Gene 527 (2013) 248–258

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Molecular characterization of major and minor rDNA repeats and genetic variability assessment in different species of mahseer found in North India

Ravindra Kumar *, Mamta Singh [1], Basdeo Kushwaha [1], N.S. Nagpure [1], Indra Mani [1], W.S. Lakra [1]

National Bureau of Fish Genetic Resources, Canal Ring Road, P.O. Dilkusha, Lucknow 226 002, U.P., India

ARTICLE INFO

Article history:
Accepted 4 June 2013
Available online 22 June 2013

Keywords:
Mahseers
Neolissochilus
Phylogeny
rDNA
Tor

ABSTRACT

Relationship among the mahseer species (Family: Cyprinidae) has long been debated in fish systematics. Present study concentrates on the nature of the phylogenetic relationship among the five mahseer species using the sequence of major ribosomal DNA (45S rDNA). We have covered rDNA sequence of approximately 5.2 kb per individual, 26.0 kb per species and 130.0 kb as a whole. We also characterized the 45S and 5S rDNA regions with respect to their nucleotide composition. For phylogenetic analyses, nucleotide sequences were divided into four datasets. First and second datasets contained 18S rDNA and ITS1 sequence, whereas third and fourth datasets consisted of ITS2 and complete 18S-ITS1-5.8S-ITS2-28S, respectively. The NJ tree was constructed for all the datasets. The mahseer species under study formed a monophyletic group well separated from the outgroup species. Similarly, the individuals of *Neolissochilus hexagonolepis* form monophyletic group with *Tor* species, indicating *Neolissochilus* as a sister genus of *Tor*. The findings from the present study provide greater insights into taxonomic status of mahseer, and set the stage for future investigations dealing with phylo-geography, taxonomy, conservation and co-evolution within this interesting and important group of fish.

1. Introduction

Family Cyprinidae is the largest, abundant and globally widespread family of freshwater fishes, which comprises 220 genera and about 2420 species (Nelson, 2006). The cyprinid fishes of the three genera *viz*., *Tor*, *Neolissochilus* and *Naziritor*, constitute an important group referred to as 'mahseer'. They are endemic to Asia with natural distribution encompassing the trans-Himalayan region in the North-West to Sumatra and Borneo islands in the South-East, across a number of Asian countries. Mahseer species are large-scaled barbells (Subfamily: Barbinae), live in upstream, clear, running waters and have high demand as food and attraction for anglers as a sport fish (Ng, 2004; Shrestha, 1990). Some of the species are of great economic value and conservation concern (Nguyen et al., 2006; Sarkar and Srivastava, 2000) as well as of aquaculture potential (Haque et al., 1995; Ingram et al., 2005). At present 46 mahseer species have been recognized; of which 23 species belong to genus *Tor* Gray, 22 to genus *Neolissochilus* Rainboth and one species to genus *Naziritor* Mirza (Eschmeyer et al., 2004). There still exists confusion with regard to taxonomy and systematics, and uniformity in identification of this group. Fishes of the genus *Tor* with the presence of the median lobe are considered to be 'true mahseer', as opposed to *Neolissochilus* and *Naziritor*, where the median lobe is absent. The features such as shape, size and length of the median lobe, often used to distinguish species of *Tor* (Zhou and Cui, 1996), are highly variable (Roberts, 1999) and are also influenced by environmental factors, leading to confusion and as such its reliability as an indicator of species is questionable (Ng, 2004).

DNA sequence data play an indispensable role in reconstruction of evolutionary relationships among the organisms, resulting in insights in genetic affinities that may confirm or conflict with traditional taxonomy. Because of the attractive properties, ribosomal DNA (rDNA) is popular source for examining phylogenetic relationships and for studying genetic variability and divergence within and between species. These properties are secondary structure features, differential

Abbreviations: $AgNO_3$, silver nitrate; AIC, Akaike Information Criterion; BCIP, 5-bromo-4-chloro-3′-indolyphosphate p-toluidine salt; CMA_3, Chromomycin A_3; DNA, deoxyribonucleic acid; dNTPs, deoxynucleoside triphosphates; EMBL, European Molecular Biology Laboratory; FISH, fluorescence *in situ* hybridization; IE, intermediate element; ITS1, internal transcribed spacer 1; ITS2, internal transcribed spacer 2; KCl, Potassium chloride; $MgCl_2$, Magnesium chloride; NBT, nitro-blue tetrazolium chloride; NCBI, National Center for Biotechnology Information; NJ, neighbor-joining; NH, *Neolissochilus hexagonolepis*; NTS, Non-transcribed spacer; PCR, polymerase chain reaction; rDNA, ribosomal deoxyribonucleic acid; SM, supplementary materials; TC, *Tor chelynoides*; TP, *Tor putitora*; TPr, *Tor progeneius*; TT, *Tor tor*; Tris–HCl, Tris hydrochloride.

* Corresponding author at: Division of Molecular Biology and Biotechnology, National Bureau of Fish Genetic Resources, Canal Ring Road, P.O. Dilkusha, Lucknow 226 002, U.P., India. Tel.: +91 522 2442440, +91 522 2442441, +91 522 24421735; fax: +91 522 2442403.

E-mail addresses: Ravindra.Scientist@gmail.com, rkumar@nbfgr.res.in (R. Kumar).

[1] Tel.: +91 522 2442440, +91 522 2442441, +91 522 24421735; fax: +91 522 2442403.

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.06.032

rates of evolution among divergent regions and tandemly repeated sequences (Arnheim, 1983; Gebri, 1985). As in other eukaryotes, the 45S rDNA of mahseer also contains tandemly repeated transcriptional units, which are separated by intergenic spacers. Within each transcriptional unit, the internal transcribed spacer 1 (ITS1) separates the 18S small subunit (18S rDNA) from the 5.8S large subunit (5.8S rDNA); while the internal transcribed spacer 2 (ITS2) separates the 5.8S from the 28S large subunit (28S rDNA). Each rRNA gene is organized into several divergent domains, also called expansion segments, interspersed between slow evolving, highly conserved cores (Hassouna et al., 1984). Divergent domains are commonly subjected to insertion and deletion events, coupled with substitution rates, that are at least two orders of magnitude higher than that for cores (Olsen and Woese, 1993), and account for the large overall size variation of rRNA genes among eukaryotes. A clear bias toward nucleotide substitutions (especially transitions) rather than insertions or deletions is found in the highly conserved core sequences. These dual modes of evolution in the expansion domains and the cores of nuclear rRNA molecules (particularly 28S) make these genes useful for phylogenetic analyses. The coding regions show little sequence divergence among closely related species, whereas the spacer regions may exhibit higher variability as mutations occur at a relatively rapid rate in internal transcribed spacers. Therefore, these regions may resolve the relationships between closely related species that otherwise show little genetic divergence (Fritz et al., 1994; Porter and Collins, 1991; Tang et al., 1996). Due to above mentioned reasons, several authors used these regions for phylogenetic analyses, *i.e.* in African anguilliform catfishes (Siluriformes: Clariidae), rDNA and spacer regions were used to established the phylogenetic relationship and divergent time estimation (Jansen et al., 2006). Verma and Serajuddin (2012) also used divergent domain 8 of 28S rDNA for phylogenetic analysis of freshwater catfish *Ompok pabda*, *Ompok pabo*, *Ompoc bimaculatus* and *Wallago attu*. There are several reports on phylogenetic studies based on 5S rDNA in fishes belonging to family Anostomidae, Cyprinidae, Merlucciidae, Mugilidae (Campo et al., 2009; Ferreira et al., 2007; Fujiwara et al., 2009; Imsiridou et al., 2007). Till date, no molecular genetic studies based on nuclear DNA or 45S and 5S rDNA sequences, which address questions relating to genetic relationship among mahseer species, were performed. But there are some reports on phylogeny of mahseer using mitochondrial

