Int. J. Aquat. Biol. (2021) 9(1): 1-10
ISSN: 2322-5270; P-ISSN: 2383-0956
Journal homepage: www.ij-aquaticbiology.com

Original Article

Genetic diversity and population structure of *Barilius barna* (Hamilton, 1822) in the sub-Himalayan Dooars region of West Bengal, India through Mitochondrial Cytochrome Oxidase I Sequence analyses

Ajoy Paul[1,2], Tanmay Mukhopadhyay[1,3], Soumen Bhattacharjee[*1]

[1]Cell and Molecular Biology Laboratory, Department of Zoology, University of North Bengal, P.O. North Bengal University, Raja Rammohunpur, Siliguri, India.
[2]Protein Engineering and Neurobiology Laboratory, Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Powai, Mumbai, India.
[3]Department of Zoology, North Bengal St. Xavier's College, Rajganj, India.

Abstract: The genetic diversity and the population structure of *Barilius barna* (Hamilton, 1822) wild population from the Teesta River were assessed through mtDNA cytochrome oxidase I (COI) sequence analyses. The haplotype and nucleotide diversity analyses revealed low level of genetic diversity in the *B. barna* wild populations, especially in the lower reaches of Teesta (Bholarhat). The genetic differentiation and gene flow between the two study sites were 0.08434 and 2.71, respectively. Tajima's *D*, Fu and Li's *D* and Fu and Li's *F* analyses were used to assess population differentiation in the two study sites. Haplotype networking and phylogenetic analyses clearly distinguished the two populations from each other, as well as from other populations from other parts of the country. Nature and implications of the genetic and haplotype diversities among the populations are discussed. Phylogenetic analyses also indicated that the Gajoldoba population is genetically closer to north Indian river populations, than that to Bholarhat population.

Article history:
Received 8 June 2020
Accepted 22 January 2021
Available online 25 February 2021

Keywords:
Teesta
COI
Haplotype
Phylogeny

Introduction

Barilius barna (Hamilton, 1822) (Cyprinidae) is an economically important tropical fresh water fish found in the Teesta River and adjoining rivers of North Bengal, India. *Barilius barna* has a "NE" or "not evaluated" status according to the IUCN but has noteworthy ornamental and market values. The conservation status of this species is considered as "lower risk near threatened" (LRnt) according to CAMP (Conservation Assessment and Management Plan) report for freshwater fishes of India (Molur and Walker, 1998). However, during the past few years, its population has dwindled substantially and the fish has become increasingly rare in rivers of North Bengal, India. The dwindling population of *B. barna* may be ascribed to over-exploitation, dam/hydroelectric power plant constructions, urban effluents and/or pesticide run-offs from the nearby tea gardens. The loss of genetic diversity of any economically valuable species in the hotspot region is a great threat to biodiversity locally, as well as globally.

The study region is situated at the north-eastern sub-Himalayan region of India which is known for its proximity to Himalayan biodiversity hotspot. Previously we have studied the genetic architecture of *B. barna* by RAPD and ISSR markers (Paul et al., 2018) and have found that the genetic diversity of this fish has dwindled. Owing to the inherent nature of those dominant and arbitrary markers, we set out to ascertain the present status of genetic diversity through mitochondrial DNA-based assessment of the local wild populations of *B. barna* in the Teesta River of North Bengal, India.

Application of molecular data in conservation of endangered species has been practiced in various organism groups at the national and international level for a long time. Mitochondrial DNA (mtDNA) has been widely adopted for population genetic studies and it being almost exclusively maternally inherited has some advantages over the nuclear DNA

[*]Correspondence: Soumen Bhattacharjee
E-mail: soumenb123@rediffmail.com
DOI: https://doi.org/10.22034/ijab.v9i1.901

(Billington, 2003). Mitochondrial DNA is highly variable in natural populations because of its elevated mutation rate, which can generate some information about population history over short time frames. Being involved in basic metabolic functions (cellular respiration), mitochondrial genes have been considered as less likely than other genes to be involved in adaptive processes (Galtier et al., 2009). One consequence of maternal transmission is that the effective population size for mtDNA is smaller than that of nuclear DNA, therefore, mtDNA variation is a more sensitive indicator of population phenomena (Avise, 1994). Among the most frequently used mitochondrial genes for detecting genetic data, the one coding for cytochrome oxidase subunit I (COI) is amplified using polymerase chain reaction methods using conserved primers (Folmer et al., 1994). In this regard, the primary goal of the present study was to ascertain the genetic diversity available at mitochondrial COI and also to investigate the phylogenetic relationships of the isolates with other *B. barna* populations, based on publicly available sequences.

