance of pectoral fin to pelvic fin origin, distance of pelvic fin to anal fin origin, least height of caudal peduncle (LCPD), caudal peduncle length (CPL), and number of fin rays in dorsal, pectoral, pelvic, anal, and caudal fins. All the measurements were taken using electronic digital calipers (Mitiyo make) to the nearest 0.01 mm.

Morphomteric data were allometric transformed prior to further analysis to stabilize variances in specimens of different sizes (Elliott et al. 1995):

$$\text{Madj} = M\left(\frac{\text{LS}}{\text{LO}}\right)^{b},$$

where $M$ is the original measurement, Madj is the size-adjusted measurement, LO is the SL of the fish, and

Table I. Summary of sampling sites with total number of samples.

| Species | River/ stream | Sampling months | | | GPS coordinates | Sample sizes | | GenBank Accession Nos |
|---|---|---|---|---|---|---|---|---|
| | | February | June | October | | Morphometric study | Molecular analysis | |
| *A. botia* | Yamuna | 12 | 06 | 10 | 30°30′13.02 N 77°47′50.91 E | 28 | 28 | FJ711332–FJ711352 and GQ478433–GQ478439 |
| *N. corica* | Kosi | 10 | 08 | 12 | 29°23′47.60 N 79°07′54.26 E | 30 | 26 | FJ711353–FJ711375 and GQ478440–GQ478442 |
| *S. beavani* | Kosi | 09 | 17 | 06 | 29°25′09.60 N 79°07′59.54 E | 32 | 32 | FJ711376–FJ711401 and GQ478443–GQ478448 |
| *S. montana* | Khanda | 18 | 03 | 07 | 30°12′04.89 N 78°46′08.34 E | 28 | 17 | FJ711402–FJ711418 |

LS is the overall mean of the SL for all fishes from all samples in each analysis. Parameter *b* was estimated for each character from the observed data as the slope of the regression of log *M* on log LO, using all fishes combined for both sites.

The coefficient of variation (CV) was computed for each character according to

$$CV = \frac{(100 \times SD)}{X_m},$$

where SD is the standard deviation and $X_m$ is the mean of the transformed measurements of characters in each species. In each species' sample group, morphological variability was estimated by the multivariate generalization of the coefficient of variation (CV*p*) according to

$$CV_p = 100 \times \sqrt{\frac{\sum SD_x}{\sum M_x}},$$

where $SD_x$ is the variance of each morphometric variable and $M_x$ is the mean square (Van Valen 1978).

### *mtDNA marker analysis*

The tissues were collected by removing either of the paired fins using sterile razors and forceps and preserved in 95% *v*/*v* ethanol at 4°C. Total genomic DNA was isolated by using a standard phenol–chloroform procedure (Sambrook et al. 1989) followed by digestion with proteinase-K at 37°C for overnight, alcohol precipitation, and wash with 70% ethanol. The final pellet of DNA suspended with 50 µl of Tris-EDTA (TE) buffer.

A 1127 bp region of mitochondrial Cyt *b* region was amplified (Irwin et al. 1991) by polymerase chain reaction (PCR) using an upstream primer L14724 (5′-TGACTTGAARAACCAYCGTTG-3′) and a downstream primer H15915 (5′-CTCCGATCTCCGGATTACAAGAC-3′). Amplification was carried out in a 50 µl volume, each containing 5 µl 10 × PCR buffer, 2 µl (50 mM) $MgCl_2$, 0.25 µl (0.05 mM) of each deoxynucleotide triphosphate (dNTP), 0.4 µl 0.6 U of Taq DNA polymerase, 0.5 µl (0.01 mM) each of forward and reverse primers, and 2 µl genomic DNA. Amplification was carried out by a thermocycler (MJ PTC-200, Watertown, MA, USA) at 95°C for 4 min for initial strand separation, followed by 40 cycles of 1 min at 94°C, 1 min at 37°C, 2 min at 72°C, and a final 10 min extension at 72°C. An automated DNA sequencer (ABI 3730 DNA sequencer, Genewiz, NJ, USA) was used to generate the sequences of the amplified PCR products.

### *Statistical analysis*

The descriptive statistics for each morphometric and meristic characters were obtained. A one-way analysis of variance (ANOVA) was carried out on each transformed morphometric variable. Significant variables were retained and then subjected to stepwise discriminant analysis (DA) and cluster analysis. The meristic characters among species were compared using nonparametric Kruskal–Wallis test (*H*-test). All statistical analyses were done using the statistical analysis system (StatSoft 1995).

Light and heavy strands of mitochondrial Cyt b gene were sequenced for all the species. Sequences were inverted (reversed and complimented) to match the heavy strand sequence using the EditSeq and MegAlign subprograms of DNASTAR software package (Burland 2000). Ambiguities were referenced against the sequencing electropherograms and corrected as required. Overlapping sequence fragments from each individual were completely assembled to create consensus mitochondrial sequences. The sequences were then aligned using ClustalW (Thompson et al. 1997) for comparison.

The extent of sequence difference between species was calculated by averaging pairwise comparisons of sequence difference across all individuals. Pairwise evolutionary distance among haplotypes was determined by *P*-distances. To conduct the phylogenetic analysis, we downloaded Cyt b sequences of 14 species (genera: *Schistura*, *Acanthocobitis*, *Nemacheilus*, *Barbatula*, and *Triplophysa*) from GenBank database. As outgroups, from NCBI database, we also downloaded sequences of *Balitora elongata* and *Crossostoma lacustre* for subfamily-level comparative study and *Botia banarescui* for family-level comparative study. The details of the sequences studied are given in Tables I and II. The maximum likelihood (ML) and maximum parsimony (MP) trees were constructed to know the phylogenetic relationship among the species. Prior to ML analysis, the Modeltest v3.6 program

Table II. GenBank Accession Numbers for the mitochondrial Cyt b gene of the specimens investigated here.