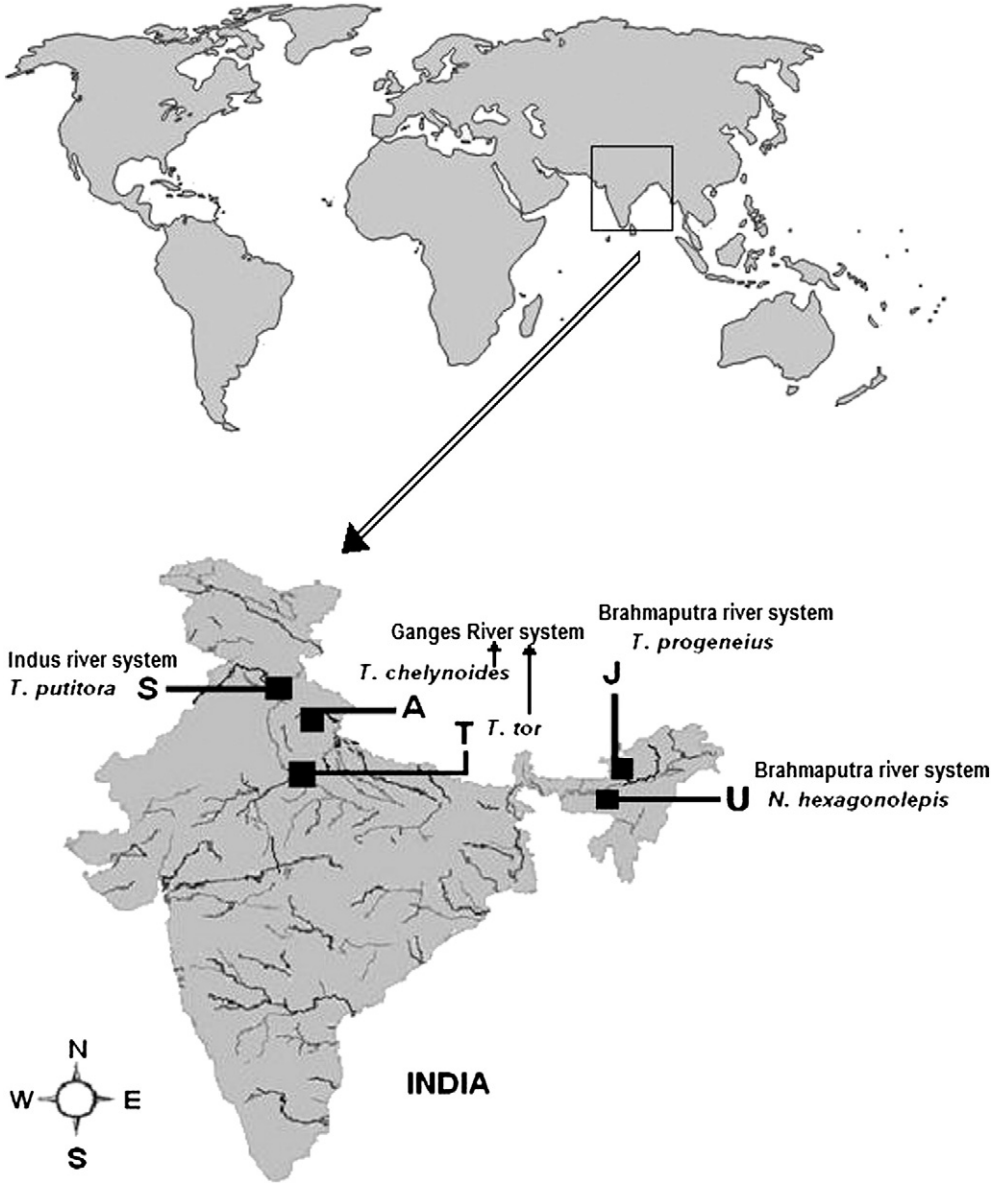

Fig. 1. Collection sites of fish specimens from India.

DNA sequences which includes 11 species (Nguyen et al., 2006, 2008). Phylogenetic hypotheses of the evolutionary relationships among members of the same genus provide frameworks for comparative research on mechanisms of diversification and speciation. The results from the present study will provide valuable resources for people concerned with conservation, where it provides a relatively objective means of quantifying evolutionary distinctiveness and resolving taxonomic ambiguities involving rare taxa.

2. Materials and methods

2.1. Sample collection and DNA isolation

A total of 25 live specimens of four species of *Tor* and one species of *Neolissochilus*, namely *T. chelynoides* (5 specimens), *T. putitora* (6 specimens), *T. progeneius* (4 specimens), *T. tor* (5 specimens) and *N. hexagonolepis* (5 specimens) were collected from Alakhnanda River (30°13′31.96″N and 78°46′48.22″E), near Garhwal, Uttarakhand; Satluj River (31°22′31.61″N and 76°48′30.74″E), near Bilaspur, Himachal Pradesh; Jia-Bhoreli River (26°59′43.79″N and 92°41′19.21″E), near Bhalukpong, Arunachal Pradesh border; Tons River (25°01′57.51″N and 81°44′07.50″E), near Rewa, Madhya Pradesh and tributaries of Umiam Reservoir (25°40′00.26″N and 91°53′07.27″E), near Shillong, Meghalaya, India, respectively (Fig. 1), with the help of local fishermen. The genomic DNA was extracted from the whole blood using the standard phenol-chloroform-isoamylalcohol method described by Sambrook and Russell (2001).

2.2. PCR amplification, cloning and sequencing

The details of primers *viz*., their sequences, region they amplified and the approximate amplicon size, used in this study are given in Table 1. Complete 18S, ITS1, 5.8S, ITS2, partial 28S and 5S rDNAs in 25 individuals (representative individuals of each species) were amplified using the primers listed in Table 1. Primers of 18S, ITS1, 5.8S and ITS2 were designed from the sequence of *Cyprinus carpio* (Accession Number AF133089) available in the NCBI database with the help of software Oligo 4.0 while primers of 28S regions were taken from Zardoya and Meyer (1996) and 5S from Moran et al. (1996). A standard PCR reaction was performed in 25 samples with the primers using 10 mM Tris–HCl (*p*H 8.3), 50 mM KCl, 2 mM $MgCl_2$, 200 μM dNTPs mix, 10 pmol of each primer (forward and reverse), 2 U *Taq* DNA polymerase and 50 ng of genomic DNA in a final reaction volume of 50 μl. PCR cycling conditions were: initial denaturation at 94 °C for 5 min; followed by 35 cycles of denaturation at 94 °C for 1 min, primer annealing at 53–58 °C (depending on the primers used) for 1 min, primer extension at 72 °C for 2 min; with post-cycling extension at 72 °C for 10 min. Amplified products were run on 1.5% agarose gel stained with ethidium bromide.

Amplified products with sizes more than 700 bp (except 5S rDNA) were purified by Qiagen PCR purification kit, according to the manufacturer's instructions, ligated with pDrive T/A cloning vector (Qiagen) and incubated overnight at 16 °C. The ligated products were transformed in *Escherichia coli* strain DH5α competent cells (Invitrogen). Recombinant clones were screened with the help of blue/white colony selection and confirmed by restriction digestion. We found three intense bands of 5S rDNA amplified products, which was approximately 200, 400 and 600 bp in lengths. Three bands were eluted from the gel using gel extraction kit (Qiagen), according to the manufacturer's instructions, were cloned and sequenced from one individual of each species. Cloned products of partial 18S rDNA and ITS1 were sequenced by primer walking, whereas other products were sequenced directly using custom services. Sequencing was performed for both the strands in all the samples.

2.3. Sequence alignment and data analyses

Nucleotide composition, GC percentage and number of repeats in all the sequences were calculated with the help of BioEdit (version 1.3). The sequence homology with other fish species listed in the NCBI database was done with the help of nBLAST search (Altschul et al., 1997). For phylogenetic analyses, sequences were analyzed in four data sets (individual gene/region and combined gene/region) as: 18S data, ITS1 data, ITS2 data and combined data (18S + ITS1 + 5.8S + ITS2 + 28S partial). All the sequences of 18S data set and combined data set (18S + ITS1 + 5.8S + ITS2 + 28S partial) were aligned first using the ClustalW software (Thompson et al., 1994) using default setting, resulting in an initial dataset. The 18S sequences showed less variability, so it was easy to align, which was further fine-tuned by manual adjustment. In combined data set, presence of conserved regions on both ends (5′-18S and 3′-28S) helped in alignment which was also adjusted manually for final analyses. The small and highly conserved 5.8S gene region and the small portion of the 28S gene were easy to align, and were used to position

Table 1
List of primers used in this study.

Primer code (for Fig. 2)	Sequence (5′-3′)	Product size (~bp)	Region	Reference
18SF (1F)	AGCATTATGCTTGTCTCAAAGAT	1870	18S complete	Present study
18SR (1R)	CCTTGTTACGACTTTTACTTCCTC			
ITS1F (2F)	TAGGTGAACCTGCGGAAGGATCATTA	880	ITS1 complete	Present study
ITS1R (2R)	CGAGTGATCCACCGCTAAGAGTTG			
5.8SF (3F)	TACAACTCTTAGCGGTGGATCA	161	5.8S complete	Present study
5.8SR (3R)	CACGCCTGTCTGAGGGTC			
ITS2F (4F)	CACTTTGCGGCCCCGGGTTCCT	392	ITS2 complete	Present study
ITS2R (4R)	CCTCTTACCGGTTTCACGCCCT			
28S1F (5F)	CCCGCTGAATTTAAGCATATAAGTAAGCGG	400	28S partial	Zardoya and Meyer (1996)
28S1R (5R)	AACGGTTTCACGCCCTCTTGAACT			
28S2F (6F)	CCCACCCCGACCCGTCTTGAA	600	28S partial	Zardoya and Meyer (1996)
28S2R (6R)	TCGGTTCATCCCGCAGCG			
28S3F (7F)	GGTAAAGCGAATGATTAGAGGTCTT	450	28S partial	Zardoya and Meyer (1996)
28S3R (7R)	GACTGACCCATGTTCAACTGCTGT			
28S4F (8F)	AAGTGGAGAAGGGTTCCATGTGA	500	28S partial	Zardoya and Meyer (1996)
28S4R (8R)	AGAGCCAATCCTTATCCCGAAGTT			
28S6F (9F)	CGGCGGGAGTAACTATGACTCTCTTAAGGT	550	28S partial	Zardoya and Meyer (1996)
28S6R (9R)	CCGCCCCAGCCAAACTCCCCA			
28S7F (10F)	TGAAATACCACTACTCTTATCGTT	661	28S partial	Zardoya and Meyer (1996)
28S7R (10R)	GGATTCTGACTTAGAGGCGTTCAG			
5SF	TACGCCCGATCTCGTCCGATC	201	Complete 5S coding & NTS	Moran et al. (1996)
5SR	CAGGCTGGTATGGCCGTAAGC			