Materials and Methods

A detailed survey has been carried out in two spots (at two different altitudes) of the Teesta River of the sub-Himalayan, Dooars Region of West Bengal, India. *Barilius barna* were collected from the river during October, 2015 and March, 2017. Geographic coordinates were recorded using handheld GPS (eTrex Vista HCx, Garmin, USA). The fishes were carried to the laboratory immediately after collection in ice bucket and identified (Shaw and Shebbeare, 1934; Talwar and Jhingran, 1991). The collection spots were as follows: GBb (Gajoldoba *B. barna*) and BBb (Bholarhat *B. barna*). The location and geographical co-ordinates of the collection spots are mentioned in Figure 1.

Isolation of genomic DNA and quantification: Genomic DNA (gDNA) was isolated from small amount of tissue samples (10-15 mg of fin clips) using commercial DNA isolation Kit (DNeasy Blood and Tissue Kit, Qiagen). The isolated gDNA samples were

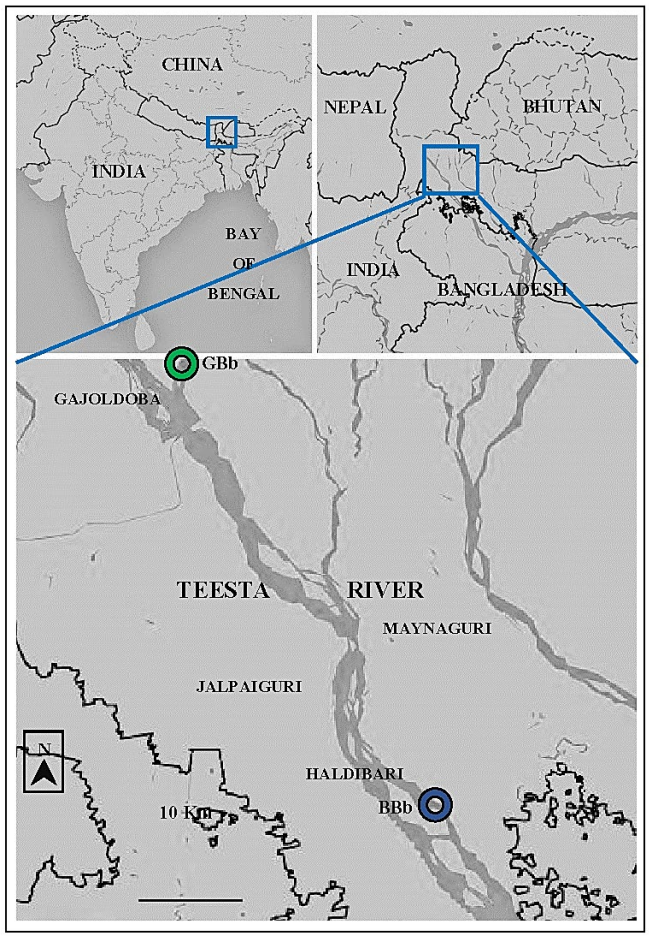


Figure 1. Map showing two collection spots of *Barilius barna* from the Teesta river of Dooars region of West Bengal, India (GBb= Gajoldoba (N 26°44´584, E 88°35´314 Elev 354 AMSL) and BBb= Bholarhat (N 26°21´423, E 88°50´485 Elev 193 AMSL).

stored in 1.5 ml microcentrifuge tubes at -20ºC. The gDNA samples were quantified using spectrophotometer (Rayleigh UV-2601 Spectrophotometer, Beijing, China). The concentration of the extracted gDNA was adjusted to 50 ng/µl for subsequent PCR amplifications. The used primer sequences are follows: F1: 5′- TCAACCAAC CACAAAGACAT TGGCAC-3′ (GC Content 42.30%); R1: 5′- TAGA C TTCTGGGTGGCCAAA GAATCA-3′ (GC Content 46.15%) (Ward et al., 2005). They were synthesized by Imperial Life Science (Gurgaon, India).