| Family/subfamily | Species | NCBI, GenBank Accession Nos |
|---|---|---|
| Balitoridae/ Nemachelinae | *A. botia* | HQ184046 |
| | *A. zonalternans* | GQ174374 |
| | *Nemacheilus putaoensis* | DQ105226 |
| | *N. subfuscus* | DQ105224 |
| | *N. polytaenia* | DQ105227 |
| | *S. spilota* | EF508596 |
| | *S. thai* | DQ105202 |
| | *S. fasciolata* | DQ105201 |
| | *S. longa* | AY625698 |
| | *Triplophysa orientalis* | DQ105251 |
| | *Barbatula barbatula* | DQ105255 |
| | *Barbatula nuda* | DQ105253 |
| Balitoridae/ Balitorinae | *C. lacustre* | AY392467 |
| | *B. elongata* | DQ105217 |
| Cobitidae/Botiinae | *Bo. Banarescui* | AY281262 |

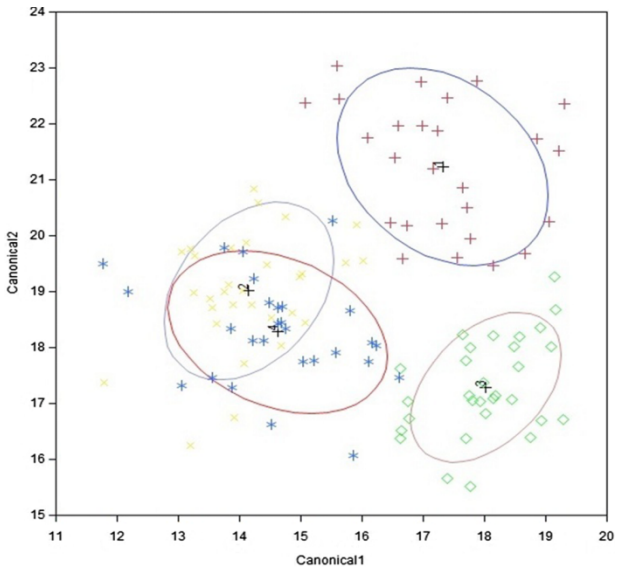


Figure 2. Discriminant analysis plot where the eight morphometric variables were used. 1, *A. botia*; 2, *S. beavani*; 3, *N. corica*; 4, *S. montana*. Ellipses include 95 % of the specimens.

(Posada and Crandall 1998) was run in order to calculate the best substitution model for analysis. The model with the highest Akaike information criterion value, i.e. Jukes–Cantor model, was selected for further analysis using the nearest neighbor interchange algorithm and a bootstrap (BP) of 1000 replicates. The estimation of Cyt b gene diversity and phylogenetic reconstruction was carried out using MEGA v5 software (Tamura et al. 2011). Number of transition and transversion was plotted separately versus the nucleotide distance using Data Analysis and Molecular Biology Evolution (DAMBE 5.12) software package (Xia and Xie 2001).

## Results

### *Morphometric analysis*

ANOVA ($P < 0.05$) showed significant differences in the mean values of eight morphometric characters including HL, BD, SntL, AFL, PrPlvL, PrAL, LCPD, and CPL. Whereas in the Kruskal–Wallis test (*H*-test), apart from the number of anal fin rays, the mean values of the remaining four meristic characters differed significantly (at 5% level) among the four species.

Generally, low to high range of coefficient of variations (CVs) for the morphometric characters (0.7–29.24%) were obtained. The multivariate generalization of the coefficient of variation (CVp) for the 21 transformed morphometric variables was also relatively low. Although *N. corica* shows the highest CVp (6.01%), followed by *S. beavani* (5.11%), *S. montana* (4.77%), and *A. botia* (4.66%), these values are relatively low.

Forward stepwise DA of the eight significant variables produced five discriminating variables. The first and second canonical variables contributed 54.6% and 41.5%, respectively, while the third canonical variable contributed 3.9% to the total variance. The characters of primary importance in distinguishing groups were the BD and HL for the CaV-1, the LCPD and SntL for the CaV-2, and AFL for the CaV-3. These discriminant functions identified the membership (classification) of individual fish in the data with one of the four species with a success rate of 83.5%. The graphical presentation of the first and second canonical variables is shown in Figure 2.

The Unweighted Pair Group Method with Arithmetic Mean (UPGMA) cluster analysis based on the Mahalanobis distance between group centroids showed that the four species were clustering in two clusters. *A. botia* and *N. corica* belong to the first cluster (cluster I), while *S. beavani* and *S. montana* belong to the second one (cluster II; Figure 3). This clustering suggests closer morphological similarity between fish species *A. botia* and *N. corica*, and between species *S. beavani* and *S. montana*.

### *mtDNA marker analysis*

The complete sequence of 1127 bp of Cyt b was determined for 103 individuals of all the species. A total of 840 sites (75% of total sites) were constant and 287 sites were variable (25% of the total sites). Among the informative sites, 82.99% came from the third codon position. Third codon position bases were more variable (137 constant, 239 informative) than first codon position bases (330 constant, 45 informative) and second codon position bases (372 constant, 4 informative). Transitions (30%) were more common than transversions (12%). Informative sites were primarily transitions in the third codon position (Ts = 79.76%, Tv = 94.11%) and first codon position (Ts = 19.04%, Tv = 2.94%). Few of the informative sites were transitions in the second codon position (Ts = 1.19 2% of all substitution). The mean nucleotide diversity ($P_i$) was found to be 0.10.

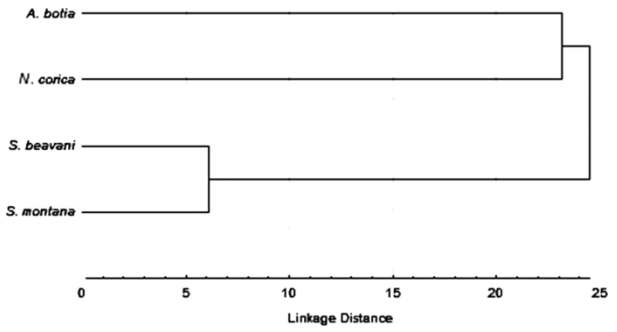


Figure 3. UPGMA cluster analysis based on the Mahalanobis distance between the species Centroids.