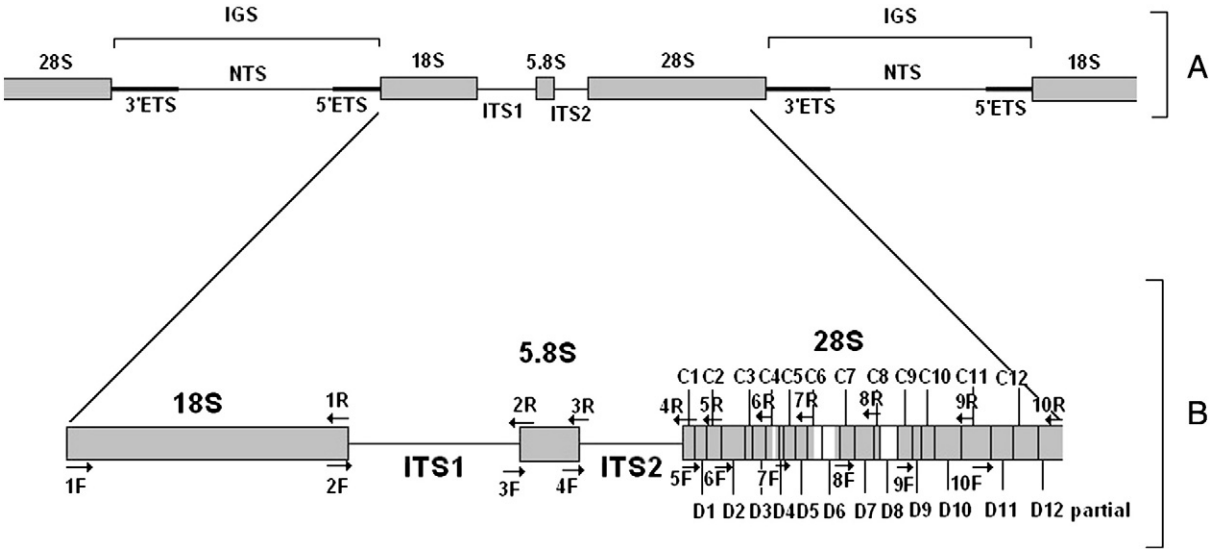

Fig. 2. Diagrammatic representation of 45S rDNA repeat unit. Part A of the image shows the arrangement of 45S rDNA unit and Part B shows the regions that are sequenced in mahseers. White blocks in 28S rRNA gene in part B shows the gaps or non-sequenced regions. Black arrows in Part B show the primer annealing site. This diagrammatic presentation is not to scale.

the highly variable ITS1 and ITS2 regions during the alignment in ClustalW. The boundaries of the ITS1 and ITS2 were determined by comparison of the aligned dataset with ITS sequences of fish taxa available in the NCBI databank. The positional homology, due to the presence of indels and repeats in the 18S, ITS1, ITS2 and combined data sets, was difficult to determine. As a result, ambiguously aligned characters were excluded from the analyses. Modeltest v3.6 (Posada and Crandall, 1998) was used to select a model of sequence evolution for each data set (individual region data and combined data) for neighbor-joining (NJ) analyses. For each data set, we used the best-fit model chosen by the Akaike Information Criterion (AIC) method (Posada and Buckley, 2004). The NJ analyses were implemented in PAUP v.4.0b10 (Swofford, 2003). The models corresponding to each data set were as follows: TrNef + I (18S), TVM + I + G (ITS1 and ITS2) and GTR + I + G (18S + ITS1 + 5.8S + ITS2 + 28S partial). NJ tree for each data was constructed and support for monophyly was assessed with 1000 bootstrap pseudo-replicates (Felsenstein, 1985). We considered nodes with bootstrap values ≥70% (Hills and Bull, 1993), to be well-supported. *Ictalurus punctatus* and *Clarias gariepinus* were used as outgroup species for construction of a phylogenetic tree. Kimura genetic distances for 18S, ITS1, ITS2 and complete set of 18S-ITS1- 5.8S-ITS2-28S partial were also calculated with the help of BioEdit (version 1.3). The correlation coefficient of genetic distance from all the data sets were calculated by Mantel test between species using TFPGA (version 1.3) and significance was tested with two tailed *t*-test to see the relation between the genetic distance of all the analyzed data sets. After sequencing of three bands of 5SrDNA, band I was 201 bp in length, band II was 402 bp in length and band III was 603 bp in the length. When we compared these three sequences manually in all the species, we found that the sequence of band II actually contained two exact repeats of band I while band III contained three exact repeats of band I in all the studies of mahseer species, so 201 bp band sequences were used in further analyses. Nucleotide composition, GC percentage and repeats in all the sequences were calculated with the help of BioEdit (version 1.3) and sequence homology with other fish species listed in the NCBI database was done with the help of nBLAST search.

2.4. Southern blot hybridization

To know the types of repeats of 5S rDNA (1 or 2) present in the genome, we performed the Southern blot hybridization in all the studied species. Genomic DNA (10 μg) from all the species were partially and completely digested with *HindIII* at 37 °C for 30 min and overnight, respectively. Partially to completely digested DNA was subjected to gel electrophoresis in 1% agarose gel and transferred to positively charged nylon membrane, according to Sambrook and Russell (2001). Probe of 5S rRNA gene of *T. putitora*, collected from the river Satluj, was labeled with biotin, according to the manufacturer's instructions (Vector Labs, Burlingame, USA). Hybridization of filter immobilized DNA and post-hybridization washes were performed according to Sambrook and Russell (2001). Hybridized DNA was detected by BCIP/NBT substrate kit (Vector Labs, Burlingame, USA).

3. Results

A total of 25 individuals for 45S rDNA and 5 individual for 5S rDNA (one individual per species) were analyzed for molecular characterization and genetic variability assessments. Species name, individual code, sampling location and GenBank accession number for 45S and 5S rDNAs are provided as Supplementary materials (SM1 and SM2, respectively). Arrangement of major rRNA genes, namely 18S, 5.8S, 28S as well as ITS1 and ITS2 spacers, are presented in Fig. 2. Length and GC content of 45S rDNA regions studied in all the undertaken species are provided as Supplementary material (SM3). The overall base composition of 18S gene is characterized by an over-representation of G (29.82%) and under-representation of T (21.16%), whereas A and C nucleotides were almost distributed in between (A: 22.46% and C: 26.57%). Small differences in terms of single base substitutions, deletions or insertions were found in the nucleotide sequences when compared with other cyprinid sequences, and average similarity of 96%, to other fish species listed in the NCBI database, was observed.