PCR amplification and documentation of amplified products: Mitochondrial COI gene amplification was performed in a 96 wells Eppendorf® (Germany) Peltier thermal cycler. The final reaction volumes were 30 µl, each containing a final concentration of ~100-150 ng of isolated gDNA, 1 pM

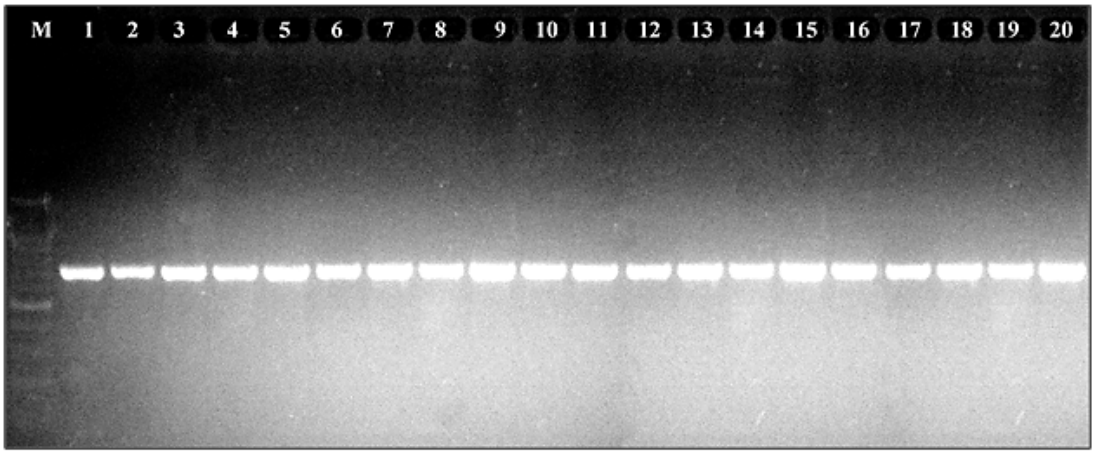


Figure 2. Representative gel photograph of the mtDNA CO1 PCR amplicons. M=100 bp DNA size ladder, Lane no. 1-10= individuals from GBb and Lane no. 11-20= individuals from BBb.

of each oligonucleotide primers (Forward and Reverse), 1X standard Taq Polymerase buffer (10 mM Tris-HCl, pH 8.3, 50 mMKCl, 1.5 mM $MgCl_2$) (NEB, USA), 200 µM of each dNTPs (dATP, dTTP, dCTP, dGTP) (NEB, USA), and one unit of Taq DNA Polymerase (NEB, USA). PCR cycling programs were as follows: initial denaturation at 94ºC for 5 min followed by 40 cycles of 94ºC, 1 min for denaturation; 50ºC, 45 sec for annealing; 72ºC, 1 min for elongation and finally an extension at 72ºC for 10 min. The amplified products were electrophoresed in an ethidium bromide (0.5 µg/ml) pre-stained 1.5% (w/v) agarose gel (Lonza, Switzerland) at a constant voltage 100 V and current 100 mA in TAE buffer (40 mM Tris-HCl, pH 8.0; 20 mM acetic acid; 1 mM EDTA, pH 8.0) using BenchTop Labsystems BT-MS-300 (Taiwan) electrophoretic apparatus. Molecular weight of each band was estimated using a standard 100 base pair ladder (NEB, USA). The gels were visualized on the UV-transilluminator (Spectroline BI-O-Vision®NY, USA) and photographed using a Nikon D3100 camera (Fig. 2).