Table III. Pairwise values of mean % sequence divergence derived from uncorrected "*P*" genetic distance within and between four balitorid species using Cyt b sequence.

| | *A. botia* | *N. corica* | *S. beavani* | *S. montana* |
|---|---|---|---|---|
| *A. botia* | | | | |
| Mean % divergence | 0.1 | | | |
| *N. corica* | | | | |
| Mean % divergence | 17.6 | 0.7 | | |
| *S. beavani* | | | | |
| Mean % divergence | 13.9 | 14.2 | 0.0 | |
| *S. montana* | | | | |
| Mean % divergence | 15.9 | 15.5 | 7.7 | 0.1 |
| Groups/species | Uncorrected distance | % Divergence | Number of comparisons | Standard deviation |
| All species | 0.10 | 10% | 103 | 0.006 |

Altogether 18 unique haplotypes (H1–H18) were identified in 1127 bp sequences of 103 samples. It includes four of *A. botia*, five of *N. corica*, five of *S. beavani*, and four of *S. montana*. Haplotype H2 present in *A. botia* was found to be the most prevalent (13.59%) followed by H11 in *S. beavani* (12.62%) while H17 in *S. montana* was least prevalent (1.94%). No haplotypes were shared between the species and the haplotype diversity ($H_d$) was found to be 0.93. Mean pairwise uncorrected sequence divergence (*P*-distance) calculated among and between the four balitorid species is given in Table III. The mean base composition showed low G content (14.9%) and almost equal A, T, and C contents (28.9%, 29.3%, and 26.9%). The mean percentage of A + T for the first codon and second codon positions was 49.1% and 62.3%, which was just a bit higher in the third codon (63.4%) position. The mean transition/transversion ratio was 5:21. The overall positional bias was 0.82 (82%). The highest bias was found in the third codon position that had a bias of 1.95%, and there was a strong anti-guanine bias for the third codon (G = 5.4%) position. The transitions and transversions of the Cyt b sequences for all the codon position were plotted versus nucleotide divergence (*P*-distance), a relatively linear relationship with effect to first and second codon position results with transitions exceeding transversions upto 0.12 (Figure 4). The number of transitions due to the third codon position increases linearly with genetic distance, but beyond 0.12, it started decreasing, thereby indicating the saturation level. Nucleotide base pairs were translated into amino acid residues. Of 375 total residues, 357 (95.2%) were constant across all four balitorid species and only 18 (4.8%) were parsimony informative.

Phylogenetic analysis using two reconstruction methods (ML and MP) gives the same general topology. Analyses using ML (Figure 5) method revealed that subfamily *Nemachelinae* forming monophyletic group was highly supported statistically (BP, 98%). The phylogram can be divided in two groups (named A; BP: 93%) which contains the four balitorids (*S. montana*, *S. beavani*, *A. botia*, and *N. corica*) plus the other species of genera *Schistura*, *Nemacheilus*, and *Acanthocobitis*, while the second group (B) contains the genera *Triplophysa* and *Barbatula*. Within the group A, four monophyletic subgroups can be shown which are well statistically supported. The *Schistura* species (*Schistura spilota*, *Schistura thai*, *Schistura fasciolata*, and *Schistura longa*) constitute one monophyletic subgroup C (BP: 100%). Similarly, the *A. botia* haplotypes (including Ab082 sequence) form subgroup (D) with *Acanthocobitis zonalternans* (BP: 100%), and the *N. corica* haplotypes form a monophyletic group E (BP: 100%) not shared with others. The *S. beavani* haplotypes and *S. montana* haplotypes constitute a monophyletic group (F) in which the *Nemacheilus putaonesis* and *Nemacheilus subfuscus* form sister subgroups with *S. beavani* (BP, 100%) and *Nemacheilus polytaenia*, a separate clade (BP: 100%). The MP analyses (Figure 6) also resulted in the similar tree topology with good BP values for the inner nodes; however, the monophyly of the subfamily *Nemachelinae* is not supported here (BP: <50%). Although the subgroups D (BP: 75%), E (BP: 79%) and F (BP: 98%) of group A are well statistically

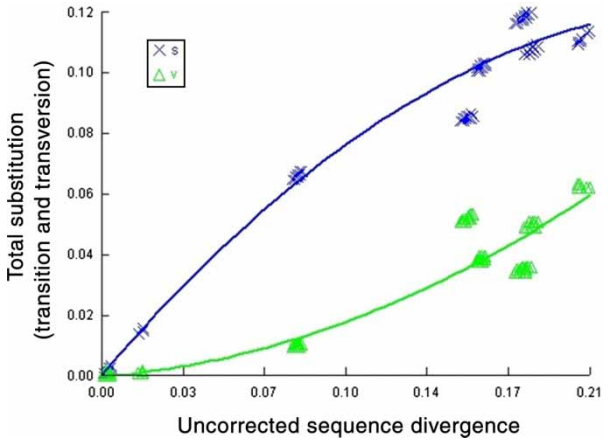


Figure 4. Total substitutions of Cyt b plotted as a function of pairwise sequence divergence for four balitorid species.