ITS1 shows a high level of length (710–909 bp) and sequence variability (Fig. 3). Overall base composition showed slight bias towards the under-representation of T nucleotide (14.39%), remaining three bases showed almost equal distribution (A: 28.78%, C: 29.49% and G: 27.35%). ITS1 is also characterized by the presence of 36 bp long stretch of GAAA repeat (9 times) in *N. hexagonolepis* and 15 bp long stretch of GAAAT repeat (3 times) in *T. progeneius*. 5.8S rRNA gene was 157 bp long without length variation and GC rich (56.05%). ITS2 also showed a high degree of inter-species variation in the form of length (293–389 bp) and sequence (Fig. 4). ITS2 was found GC rich (66.85%) and showed under-representation of A base (13.09%). Partially sequenced regions of 28S rRNA gene was 2065 bp long in *T. putitora*, 2065 bp long in *T. tor*, 2083 bp long in *T. chelynoides*, 2062 bp long

```
TP-ITS1   CGTGTGGCCCGCCTTGGTCCCGGGAACCCCTGGGGAACCCACCCAAAACAACCAGAAAAA
TPr-ITS1  CGTGTGGCCCGCCTTGGTCTCCCCCACCGTCGAGGCTTCACTCCGGC----CCGGGGGAA
TR-ITS1   CGTGTGGCCCGCCTTGGTCCCGGGAACCCCTGGGGAACCCACAGCCAA-----GGAGAGC
NH-ITS1   CGTGTGGCCCGCCTTGGTCCCGGGAACCCGAGCCTTTGTGACAGCCAA-----GGTGG--
TC-ITS1   CGTGTGGCCCGCCTTGGTCCCGGGAACCCCTGGGGAACCAAGCGGCAA-----AGGCA--
          ******************* *    ***   *                    *

TP-ITS1   GAGAGCCTGCGACGAACAGCGCAGG-ACGAGGAGGCCCCGGAGGAAGGTCGGGACGGGTT
TPr-ITS1  GGGGGGGGGCGCCTCCCCCTACGGGTACCCTGTTGGCCCGGCGCCCTGCCGGGACCGTGG
TR-ITS1   CTGCGACGAACAGCGCAGGCTCCGGCGCGGCCTTCCCGCACCCCGTCCCGCCGGGCGCCA
NH-ITS1   -CGGCAGAGTCAGCCGGGGCTCCGGCGCGGCCTTCCC-CCCCTCTCCGCACGAGACGAAA
TC-ITS1   ---GGACCCCGAAACGTACCTTCGGCGC----TCTCCCCCACCGTCGAGGCCGTGCGGCA
                              **  *       * *                 *

TP-ITS1   TGTTTGTGACAGCCAAGGGTGGCGGGCACCCGTCCCGCCGGGCGCCACAGAAAAAAGTCG
TPr-ITS1  GTTCCAAGACCCCCCGTAATAGGGGGCGCCCGTCCGG--GGTTACAACCCCACTTGTTAA
TR-ITS1   CATTTGTT----TGTGAGAGAGAAGGCATCGGTCCGCGAGAATCCTTCATCAAAAACTGG
NH-ITS1   CAAAACTCACAACCGGAGAGAGAAGGCATCGGTCCGCGAGAATCCTTCATCAAAAACTGG
TC-ITS1   AAGCCCTT-------GGGGACGAA-----CGACCC-CAAAAAAATTCAAACTGAAGTCGT
                         *       *   **

TP-ITS1   TAACAAAAACTGGAAAGGATCATTACCACGCGGCAAAGGCAGGACCCC--GAACGTACCT
TPr-ITS1  CAGAAGACATAGGTAATGGATGAGCGGACGTGCCTAAGGAGTGAAGGC--ACAGAGGCCC
TR-ITS1   GGCAAAAACCTGGGAAGGATCATTATC---CGGCAAAGCTATCGGAACTGATACGTACCT
NH-ITS1   GGCAAAA------GAAGGATCTGTGTC---CGGCAAAGCTATCGGAACTGATACGTACTT
TC-ITS1   AACAGAAAAACTGGAAGGATCATTGCAAAGCGGCAAAGGCAGGAACCCCGATACGTACCT
                *       ** *             * * ***         *    *    *

TP-ITS1   TCGGCGCTCTCCCCCACCGTCGAGGAGGGGGGGGGCGCCGCCCCCGACGGGTACCCTGTA
TPr-ITS1  TGG----CTTGACTCGGCCTCGAGCTTGGTGCAGCCTCAGGTCT-AAGGACCACCCTGGA
TR-ITS1   TCGGTACTCTCTTCCTCACCCCGGG-GGAGGGAGACGCCGCCCCCGATGTTTACCCTGCA
NH-ITS1   TCGGTGCTCTCTCCCTCACCCCGGG-GGAGGGAGACGCCGCCCCCGATGTTTACCCTGCA
TC-ITS1   TCGGCGTTCTCCCCC--------GA-GGGGGGGGGCGCCGCCCCCGACGGGTACCCTGTT
          * *      *    *        *   *  *  * * * *  *   * *   ******

TP-ITS1   GGCCCGGCGCCCTGCCGGGACCGTGGGTTCAAAGACCCCCCGTAATAGGGGGCGCCCGTC
TPr-ITS1  GCCGTGAAGGCCAACGGGAAAAAGAAAGACAGAAAAACCCAACCCAAAGAGGCCCTCCTC
TR-ITS1   GGCCCAGCGCCCTGCCGGGACCGTGGGTTCAAAGACCCCCCGTAATAGGGGGCGCCCGTC
NH-ITS1   GGCCCAGCGCCCTGCCGGGACCGTGGGTTCAAAGACCCCCCGTAATAGGGGGCGCCCGTC
TC-ITS1   GGCCCGGCACCCCGCCGGGACCGTGGGTTCAAAGACCCCCCGTAATAGGGGGCGCCCGTC
          * *       **  * ** *         ** * *  ***      * * *** * * **

TP-ITS1   AAGGGGGAAAAAGAAACAAAGCAAGAAAGAAATGAAATGAA-ATGAAATATGAGCAAAAG
TPr-ITS1  GGGTGGGTCCGAGAGACCCGGCCTGGAGTGGAGCAGGCAAACGGGAAACCCAACCCAAAG
TR-ITS1   CGGGGTTACGAAGAAACAA----AGTCCAAGACAAGGTAAC--GGACGGACGTGCCTAAG
NH-ITS1   CGGGGTTACGA-----CCC----CTTCCAAGACAAGGTAAC--GGACGGACGTGCTTAAG
TC-ITS1   CGGGGTTACAACCCCACTTGTTTTTTTCGAGACAAGGTAAC--GGA----CGTGCCTGAG
            * *           *              *  *    *    **        *   **

TP-ITS1   AGGTCTTGGT-----GCAGCCTCAGGTCTAAGATGACC-ACCCTGGAGTGGCGAAGGCAA
TPr-ITS1  AGGCCCTGACC----ACAGCCTCGGGT-----------------GAGCGGAGAAGGCAG
TR-ITS1   GAGCCCTGGC------CGGCCTCGGGTCTAAGAGGCCTGGACTTGGAGCGGCGAAGGCGA
NH-ITS1   GAGCCCTGGCTTGACTCGGCCTCGGGTCTAAGAGGCCTGGACTTGGAGCGGCGAAGGCGA
TC-ITS1   GAGTGGAGGCACGG--AGGCCCTGGCTTGACTCGGCCTCGAGCGGCAAAGGCAAAAG-AA
            *    *          ***   * *                  *  **  ** *
```

```
TP-ITS1   ACGGGGGAAAA-------AGAAAGACAGAAA-------AACCTAAAAAAAAGAGGCACT-
TPr-ITS1  ACGGGAGAAA---------GAAAGACAGAAA-------AGCCCACATGAAAGAGGCACT-
TR-ITS1   ATGGGAAAAAGAAAGAAAAGAAAGAAAGAAAGAAGGAAAACCTAAGCAAAAGAGGCCCT-
NH-ITS1   ATGGGAAAAA--------AGAAAGAAAGAAAGAAAGAAAACCTAAGCAAAAGAGGCCCTA
TC-ITS1   ACAAACGCACAAAGCGAGCGAAAGACATAAAGAAATGAAATATAAACGAAAGAGGCCAT-
          *       *          ****** * ***       *    *    ********  *

TP-ITS1   -----GCCTCGGGTCTCGTACCCATCTGAGAGAGGGGAGGCCCCCCTCTCCAAAAAGC--
TPr-ITS1  -----GCCTCGGGTCTCGTACCCATCCGAGAGAGGGGAGGCCCCCCCCTCTCCAAAGCGA
TR-ITS1   -----GCCTCGGGTCTCAGACCCCTCCGAGAGAGGGGAGGTCCCCCCCCCCCCCCCAAAAA
NH-ITS1   ACTCAGCCTCGGGTCTCAGACCCCTCCGAGAGAGGGGAGGTCCCCCCCCCCCCCCCAAAAAA
TC-ITS1   -----------GGTGGTGGAAAGCCTCGGGTCCAGGAAA--CCCACCCTGAAGCGGCGAA
                    ***     *       * *    ** *   *** * *

TP-ITS1   -----TCAAAGCGACACAGAGCA--AAAG--CTCGAAACGAAACACGGC--AATGCACCG
TPr-ITS1  TGTGCTCAAAGCGATAAAAACGAGCAATG--CTCGAAACAAAACACGGCATAACACACCG
TR-ITS1   AG---CCAAAGCGACACAGAGC--AAAAGCACTCGGAACAAA---CGGC--AAAGAACCG
NH-ITS1   AG---CTCAAACGAAAAAGCGATAAAGAGCACTCGGAACAAA---CGGC--AAAGAACCG
TC-ITS1   GG---AAAAAGAAAGACAGAAACACAAAGCTCTAAACACAAAACACGGC--AATGCTCCG
                  **   * * *       *  *  **    ** **   ****  **    ***

TP-ITS1   TGAGA-------GGCCCACAGAAGGCCCCAAAAAGCTACAATGCTCAAGCCGTCAAAGAC
TPr-ITS1  TGAGA-------GGCCCACTGAAGGCCCCAAAAAGCTACCATGCTCAAGCCGTCAAACAC
TR-ITS1   TGATA-CCAAA-GGCCCAGAGAAGGCCTCAAAA---------GCTCAAACTGTGAAACAA
NH-ITS1   TGATA-CCAAAAGGCCCAGAGAAGGCCTCAAAA---------GCTCAAACTGTGAAACAA
TC-ITS1   TGATAACTGAAAGGCCCACAGAAGGCCCCGAAAAGCTACAACGCTCAAGCCGTCAAACAA
          *** *       ******  ******* * ***         ****** * ** *** *

TP-ITS1   TCT-AATAACACGGCATAA-----AAACGGCGAGACGTCGGGGTCCT------GGGGACG
TPr-ITS1  TCTTGATAACACGGCATAA-----AAACGGCGAAACGTCGGGGTCCTCCCCC-GGGGACG
TR-ITS1   AGTCTCTAACACGACACAAGAATAAAACCTCAAACCATGGGG---------------ACA
NH-ITS1   AGTCTCTAACACGACACAAAAATAAAACCTCAAACCATGGGGGACCC------GGGGACA
TC-ITS1   ACACTCTGACACGGCATAA-----AAACGTCGAAACGTGGGGGGTCCTCCCCTGGGGACG
                 * ***** ** **     ****  * *  * * ***             **

TP-ITS1   GG-TGGACGCCCCAAGACAAGGCACCGGTCCCCGCCAACCCCCCC--TCTGTTGTGTCTG
TPr-ITS1  GGGTGGACGGCCCAAGACAAGGCACCGGTCCCCGCCAACCCCCCCCCTCTGTTGTGTCTG
TR-ITS1   GGGTGGATGCCCGAAGACAAGGCACCGGTAAACACCCTTCCCCCCCTG---ATGTCTGTA
NH-ITS1   GGGTGGATGCCCGAAGACAAGGCACCGGTAAACACCCTTCCCCCCCTTTGCATGTCTGTA
TC-ITS1   GGGTGGACGCCCCGAGAAAAGGCACCGGTCCCCGCCGACAACCCCCCCCCCCTGTTTGTC
          ** **** * **  *** ***********   * **     ****       *** * *

TP-ITS1   IAT--TGGCCTATTACCTCAAACAAA-GAGTA--
TPr-ITS1  TAT--CGGCCTATTACTTCAAACAAA-AAGAGTA
TR-ITS1   TCT--CAGCCTATTACCTCAAACAAA-GAGTA--
NH-ITS1   TCT--CAGCCTATTACCTCAAACAAA-GAGTA--
TC-ITS1   TTTACTGGCCAATTACCTCAAACAAAAGAGTA--
          * *    *** ***** *********  **
```

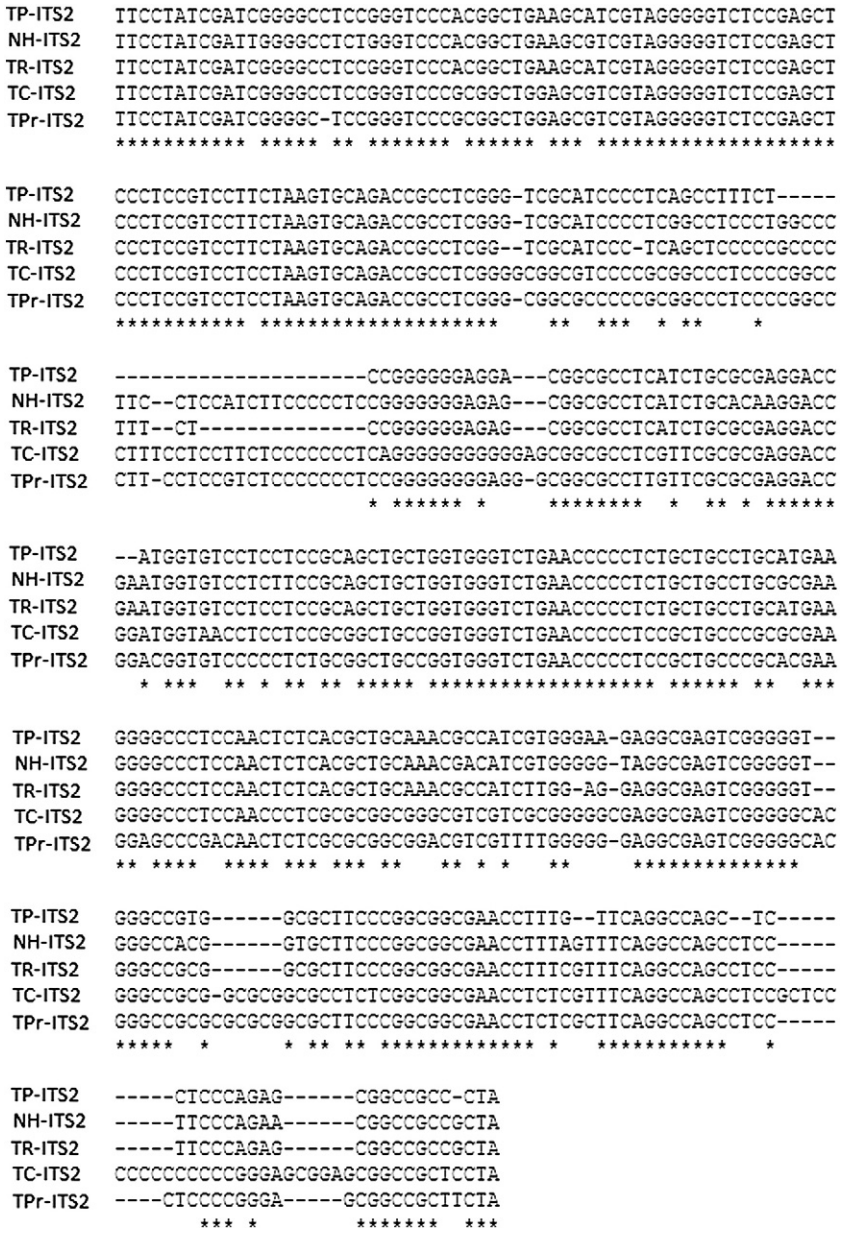

Fig. 4. Nucleotide sequence alignment of internal transcribed spacer 2 (ITS2) region from mahseer species reported in present study (*show the conserved region). The consensus sequence was generated from all the individuals of a species for final alignment which represent the sequence of one species. (TR = *T. tor*, TP = *T. putitora*, TC = *T. chelynoides*, TPr = *T. progeneius* and NH = *N. hexagonolepis*).

in *T. progeneius* and 2053 bp long in *N. hexagonolepis*. 28S rRNA gene contained several conserved and divergent domains. We sequenced conserved domain C1, C2, C7, C9, C10, C11 and divergent domain D1, partial D6, D7, D9, D10, D11, and partial D12 of 28S gene. Overall average base compositions were: A: 23.03%, C: 25.07%, G: 32.33% and T: 19.57%. The divergent domains of 28S gene were characterized by a slight bias to exclude T and for high GC content. The GC rich composition of 28S rRNA gene pattern reflects their potential capability to fold into secondary structures with particularly large GC rich stems.

After sequencing, the length of single repeat of minor 5S rDNA family was found to be 201 bp in all the mahseer species studied. The coding 5S rDNA region shared a highly conserved gene of 120 bp long contained three elements, namely Box A, Box C and IE, of the internal control regions (Fig. 5), which functions as a promoter for the gene (Hallenberg et al., 1994). Nucleotide compositions of Box A, Box C and IE are AGCTAAGCAGGGTCG (location: +48 bp to +62 bp), TGGATGGGAGACCGCCTG (+78 bp to +98 bp) and CCGGGT (+65 bp to +70 bp), respectively. Non-transcribed spacer (NTS) of 5SrDNA was 81 bp long variable region and contained TATA box like sequences in all five mahseer species, but the number of these boxes were found to be three in *T. putitora* and two in *T. tor*, *T. chelynoides*, *T. progeneius* and *N. hexagonolepis*. The comparative analyses of 5S rDNA sequences among the five species under study showed 27.93% variability in NTS region. The coding sequences of

Fig. 3. Nucleotide sequence alignment of internal transcribed spacer I (ITS1) region from mahseer species reported in present study (*show the conserved region). The consensus sequence was generated from all the individuals of a species for final alignment which represents the sequence of one species. (TR = *T. tor*, TP = *T. putitora*, TC = *T. chelynoides*, TPr = *T. progeneius* and NH = *N. hexagonolepis*).

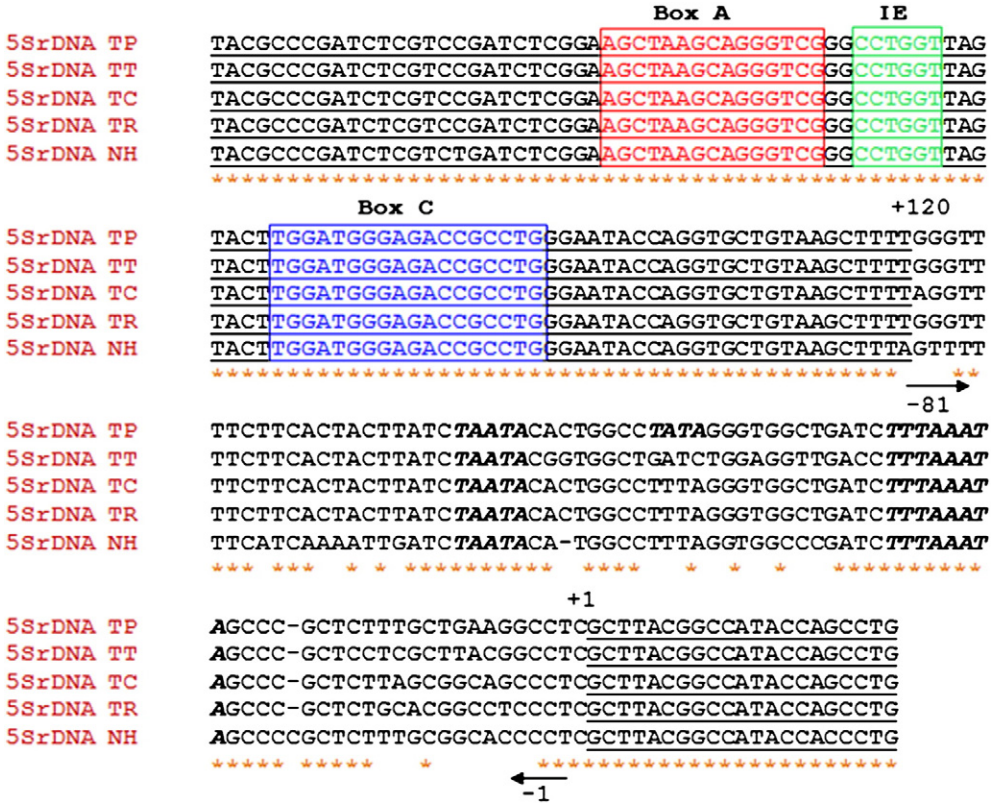

Fig. 5. Diagrammatic representation of 5S rDNA repeat unit. Coding region of 5S rRNA gene (underlined region) shows the internal promoter region, namely Box A, IE (intermediate element) and Box C. NTS region of the gene shows TATA like elements. (NH= *N. hexagonolepis*, TP= *T. putitora*, TT= *T. tor*, TC= *T. chelynoides*, TR= *T. progeneius*).

5SrDNA in *N. hexagonolepis*, after searches using BLASTn program (Hallenberg et al., 1994), showed an average sequence similarity of 95.74% with other fishes belonging to the order Cypriniformes, Characiformes, Salmoniformes, Gadiformes, Perciformes and Tetraodontiformes, while NTS showed extensive variations. To examine that how many types of repeats of 5S present in mahseers, genomic DNA of all the five specieswere partially and completely digested with *HindIII*, which cleaves only once in the 5S rRNA gene, and then transferred to positively charged nylon membrane by Southern blotting and hybridized with 5S rRNA gene probe. Analysis of band pattern of partially digested DNA showed identical results in all the species examined. Southern blot showed ladder of exact integers of 201 bp with partially digested DNA (Fig. 6A), while complete digestion showed only one 201 bp long band after southern blot detection (Fig. 6B) and conformed that only one type of 5S rDNA repeat was present in all the studied species.