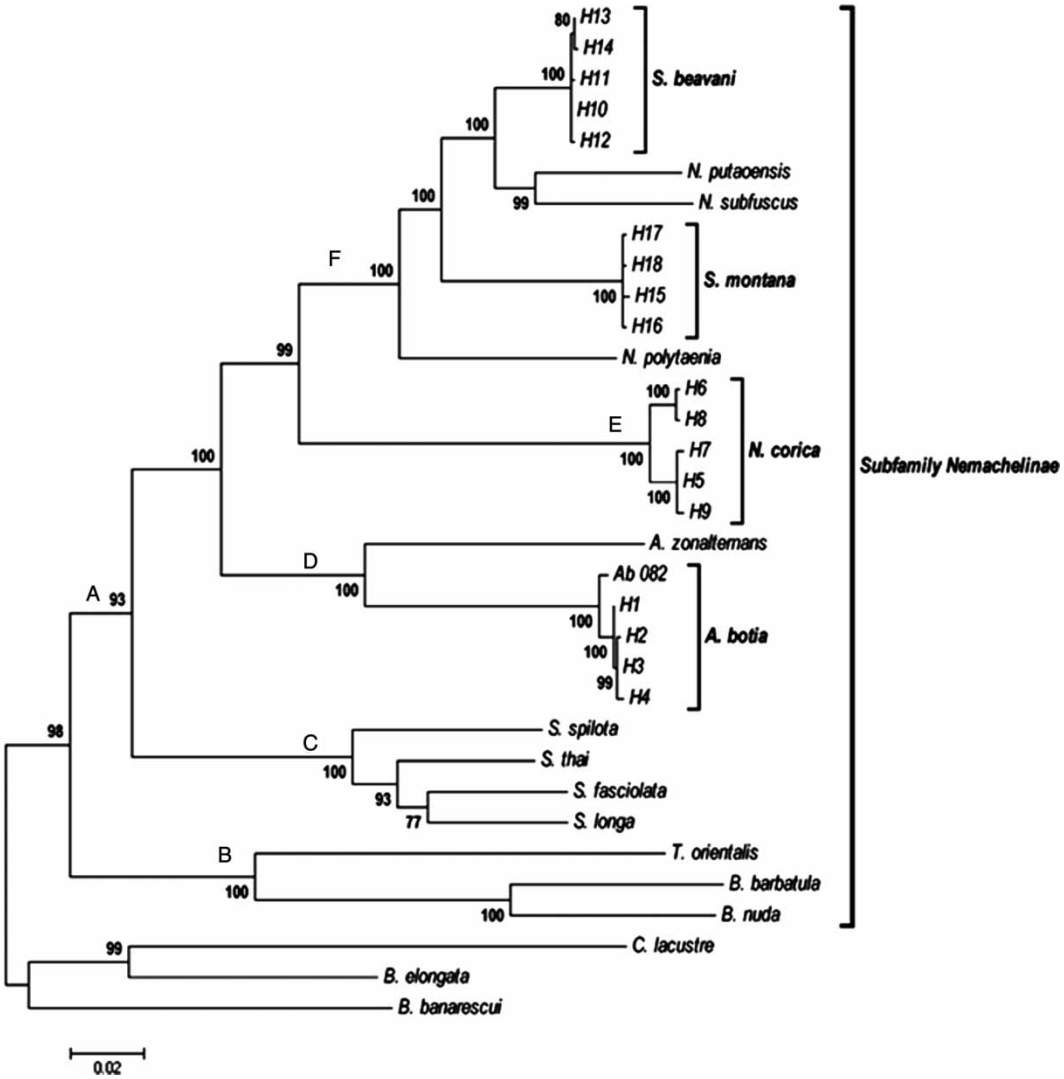


Figure 5. Phylogenetic tree using ML method showing the relationship among Cyt b haplotypes of the four balitorids and other related species. Only representative haplotype sequences are taken for the analysis.

supported, the subgroup C has shown low BP values thus not supported (BP: $<50\%$) with this method.

## Discussion

The present survey represents the first attempt to integrate morphological and molecular marker analyses together in order to assess the species identification and evolutionary relationships among these balitorid species. Although only a limited number of Himalayan *balitorid* species of subfamily *Nemachelinae* (Menon 1987) were analysed, several issues on biology and ecology outline the usefulness of such multi-approach analysis to address most of the questions concerning the biological conservation of several threatened loach species throughout their natural range of distribution along Himalayas. This type of experimental analysis has successfully applied earlier on numerous marine groups at multiple taxonomic levels, focusing on the assessment of intraspecific and interspecific diversity, ecological interactions, and conservation (Kocher and Stepien 1997; Feral 2002).

The number of measurements and counts used in the present study was considered to be of adequate number for this type of fishes compared with earlier reports (Klingenberg 1996; Elvira and Almodovar 2000; Iguchi et al. 2003; Poulet et al. 2004; Chen et al. 2005; Champasri et al. 2007). The within-species variation in most cases was less evident as indicated by the relative low CV values ($CV < 20\%$) for most of the morphometric characters along with low intraspecific variation and it suggests that each species consists of a phenotypically and genetically homogeneous group.

The various multivariate techniques are being increasingly employed in fish systematic using the morphometric and meristic characters, the commonly utilized procedures are discriminant function analysis (DFA) or DA and cluster analysis. The DFA maximizes differences between groups (species or