For phylogenetic analyses, nucleotide sequences were divided into four datasets and *C. gariepinus* (GenBank Accession no. AJ876383) and *I. punctatus* (GenBank Accession no. AF021880) were used as outgroup species. The first dataset contained a total of 27sequences of 18S rDNA from five mahseer species (*T. putitora*: 6, *T. tor*: 5, *T. chelynoides:* 5, *T. progeneius*: 4, *N. hexagonolepis*: 5), one each of *C. gariepinus* and *I. punctatus*. A phylogenetic tree of 18S rDNA data set was constructed by the NJ method (Fig. 7). The studied in-group mahseer are a monophyletic group, well separated from the outgroup species. All the specimens of *N. hexagonolepis* form monophyletic group from *Tor* species, indicating that *Neolissochilus* is a sister genus of *Tor*, and they are sufficiently different to warrant generic status. However, monophyletic status of these must be cautiously interpreted because not all described species of *Neolissochilus* were examined in the present study. The second dataset contained ITS1 sequences in all in-group species as well as two outgroup species and NJ phylogenetic tree was constructed (Fig. 8). NJ tree constructed using ITS1 sequences showed almost identical homology with 18S rDNA NJ tree. ITS1 NJ tree also showed monophyletic mahseer group that clearly separated from outgroup species. Monophyletic arm of mahseer further divided into two arms, where one arm contained all individuals of *N. hexagonolepis* and second arm contained the individuals of *T. tor*, *T. putitora*, *T. chelynoides* and *T. progeneius.*

The third data set contained ITS2 sequences and the fourth dataset contained 18S-ITS1-5.8S-ITS2-28S (partial) sequence in all in-group species and two outgroup species, respectively, and the NJ phylogenetic trees were constructed (Figs. 9 and 10). The NJ tree based on ITS2 sequence showed somewhat different topology from

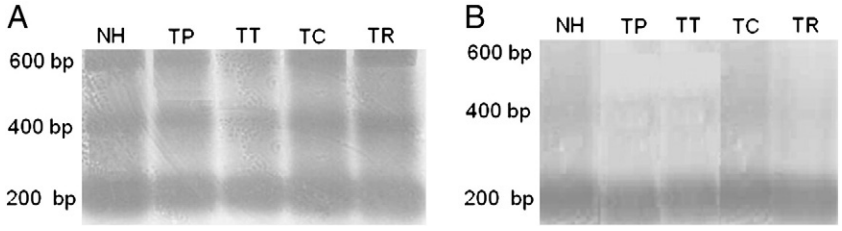

Fig. 6. Southern blot of 5S rDNA in different mahseer species: (A) partially digested, and (B) completely digested (*HindIII*) genomic DNA in specimen of *N. hexagonolepis* (NH), *T. putitora* (TP), *T. tor* (TT), *T. chelynoides* (TC) and *T. progeneius* (TR).

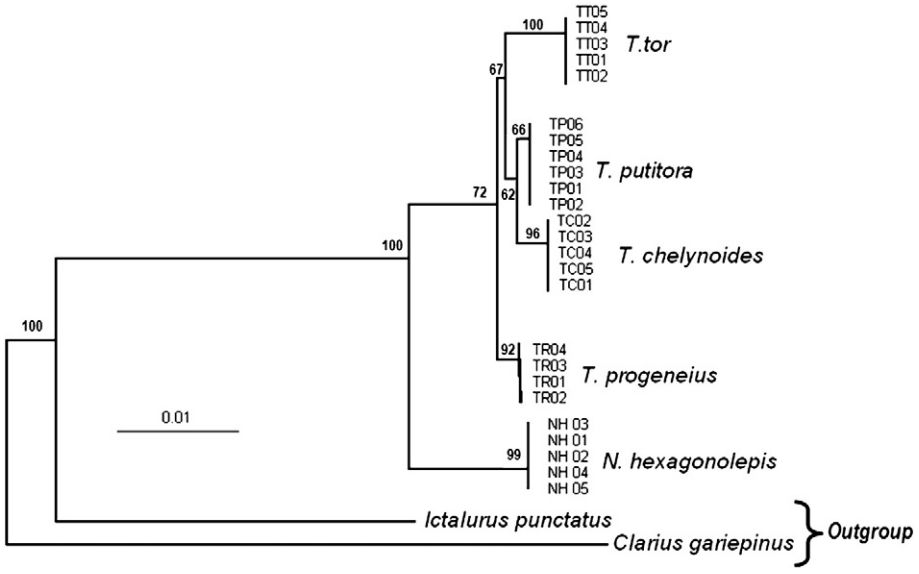

Fig. 7. NJ phylogenetic tree constructed using 18S rDNA sequences in undertaken species. 18S sequences of *I. punctatus* and *C. gariepinus* were taken from NCBI database.

the 18S and ITS1 NJ trees, while the NJ tree of combined 18S-ITS1-5.8S-ITS2-28S sequences showed almost identical topology with 18S and ITS1 NJ trees. In ITS 2 based NJ tree, *T. chelynoides* showed close relation with *T. tor*, while in NJ tree based on 18S, ITS1 and combined 18S-ITS1-5.8S-ITS2-28S sequences, the position of *T. chelynoides* was taken by *T. putitora*. Kimura genetic distances were calculated between all the species using each data set. Three genetic distance matrices, namely 18S, ITS1 and ITS2, were compared with 18S-ITS1-5.8S-ITS2-28S (partial) matrix and correlation coefficients were estimated by Mantel test to see the correlation between all the datasets. Significant positive correlation (between 18S and 18S-ITS1-5.8S-ITS2-28S $r = 0.837$, $p < 0.05$; between ITS1 and 18S-ITS1-5.8S-ITS2-28S $r = 0.653$, $p < 0.05$; and between ITS2 and 18S-ITS1-5.8S-ITS2-28S $r = 0.893$, $p < 0.05$) were found between all the compared genetic distance matrices.

4. Discussion

4.1. Structure of rDNA in mahseer fishes

As expected, 18S rRNA gene showed a high degree of similarity among all the mahseer species studied as well as with other fishes listed in the NCBI database, whereas both the spacers (ITS1 and ITS2) showed great variability when compared between the mahseer species and with closely related group of fish species. The length and GC content of the 18S and 5.8S genes of the studied mahseers fall within the range of known sequences available in GenBank/EMBL. The same holds true for the length and GC content of the spacers. A comparably high GC content in ITS1 and ITS2 is considered to be a general property, as suggested by Torres et al. (1990). The divergent domains of 28S rRNA gene are GC rich. This base composition pattern

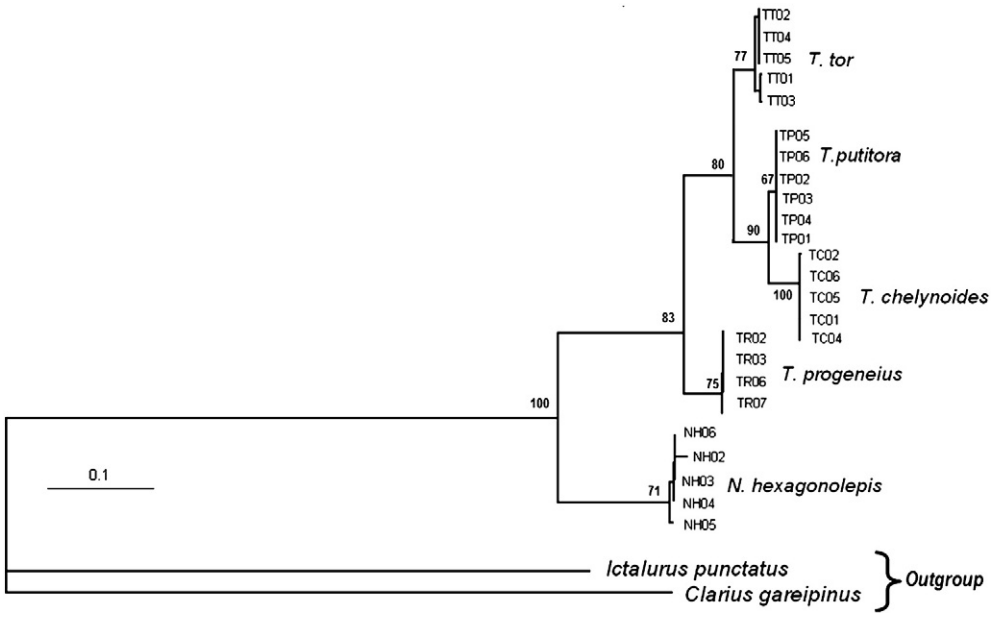

Fig. 8. NJ phylogenetic tree constructed using ITS1 sequences in undertaken species. ITS1 sequences of *I. punctatus* and *C. gariepinus* were taken from NCBI database.

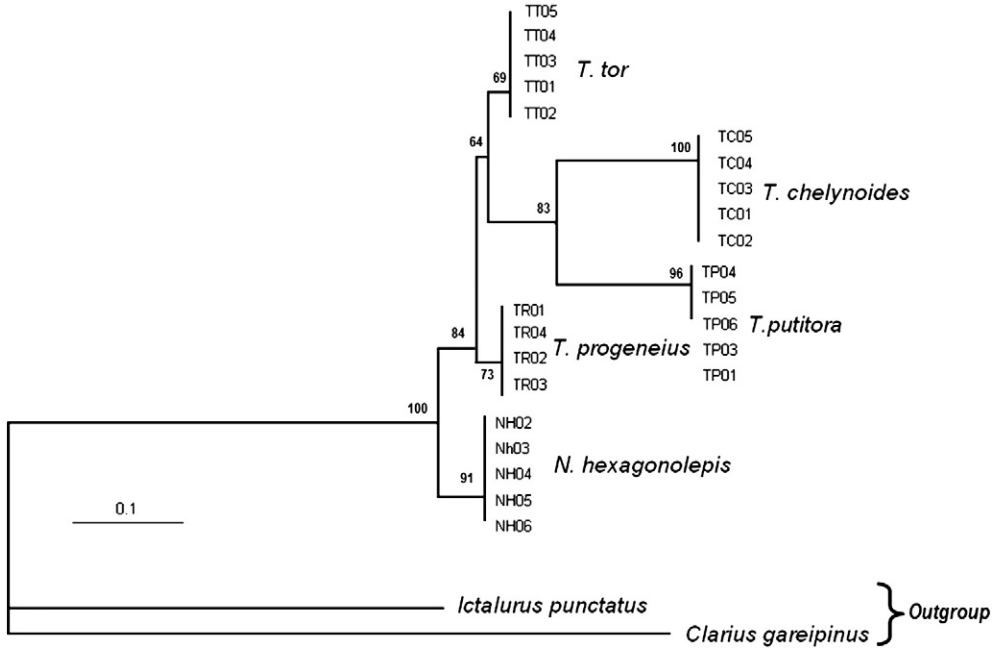

Fig. 9. NJ phylogenetic tree constructed using ITS2 sequences in undertaken species. ITS2 sequences of *I. punctatus* and *C. gariepinus* were taken from NCBI database.

reflects their potential capability to fold into secondary structures with particularly large GC rich stems (Vawter and Brown, 1993). However, the conserved core of the molecule has an almost equal distribution of the different types of nucleotides that is correlated with the presence not only of stem loop structure but also long single-stranded A rich regions (Gutell et al., 1985), which are thought to interact with other rRNA subunits and ribosomal proteins (Vawter and Brown, 1993).

Molecular organization of 5S rDNA in cyprinid fishes is still not fully known. The coding region of 120 bp long 5S was highly conserved even in distinct taxa, while NTS region was highly variable and species specific. Moreover, it has been demonstrated earlier that conserved sequences, which influence the transcription level of 5S rDNA in mammals, can also be located in the NTS region (Nederby-Nielson et al., 1993; Suzuki et al., 1996). Although little is known about the NTS sequences among fishes, a TATA like sequence has also been observed in the NTS region of 5S in some other fishes (Inafuku et al., 2000; Martins and Galetti, 2000; Pendas et al., 1994; Tigano et al., 2004; Wasko et al., 2001) that may be involved in the transcription of these genes. An AT rich spacer region in mahseers showed clear agreement with other fish species studies, such as Atlantic salmon and bitterling (Acheilognathinae) (Fujiwara et al., 1998; Sajdak et al., 1998). Comparison of NTS region from selected fishes and mahseers studied here demonstrated the presence of 81 bp long NTS which have the optimum conserved sequences required for the organization and/or activity of 5S rDNA, but the work

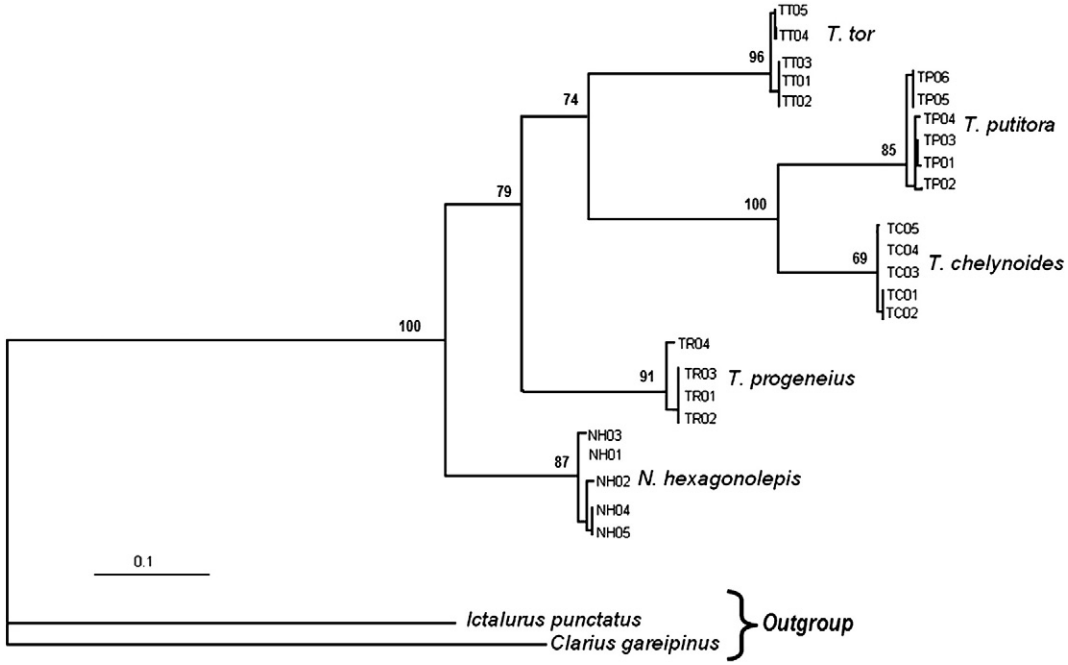

Fig. 10. NJ phylogenetic tree constructed using complete 18S, ITS1, 5.8S, ITS2 and partial 28S (approx. 5.2 kb) sequences in undertaken species. For outgroup, sequences of *I. punctatus* and *C. gariepinus* were taken from NCBI database.

on a variety of organisms has shown even shorter 5S rDNA NTS sequences. Fujiwara et al. (1998) demonstrated a conserved 120 bp sequence of 5S rRNA gene and a short 56–67 bp NTS with two distinct portion, a conserved (5′-flanking portion at position −1 to −38) and a variable part (3′-flanking portion), in six of ten species/subspecies examined from four genera: *Acheiloganathus*, *Pseudperilampus*, *Rhodeus* and *Tanakia*. Martins and Wasko (2004) and Santos et al. (2006) showed 62 bp long NTS in fish. In a report on primitive bony fish *Polypterus senegalus*, 5S rDNA copies also repeated in tandem comprises a conserved 120-nucleotide coding sequence, separated by non-transcribed spacers (NTS) sequences with variable length. Two different classes of 5S rDNA were observed in *P. senegalus*, with different NTS clones that probably represent paralogous copies that have evolved independently (Morescalchi et al., 2008). In cartilaginous fishes, 5S rRNA of *Raja montagui* and *Taeniura lymma* contained the coding region of 120 bp, which showed 100% identity with that of other fish present in the GenBank database, while lengths of NTS were variable. Like teleost fishes, the NTS region of cartilaginous fishes also have TATA-like and $(TGC)_n$ trinucleotides, $(CA)_n$ dinucleotides and $(GTGA)_n$ tetranucleotides (Rocco et al., 2005). Organization of 5S rDNA, examined by Southern blot hybridization, showed the presence of single type of 5S rDNA repeat in mahseers. Detection of a ladder of exact integers of 201 in partially digested DNA also provided the evidence that these 5S units were organized in repeats of tandem array.

4.2. Phylogenetic relationship of Indian mahseers

In present study, attempts are made to reconstruct a robust phylogeny based on nuclear DNA sequences (rRNA gene) of the mahseer species found in India. Nearly all previous attempts to estimate relationships between species of mahseer have been based on only morphological characters, except one report by Nguyen et al. (2008). After a long period of time since the first description of *Tor* (1833), Rainboth (1985) erected the new genus *Neolissochilus* based on the absence of a median lobe and other characters such as gill raker counts and pharyngeal arches. In the present study, all individuals of *N. hexagonolepis* formed a monophyletic group with each dataset analyzed, indicating that *Neolissochilus* is a sister genus of *Tor* and they are sufficiently different to warrant generic status. However, their monophyletic status must be cautiously interpreted because not all described species were examined in the present study. In phylogenetic analyses, the outgroup species were well separated from mahseers and further the mahseer species were divided into two lineages, one has all the individual of *N. hexagonolepis* and second has all the species of *Tor*. *T. tor* showed maximum distance from *N. hexagonolepis* and minimum distance with *T. putitora* or *T. chelynoides* and *T. progeneius* showed closer relation to *N. hexagonolepis*. Three phylogenetic trees based on 18S, ITS1 and combined 18S-ITS1-5.8S-ITS2-28S sequences showed almost identical topology, except ITS2.

Although the numeric chromosomal changes have not been observed in this mahseer group, the structural rearrangements might have played an important role in the speciation. These species showed different karyotype formula and location of $AgNO_3$, CMA_3 and rDNA FISH positive signals on the chromosomes. Among the undertaken *Tor* species, *T. chelynoides*, *T. progeneius* and *T. putitora* possessed 2 and 3 pairs of silver stained and 18S FISH positive signals, whereas *T. tor* possessed 4 and 5 pairs, respectively. Rearrangements such as pericentric inversions at different lengths of chromosomes have played an important role in karyotypic diversification among the mahseer species.

5. Conclusion

N. hexagonolepis formed a monophyletic group with each dataset analyzed, indicating that *Neolissochilus* is a sister genus of *Tor* and they are sufficiently different to warrant generic status. When comparing methods, phylogenetic reconstruction by individual rRNA gene regions produced identical results, except ITS2. These findings advocated the use of any rRNA sequences from above mentioned three datasets, namely 18S, ITS1 or combined 18S-ITS1-5.8S-ITS2-28S to draw the phylogeny because comparisons between all the genetic distance matrices from all the data sets also showed positive correlation values. So there may not be a need to sequence the entire repeat of 45S rDNA, even 18S rDNA sequence is sufficient to draw the accurate phylogeny, because in undertaken species the 18S NJ tree showed almost identical topology with complete 18-ITS1-ITS2-28S NJ tree. Even using ITS1 spacer sequences, which have high degree of variability, the NJ tree showed identical topology with complete 18-ITS1-5.8S-ITS2-28S sequence. Phylogenetic hypotheses of the evolutionary relationships among members of the same genus provide frameworks for comparative research on mechanisms of diversification and speciation. These phylogenies are also valuable resources for people concerned with conservation in which they provide a relatively objective means of quantifying evolutionary distinctiveness and resolving taxonomic ambiguities involving rare taxa. The findings from the present study provide useful insights into taxonomic status of mahseer, and set the stage for future investigations dealing with phylo-geography, taxonomy, conservation and co-evolution within this interesting and important group of fishes. Inclusion of additional taxa would warrant further investigations.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2013.06.032.

Conflict of interest statement

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

Acknowledgment

The authors are thankful to the Director, NBFGR, Lucknow and Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi for the financial support.