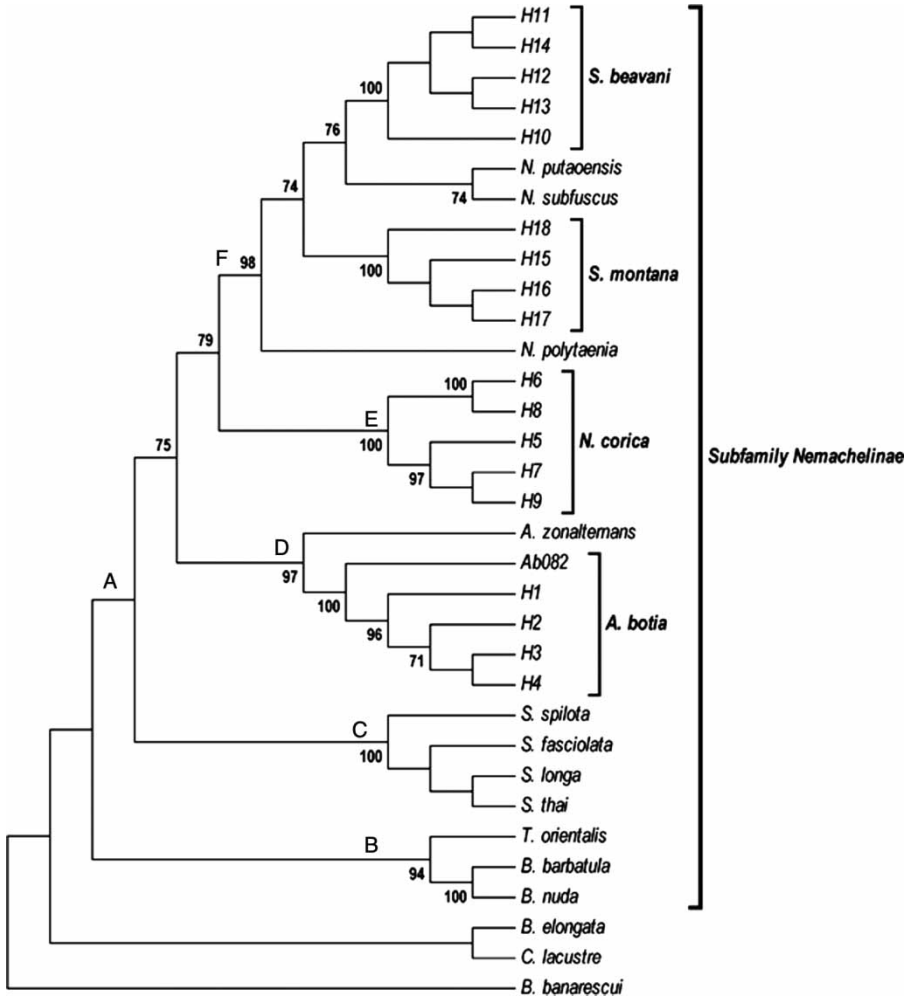


Figure 6. Phylogenetic tree using MP method showing the relationship among Cyt b haplotypes of the four balitorids and other related species. Bootstrap values (<50%) for the analysis are shown above the nodes. Only representative haplotype sequences are taken for the analysis.

population) and then classifies unknowns to these groupings (Beuttell and Losos 1999), while the cluster analysis proceeds from the same similarity (or dissimilarity) matrix and attempts to depict relationships between the groups by means of dendrogram (Cormack 1971). According to DA classification, 83.5% of the specimens examined in this study can be correctly classified into four species. The discriminatory function along CaV-1 was the BD and HL, which was comparatively deeper and larger in *A. botia* than in the other balitorid species. CaV-2 accounts for 41.5% of the variation and separates *A. botia* from the other three species based on the characteristics of SntL and LCPD. CaV-3 accounts for the remaining 3.9% of the variation and provides weak discrimination based on AFL, which was comparatively smaller in *S. beavani* than the other three species. In the cluster analysis, a dendrogram of phenetic similarity of morphometric characters revealed two major groups of species, which are morphologically similar. The first group included *S. montana* and *S. beavani* and the second included *A. botia* and *N. corica*. The lower morphologic divergence between *S. montana* and *S. beavani* indicates their closeness.

The genetic analysis supports the characterization of four *balitorid* species, as does multivariate analysis, which resulted in statistically significant differences between the four species. In Cyt b sequence analysis, several features characterized nucleotide substitution. Transitional changes occurred more frequently than transversional changes as is typical of animal mitochondrial genomes (Briolay et al. 1998). Our results are in agreement with the results recorded previously for fish Cyt b genes (Bartlett and Davidson 1991; Carr and Marshall 1991; McVeigh and Davidson 1991; Magoulas et al. 1995; Park et al. 2000; Tang et al. 2006). The interspecific divergence in the present study was greater than intraspecific variation. The uncorrected Cyt b divergence, which averaged 10%, suggests species-level differences. The ranges of sequence divergence observed between the species of different genera of Cobitidae and related loach families depicted >7.79% in Cyt b sequences

(Tang et al. 2006), while in case of Atheriniformes it is >17% among different genera of the family Melanotaeniidae (Zhu et al. 1994). In the present study, the average Cyt b divergence ranged from 7.7% to 17.0%, and these values are similar to those in other interspecific studies (Beckenbach 1991; McVeigh and Davidson 1991; Park et al. 2000) but differ from the range 0.2–9.7%, observed between species of genus *Schizothorax* (He and Chen 2006). The mean base compositions of Cyt b sequences have a low G content and almost equal A, T, and C contents. Similar nucleotide composition has been reported in several cyprinids (Briolay et al. 1998; Durand et al. 2002; Tsigenopoulos et al. 2002; Tang et al. 2006). Strong bias in the base composition is a typical feature of Cyt b and other mitochondrial genes (Irwin et al. 1991). Cyt b sequences were analyzed using two common methods of phylogenetic inference (ML and MP) and recovered congruent topologies. Tang et al. (2006) based on mitochondrial DNA Cyt *b* and control region sequence data revealed that the Balitoridae is a monophyletic group, likely derived from the ancestors of the Nemacheilidae and Cobitidae while *Schistura*, which includes several rather distinct groups of fishes, is polyphyletic in origin. Our phylogenetic inference using ML method revealed subfamily *Nemachelinae* as monophyletic group; however, the same is not conclusive with MP analysis. Neverthless, both methods of phylogenetic reconstruction revealed *A. botia*, *N. corica*, *S. beavani*, and *S. montana* as a distinct monophyletic groups. Moreover, the sequences of other *Schistura* species formed a distinct group not shared by any species. The latter result brings into question the generic status of *S. beavani* and *S. montana*. However, the existence of such disorders in the phylogenetic studies of other fish groups is common (Cataudella et al. 1974; Menezes et al. 1992; Caldara et al. 1996; Papasotiropoulos et al. 2002).

## Conclusion

Our results identified both morphological and mitochondrial DNA markers useful to discriminate the four important balitorid species. From the present morphological-based multivariate study and the mtDNA Cyt b-based sequences analysis and from the previous taxonomic comparisons (Menon 1987; Kottelat 1990; Talwar and Jhingran 1991; Jayaram 2010), it appears that four taxonomic balitorid species should be considered as valid. Based on our study, the inferred phylogenetic closeness indicates that the assignment of the species among the genera, i.e. *Nemacheilus* and *Schistura*, is improper. However, a more extensive morphological and genetic survey including other species especially from genera *Nemacheilus*, *Schistura*, and *Acanthocobitis* is needed to contribute further to this systematic debate. In addition, considering large diversity in the family Balitoridae, the possibility of obtaining an absolutely reliable phylogenetic topology from any single molecular marker seems less possible.

## Acknowledgments

The authors are thankful to Prof. J.P Bhatt, H.N.B. Garhwal University, for his guidance during this research work and the anonymous reviewers for providing constructive suggestion in improving the manuscript.

***Declarations of interest***: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

## References

Agnese JF, Teugels GG, Galbusera P, Guyomard R, Valckaerts F. 1997. Morphometric and genetic characterization of sympatric populations of *Clarias gariepinus* and *C. anguillaris* from Senegal. J Fish Biol 50:1143–1157.

Atkore VM, Sivakumar K, Johnsingh AJT. 2011. Patterns of diversity and conservation status of freshwater fishes in the tributaries of River Ramganga in the Shiwaliks of the Western Himalaya. Curr Sci 100(5):731–735.

Avise JC. 1994. Molecular markers, natural history and evolution. New York: Chapman and Hall, Inc.

Banarescu P, Mirza MR. 1972. *Noemacheilus alepidotus* nalbanti nova subsp., a new loach from Rawlakot, Azad Kashmir. Biologia 18(2):121–123.

Bartlett SE, Davidson WS. 1991. Identification of Thunnus tuna species by the polymerase chain reaction and direct sequence analysis of mitochondrial Cytochrome b genes. Can J Fish Aqua Sci 48:309–317.

Beckenbach AT. 1991. Rapid mtDNA sequence analysis of fish populations using polymerase chain reaction (PCR). Can J Fish Aqua Sci 48(1):95–98.

Beuttell K, Losos JB. 1999. Ecological morphology of Caribbean anoles. Herpet Mono 13:1–28.

Borowsky R, Mertz L. 2001. Genetic differentiation among the cave fish, *Schistura Oedipus* (Cypriniformes: Balitoridae). Environ Bio Fish 62:225–231.

Briolay J, Galtier N, Brito RM, Bouvet Y. 1998. Molecular phylogeny of Cyprinidae inferred from Cytochrome b DNA sequences. Mol Phylogenet Evol 9:100–108.

Burland TG. 2000. DNASTAR's lasergene sequence analysis software. Methods Mol Biol 132:71–91.

Caldara F, Bargelloni L, Ostellari L, Penzo E, Colombo L, Patarnello T. 1996. Molecular phylogeny of grey mullets based on mitochondrial DNA sequence analysis: Evidence of differential rate of evolution at the intra family level. Mol Phylogenet Evol 6:416–424.

CAMP. 1998. Report of the workshop on Conservation Assessment and Management Plan (CAMP) for freshwater fishes of India, Zoo Outreach Org and NBFGR, Lucknow, India, 22–26 Sept 1997 p 1–156.

Carr SM, Marshall HD. 1991. Detection of intraspecific DNA sequence variation in the mitochondrial Cytochrome b gene of Atlantic cod (*Gadus morhua*) by the polymerase chain reaction. Can J Fish Aqua Sci 48:48–52.

Cataudella SM, Civitelli V, Capanna E. 1974. Chromosome complement of the Mediterranean mullets (Pisces, Perciformes). Caryologia 27:93–105.

Champasri TR, Rapley M, Duangjinda Suksri A. 2007. A morphological identification in fish of the genus Puntius Hamilton 1822 (Cypriniformes: Cyprinidae) of some wetlands in Northeast Thailand. Pak J Biol Sci 10:4383–4390.

Chapman LJ, Chapman CA, Brazeau DA, McLaughlin B, Jordan M. 1999. Papyrus swamps, hypoxia, and faunal diversification: Variation among populations of *Barbus neumayeri*. J Fish Biol 54: 310–327.

Chen QQ, Lu DR, Ma L. 2005. Morphological differentiation between close populations discernible by multivariate analysis: A case study of genus Coilia (Teleostei: Clupeiforms). Aqua Living Resour 18:187–192.

Cormack RM. 1971. A review of classification. J Roy Stat Soc Ser A 134:321–367.

Durand JD, Tsigenopoulos CS, Unlu E, Berrebi P. 2002. Phylogeny and Biogeography of the family Cyprinidae in the Middle East inferred from Cytochrome b DNA-evolutionary significance of this region. Mol Phylogenet Evol 22:91–100.

Elliott NG, Haskard K, Koslow JA. 1995. Morphometric analysis of orange roughy (*Hoplostethus atlanticus*) off the continental slope of southern. Aus J Fish Biol 46:202–220.

Elvira B, Almodovar A. 2000. Further observations on the morphological characters of *Acipenser sturio* L., 1758 from the Iberian Peninsula: A comparison with North and Adriatic Sea populations. Bol Inst Esp Oceanogr 16:89–97.

Eschmeyer WN. 2011. Catalog of fishes electronic version. Available at: http://research.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (accessed 5 July 2011).

Feral JP. 2002. How useful are the genetic markers in attempts to understand and manage marine biodiversity. Exp Mar Bio Ecol 268:121–145.

Froese R, Pauly D. 2011. Fishbase. Available at: http://www.fishbase.org (accessed 1 July 2011).