References

Altschul, S.F., et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.
Arnheim, N., 1983. Concerted evolution of multigene families. In: Nei, M., Koehn, R.K. (Eds.), Evolution of Genes and Proteins. Sinauer, Sunderland, MA, pp. 38–61.
Campo, Danial, Machado-Schiaffino, Gonzalo, Horreo, Jose Luis, Gracia-Vazquez, Eva, 2009. Molecular organization and evolution of 5SrDNA in the genus *Merluccius* and their phylogenetic implication. J. Mol. Evol. 86, 208–216.
Eschmeyer, W.N., Ferraris Jr., C.J., Hoang, M.D., Long, D.J., 2004. The catalog of fishes, on-line, species of fishes. http://www.calacademy.org/research/ichthyology/catalog/.
Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.
Ferreira, I.A., Oliveira, C., Venere, P.C., Galetti Jr., P.M., Martins, C., 2007. 5S rDNA variation and its phylogenetic inference in the genus *Leporinus* (Characiformes: Anostomidae). Genetica 129 (3), 253–257.
Fritz, G.N., Conn, J., Cockburn, A., Seawright, J., 1994. Sequence analysis of the ribosomal DNA internal transcribed spacer 2 from populations of *Anopheles nuneztovari* (Diptera: Culicidae). Mol. Biol. Evol. 11, 406–416.
Fujiwara, A., Abe, S., Yamaha, E., Yamazaki, F., Yoshida, M.C., 1998. Chromosomal localization and heterochromatin association of ribosomal RNA genes loci and silver stained nucleolar organizer regions in salmonid fishes. Chromosome Res. 6, 463–471.
Fujiwara, M., et al., 2009. Molecular organization of 5S rDNA in bitterlings (Cyprinidae). Genetica 135 (3), 355–365.
Gebri, S.A., 1985. Evolution of ribosomal DNA. In: MacIntyre, R.J. (Ed.), Molecular Evolutionary Genetics. Plenum, New York, pp. 419–517.
Gutell, R.R., Weibser, B., Woese, C.R., Noller, H.F., 1985. Comparative anatomy of 16S-like ribosomal RNA. Prog. Nucleic Acid Res. Mol. Biol. 32, 155–216.
Hallenberg, C., Nederby-Nielson, J., Frederiksen, S., 1994. Characterisation of 5S rRNA genes from mouse. Gene (Amst) 142, 291–295.
Haque, M.Z., Mahata, S.C., Ali, M.Z., Hussain, M.G., Mazid, M.A., 1995. Seed production of Himalayan mahseer, *Tor putitora* (Ham.) in Bangladesh. J. Asiat. Soc. Bangl. Sci. 21, 279–282.

Hassouna, N., Michot, B., Bachellerie, J.P., 1984. The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. Nucleic Acid Res. 12, 3563–3583.
Hills, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.
Imsiridou, A., Minos, G., Katsares, V., Karaiskou, N., Tsiora, A., 2007. Genetic identification and phylogenetic inferences in different Mugilidae species using 5S rDNA markers. Aquac. Res. 38, 1370–1379.
Inafuku, J., Nabeyama, M., Kikuma, Y., Saitoh, J., Kubota, S., Kohno, S., 2000. Chromosomal location and nucleotide sequences of 5S ribosomal DNA of two cyprinid species (Osteichthyes, Pisces). Chromosome Res. 8, 193–199.
Ingram, B.A., Sungan, S., Gooley, G., Sim, S.Y., Tinggi, D., De Silva, S.S., 2005. Induced spawning, larval development and rearing of two indigenous Malaysian mahseer, *Tor tambroides* and *T. douronensis*. Aquacult. Res. I36, 1001–1014.
Jansen, G., Devaere, S., Weekers, P.H.H., Adriaens, D., 2006. Phylogenetic relationships and divergence time estimate of African anguilliform catfish (Siluriformes: Clariidae) inferred from ribosomal gene and spacer sequences. Mol. Phylogenet. Evol. 38, 65–78.
Martins, C., Galetti, P.M., 2000. Conservative distribution of 5S rDNA loci in *Schizodon* (Pisces, Anastomidae) chromosomes. Chromosome Res. 8, 353–355.
Martins, C., Wasko, A.P., 2004. Organisation and evolution of 5S ribosomal DNA in the fish genome. In: Williams, C.R. (Ed.), Focus on Genome Research. Nova Science Publishers, Hauppauge, pp. 289–319.
Moran, P., Martinez, J.L., Garcia-Vazquez, S., Pendas, A.M., 1996. Sex linkage of 5S rDNA in rainbow trout (*Oncorhynchus mykiss*). Cytogenet. Cell Genet. 75, 145–150.
Morescalchi, M.A., Liguori, I., Rocco, L., Archimandritis, A., Stingo, V., 2008. Karyotypic characterization and genomic organization of the 5S rDNA in *Polypterus senegalus* (Osteichthyes, Polypteridae). Genetica 132, 179–186.
Nederby-Nielson, J., Hallenberg, C., Frederiksen, S., Sorensen, P.D., Lomholt, B., 1993. Transcription of human 5S rRNA genes is influenced by an upstream DNA sequence. Nucleic Acids Res. 26, 3631–3636.
Nelson, J.S., 2006. Fishes of the World, 4th edition. John Wiley & Sons, Inc., Hoboken, New Jersey 601.
Ng, C.K., 2004. King of the rivers: mahseer in Malaysia and the region. Inter Sea Fishery.(M) SDN BHD, Kuala Lumpur 170.
Nguyen, T.T.T., Ingram, B.A., Sungan, S., Gooley, G., Sim, S.Y., Tinggi, D., De Silva, S.S., 2006. Mitochondrial DNA diversity of broodstock of two indigenous mahseer species, *Tor tambroides* and *T. douronensis* (Cyprinidae) cultured in Sarawak, Malaysia. Aquaculture 253, 259–269.
Nguyen, T.T.T., Uthairat, N.N., Srijanya, S., Chen, Z.M., 2008. A study on phylogeny and biogeography of mahseer species (Pisces: Cyprinidae) using sequences of three mitochondrial DNA gene regions. Mol. Phylogenet. Evol. 48, 1223–1231.
Olsen, G.J., Woese, C.R., 1993. Ribosomal RNA: a key to phylogeny. FASEB J. 7, 113–123.
Pendas, A.M., Moran, P., Freije, J.P., Garcia-Vazquez, S., 1994. Chromosomal mapping and nucleotide sequences of two tandem repeats of Atlantic salmon 5S rDNA. Cytogenet. Cell Genet. 67, 31–36.
Porter, C.H., Collins, F.H., 1991. Species-diagnostic differences in a ribosomal DNA internal transcribed spacer from sibling species *Anopheles freeborni* and *Anopheles hermsi* (Diptera: Culicidae). Am. J. Trop. Med. Hyg. 45, 271–279.
Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.
Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Rainboth, W.J., 1985. *Neolissochilus*, a new genus of south Asian cyprinid fishes. Beaufortia 35, 25–35.
Roberts, T.R., 1999. Fishes of the cyprinid genus *Tor* in the Nam Theun watershed (Mekong Basin) of Laos, with description of a new species. Raffles Bull. Zool. 47, 225–236.
Rocco, L., Costagliola, D., Fiorillo, M., Tinti, F., Stingo, V., 2005. Molecular and chromosomal analysis of ribosomal cistrons in two cartilaginous fish, *Taeniura lymma* and *Raja montagui* (Chondrichthyes, Batoidea). Genetica 123, 245–253.
Sajdak, S.L., Reed, K.M., Phillips, R.B., 1998. Intra individual and inter-species variation in the 5S r DNA of coregonid fish. J. Mol. Evol. 46, 680–688.
Sambrook, J., Russell, I., 2001. Molecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, Plainsveiw New York.
Santos, L.V.D.R., Foresti, F., Wasko, A.P., Oliveira, C., Martins, C., 2006. Nucleotide sequence, genomic organization and chromosome localization of 5S rDNA in two species of Curimatidae (Teleostei, Characiformes). Genet. Mol. Biol. 29 (2), 251–256.
Sarkar, U.K., Srivastava, S.M., 2000. Endangered Mahseer in North East India. In: Ponniah, A.G., Sarkar, U.K. (Eds.), Fish Biodiversity of North East India. NATP Publ, NBFGR, pp. 95–96.
Shrestha, T.K., 1990. Rare fishes of Himalayan waters of Nepal. J. Fish Biol. 37, 213–216.
Suzuki, H., Sakurai, S., Matsuda, Y., 1996. Rat 5S rDNA spacer sequences and chromosomal assignment of the genes to the extreme terminal region of chromosome 19. Cytogenet. Cell Genet. 72, 1–4.
Swofford, D.L., 2003. PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.
Tang, J., Toe, L., Back, C., Unnasch, T.R., 1996. Intra-specific heterogeneity of their DNA internal transcribed spacer in the *Simulium damnosum* (Diptera: Simuliidae) complex. Mol. Biol. Evol. 13, 244–252.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTALW: improving the sensitivity of progressive multiple alignment through sequence weighting, position specific-gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Tigano, C., Rocco, L., Ferrito, V., Costagliola, D., Pappalardo, A.M., Stinso, V., 2004. Chromosomal mapping and molecular characterization of ribosomal RNA genes in *Lebias fasciata* (Teleostei, Cyprinodontidae). Genetica 121, 95–100.
Torres, R.A., Ganal, M., Hemleben, V., 1990. GC balance in the internal transcribed spacers ITS1 and ITS2 of nuclear ribosomal RNA genes. J. Mol. Evol. 30, 170–181.
Vawter, L., Brown, W.M., 1993. Rates and patterns of base change in the small subunit ribosomal RNA gene. Genetics 134, 597–608.
Verma, J., Serajuddin, M., 2012. Phylogenetic relationship between four species using divergent domain d8 in family: siluridae. Int. J. Life Sci. Pharm. Res. 2 (3), 240–244.
Wasko, A.P., Martins, C., Wright, J.M., Galetti, P.M., 2001. Molecular organization of 5S rDNA in fishes of genus *Brycon*. Genome 44, 893–902.
Zardoya, R., Meyer, A., 1996. Evolutionary relationship of the coelacanth, lungfish and tetrapods based on 28S ribosomal RNA gene. Proc. Natl. Acad. Sci. 93, 5449–5454.
Zhou, W., Cui, G.H., 1996. A review of *Tor* species from the Lancangjiang river (Upper Mekong river), China (Teleostei: Cyprinidae). Ichthyol. Explor. Freshw. 7, 131–142.