Gunther A. 1868. Catalogue of the Physostomi, containing the families Heteropygii, Cyprinidae, Gonorhynchidae, Hyodontidae, Osteoglossidae, Clupeidae, … [thru] … Halosauridae, in the collection of the British Museum. Catalogue Fishes Br Mus 7:1–512.

Hamilton F. 1822. An account of the fishes found in the river Ganges and its branches, Edinburgh: Archibald, Constable & Co.; London: Hurst, Robinson & Co., vol. I–VII, pp 1–405, pls. 1–39.

He D, Chen Y. 2006. Biogeography and molecular phylogeny of genus *Schizothorax* (Teleostei: Cyprinidae) in China inferred from cytochrome b sequences. J Biogeol 33:1448–1460.

Hora SL. 1932. Classification, bionomics and evolution of Homalopterid fishes. Mem Indian Mus 12:263–330.

Hora SL. 1937. Distribution of Himalayan fishes and its bearing on certain palaeogeographical problems. Rec Ind Mus 39(3): 25–259.

Hora SL, Mukerji DD. 1936. Notes on the fishes in the Indian Museum. XXVII – on two collections of fish from Maungmagan, Tavoy District, Lower Burma. Rec Indian Mus 38:15–39.

Iguchi K, Yamamoto G, Matsubara N, Nishida M. 2003. Morphological and genetic analysis of fish of a *Carassius* complex (Cyprinidae) in Lake Kasumigaura with reference to the taxonomid status of two all-female triploid morphs. Biol J Linn Soc 79:351–357.

Irwin DM, Kocher TD, Wilson AC. 1991. Evolution of cytochrome b gene of mammals. J Mol Biol Evol 32:128–144.

IUCN. 2011. IUCN red list of threatened species, Version 2011.1. Available at http://www.iucnredlist.org (accessed 3 July 2011).

Janko K, Vasi'ev VP, Ráb P, Rábová M, Šlechtová V, Vasil'eva ED. 2005. Genetic and morphological analyses of 50-chromosome spined loaches (Cobitis, Cobitidae, Pisces) from the Black Sea basin that are morphologically similar to *C. taenia*, with the description of a new species. Folia Zoolog 54:405–420.

Jayaram KC. 2010. The freshwater fishes of the Indian region. Delhi: Narendra Publishing House. p 551.

Jayasankar P, Thomas PC, Paulton MP, Mathew J. 2004. Morphometric and Genetic analyzes of Indian Mackerel (*Rastrelliger kanagurta*) from Peninsular India. Asian Fish Sci 17:201–215.

Juan C, Guzik MT, Jaume D, Cooper SJB. 2010. Evolution in caves: Darwin's "wrecks of ancient life" in the molecular era. Mol Ecol 19:3865–3880.

Khan MA. 2000. An account of fishes of Uttar Pradesh plains. Rec Zool Surv India 98:101–121.

Klingenberg CP. 1996. A combined morphometric and phylogenetic analysis of an ecomorphological trend: Pelagization in Antarctic fishes (Perciformes: Nototheniidae). Biol J Linn Soc 59:143–177.

Kocher TD, Stepien CA. 1997. Molecular systematics of fishes. San Diego, CA: Academic Press.

Kottelat M. 1988. Two species of cavefishes from northern Thailand in the genera Nemacheilus and Homaloptera (Osteichthyes: Homalopteridae). Rec Aus Mus 40(4):225–231.

Kottelat M. 1988. Indian and Indochinese species of *Balitora* (Osteichthyes: Cypriniformes) with descriptions of two new species and comments on the family-group names Balitoridae and Homalopteridae. Rev Suisse Zool 95:487–504.

Kottelat M. 1990. Indochinese Nemacheilines. A revision of Nemacheiline loaches (Pisces: Cypriniformes) of Thailand, Burma, Laos, Cambodia and southern Viet Nam. Verlag Dr Friedrich Pfeil München p 262.

Kottelat M, Daniel HR, Graham SP. 2007. *Schistura papulifera*, a new species of cave loach from Meghalaya, India (Teleostei: Balitoridae). Zootaxa 1393:35–44.

Lakra WS, Goswami M, Mohindra V, Lal KK, Punia P. 2007. Molecular identification of five Indian sciaenids (Pisces: Perciformes, sciaenidae) using RAPD markers. Hydrobiologia 583:359–363.

Lakra WS, Verma MS, Goswami M, Lal KK, Mohindra V, Punia P, Gopalakrishnan A, Singh KV, Ward RD, Hebert P. 2010. DNA barcoding Indian marine fishes. Mol Biol Rep 11(1):60–71.

Magoulas A, Sophronides K, Patarnello T, Hatzilaris E, Zouros E. 1995. Mitochondrial DNA variation in an experimental stock of gilthead sea bream (Sparus aurata). Mol Mar Biol Biotech 4:110–116.

McClelland J. 1838. Observations on six new species of Cyprinidae, with an outline of a new classification of the family. J Asiat Soc Bengal 7:941–948.

McVeigh HP, Davidson WS. 1991. A salmonid phylogeny inferred from mitochondrial Cytochrome b gene sequences. J Fish Biol 39:277–282.

Menezes MR, Martins M, Naik S. 1992. Interspecific genetic divergence in grey mullets from the Goa region. Aquaculture 105:117–129.

Menon AGK. 1949. Notes on Fishes; XLIV – fish from the Koshi Himalayas. Nepal Rec Ind Mus 47:231–237.

Menon AGK. 1974. A check-list of Fishes of the Himalayan and Indo-Gangetic plains. Special Publication, Inland Fish. Soc. India, Barrackpore 1:1–136.

Menon AGK. 1987. The fauna of India and the adjacent countries. Pisces. IV Teleostei – Cobitoidea. Part 1. Homalopteridae. Zoolog Surv India p. 259.

Mihara M, Sakai T, Nakao K, Martins L, Hosoya K, Miyazaki J. 2005. Phylogeography of loaches of the genus *Lefua* (Balitoridae, Cypriniformes) inferred from mitochondrial DNA sequences. Zoo Sci 22:157–168.

Mohindra V, Anshumala Punia P, Narain L, Kapoor D, Lal KK. 2005. Microsatellite loci to determine population structure of Labeo dero (Hamilton Buchanan). Aqua Liv Resour 18:83–85.

Pant MC. 1970. Fish fauna of the Kumaon Hills. Rec. Zool Surv India 64(1–4):85–96.

Papasotiropoulos V, Klossa-Kilia E, Kilias G, Alahiotis S. 2002. Genetic divergence and phylogenetic relationships in grey mullets (Teleostei: Mugilidae) based on PCR-RFLP analysis of mtDNA segments. Biolo Genet 40:71–86.

Park JY, Lee HJ, Kim WJ, Lee JH, Min KS. 2000. Mitochondrial Cytochrome b sequence variation in Korean salmonids. J Fish Biol 56:1145–1154.

Posada D, Crandall KA. 1998. MODELTEST: Testing the model of DNA substitution. Bioinformatics 14:817–818.

Poulet N, Berrebi P, Crivelli AJ, Lek S, Argillier C. 2004. Genetic and morphometric variations in the pikeperch (*Sander lucioperca* L.) of the fragmented delta. Arch Hydrobiol 159:531–554.

Prokofiev AM. 2009. Problems of the classification and phylogeny of Nemacheiline loaches of the group lacking the preethmoid I (Cypriniformes: Balitoridae: Nemacheilinae). J Ichthyol 49(10): 874–898.

Romero A, Green SM. 2005. The end of regressive evolution: Examining and interpreting the evidence from cave fishes. J Fish Biol 67:3–32.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning a laboratory manual. 2nd ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sarkar UK, Pathak AK, Sinha RK, Sivakumar K, Pandian AK, Pandey A, Dubey VK, Lakra WS. 2011. Freshwater fish biodiversity of River Ganga (India): Pattern, threats and conservation. Rev Fish Biol Fisheries 22(1):251–272.

Sharma RC. 2004. Protection of an endangered fish *Tor tor* and *Tor putitora* population impacted by transportation network in the area of Tehri Dam Project, Garhwal Himalaya, India. In: Irwin CL, Garrett P, McDermott KP, editors. Proceedings of the 2003 International Conference on Ecology and Transportation. Raleigh: Center for Transportation and the Environment, North Carolina State University. p 83–90.

Singh S. 1990. The lotic water fisheries of Kumaun Himalaya Himalaya environment resources and development. Almora: Shree Almora Book Depot. p 276–284.

Shedko SV, Miroshnichenko IL, Nemkova GA. 2008. On the systematics and phylogeography of eight – Barbel loaches of the genus Lefua (Cobitoidea: Nemacheilidae): mtDNA typing of *L. pleskei*. Russian J Genet 44(7):817–825.

Sreekantha Gururaja KV, Remadevi K, Indra TJ, Ramchandra TV. 2006. Two new fish species of genus *Schistura* McClelland (Cypriniformes: Balitoridae) from Western ghats, India. Zoos' Print J 21(4):2211–2216.

StatSoft. 1995. Statistica for windows (computer program, a manual). Tulsa, OK: StatSoft., Inc.

Talwar PK, Jhingran AG. 1991. Inland fishes of India and adjacent countries. New Delhi, Bombay, Calcutta: Oxford & IBH Publishing Co.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739.

Tang Q, Liu H, Mayden R, Xiong B. 2006. Comparision of evolutionary rates in the mitochondrial DNA cytochrome b gene and control region and their implications for phylogeny of the cobitidae (Teleostei: Cypriniformes). Mol Phylogenet Evol 39: 347–357.

Teletchea F. 2009. Molecular identification methods of fish species: Reassessment and possible applications. Rev Fish Biol Fish 19(3):55–85.

Thangaraj M, Lipton AP. 2010. Genetic identity of three Indian populations of three spotted seahorse, *Hippocampus trimaculatus*. Adv Biolo Res 4(1):37–41.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882.

Tsigenopoulos CS, Kotlik P, Berrebi P. 2002. Biogeography and pattern of gene flow among Barbus species (Teleostei: Cyprinidae) inhabitating the Italian Peninsula and neighbouring Adriatic drainages as revelead by allozyme and mitochondrial sequence data. Biol J Linn Soc 75:83–99.

Van Valen L. 1978. The statistics of variation. Evol Theory 4:35–43.

Vishwanath W, Nebeshwar K. 2004. Schistura reticulata, a new species of balitorid loach from Manipur, India with redescription of S.chindwinica. Ichthyol Explore Freshwaters 15(4):323–330.

Vishwanath W, Nebeshwar K. 2005. A new nemacheiline fish of the genus Schistura McClelland (Cypriniformes: Balitoridae) from Manipur, India. J Bombay Nat Hiso Soc 102(1):79–82.

Vishwanath W, Shanta KM. 2006. A new nemacheiline fish of the genus *Schistura* Mcclelland (Cypriniformes: Balitoridae) from Manipur, India. J Bombay Nat Hist Soc 102(2):210–213.

Wilson PJ, Woods CM, Maina JN, White BN. 2000. Genetic structure of Lake Magadi tilapia populations. J Fish Biol 56: 590–603.

Xia X, Xie Z. 2001. DAMBE: Data analysis in molecular biology and evolution. J Heredity 92:371–373.

Yoshizaki G, Yamaguchi K, Oota T, Strüssmann CA, Takashima F. 1997. Cloning and characterization of pejerrey mitochondrial DNA and its application for RFLP analysis. J Fish Biol 51: 193–203.

Zhu D, Jamieson BGM, Hugall A, Moritz C. 1994. Sequence evolution and phylogenetic signal in control-region and Cyt b sequences of rainbow fishes (Melanotaeniidae). Mol Biol Evol 11:672–683.