Purification of PCR product and sequencing of COI: The amplified products were purified using Invitrogen Purelink™ PCR Purification Kit (Thermofisher Scientific, India), following manufacturer's protocol. The amplified products were confirmed as COI partial coding gene by restriction digestion using Hind III restriction enzyme (Hind III cutting point between 252^{th}-253^{th} bp). Each PCR product was sequenced at least twice by dye-dideoxy automated chain termination method (MWG-Biotech Pvt. Ltd., Bangalore India). All the COI partial CDS were used separately to search the GB/EMBL/DDBJ combined nr database in NCBI BLAST search through the implementation of Blastn and Blastx, optimized for highly similar sequences (Megablast), using stringent expect threshold (0.01) and excluding low complexity regions within the combined database. All the sequences identified *B. barna* and *B. bendelisis* had high scores and low expect values (data not shown). The curated sequences (approximately 606 bp) were submitted to GenBank and accession numbers were obtained (Table 1).

Genetic diversity and population structure: Levels of genetic variation within the Gajoldoba and Bholarhat populations were estimated in terms of the number of variable sites (S), total number of mutations (Eta), number of haplotypes (H), haplotype diversity (H), nucleotide diversity (π_n) and average number of nucleotide differences by DnaSP ver. 5.1 software (Librado and Rozas, 2009). The genetic differentiation (F_{ST}) and gene flow (N_m) were calculated following the method of Wright (1978), Govindaraju (1989) and Low et al. (2014). Tajima's *D* (Tajima, 1989), Fu and Li's *D* and Fu and Li's *F* (Fu, 1997) were calculated to verify selective neutrality in relation to mtDNA sequences, which would help to accumulate information regarding the demographic history of the population and used to deduce whether a population has undergone a sudden population expansion or construction.

Table 1. List of species used for molecular analysis and their GenBank accession number and deposited as result of the present study.

Accession Number	Species	Reference for Sequences
KX649920	*Barilius barna*	Paul et al. 2016
KX649921	*Barilius barna*	Paul et al. 2016
KX649922	*Barilius barna*	Paul et al. 2016
KX649923	*Barilius barna*	Paul et al. 2016
KX649924	*Barilius barna*	Paul et al. 2016
KX649925	*Barilius barna*	Paul et al. 2016
KX649926	*Barilius barna*	Paul et al. 2016
KX649927	*Barilius barna*	Paul et al. 2016
KX649928	*Barilius barna*	Paul et al. 2016
KX649929	*Barilius barna*	Paul et al. 2016
KX649930	*Barilius barna*	Mukhopadhyay et al. 2016
KX649931	*Barilius barna*	Mukhopadhyay et al. 2016
KX649932	*Barilius barna*	Mukhopadhyay et al. 2016
KX649933	*Barilius barna*	Mukhopadhyay et al. 2016
KX649934	*Barilius barna*	Mukhopadhyay et al. 2016
KX649935	*Barilius barna*	Mukhopadhyay et al. 2016
KX649936	*Barilius barna*	Mukhopadhyay et al. 2016
KX649937	*Barilius barna*	Mukhopadhyay et al. 2016
KX649938	*Barilius barna*	Mukhopadhyay et al. 2016
KX649939	*Barilius barna*	Mukhopadhyay et al. 2016
HM042158	*Barilius barna*	Mishra et al. 2010
HM042159	*Barilius barna*	Mishra et al. 2010
HM042160	*Barilius barna*	Mishra et al. 2010
HM042161	*Barilius barna*	Mishra et al. 2010
HM042162	*Barilius barna*	Mishra et al. 2010
HM042163	*Barilius barna*	Mishra et al. 2010
HM042170	*Barilius barna*	Mishra et al. 2010
HM042171	*Barilius barna*	Mishra et al. 2010
HM042172	*Barilius barna*	Mishra et al. 2010
HM042173	*Barilius barna*	Mishra et al. 2010
HM042174	*Barilius barna*	Mishra et al. 2010
HM042175	*Barilius barna*	Mishra et al. 2010
HM042164	*Barilius barna*	Mishra et al. 2010
HM042165	*Barilius barna*	Mishra et al. 2010
HM042166	*Barilius barna*	Mishra et al. 2010
HM042167	*Barilius barna*	Mishra et al. 2010
HM042168	*Barilius barna*	Mishra et al. 2010
HM042169	*Barilius barna*	Mishra et al. 2010
HM042176	*Barilius barna*	Mishra et al. 2010
HM042177	*Barilius barna*	Mishra et al. 2010
HM042178	*Barilius barna*	Mishra et al. 2010
HM042179	*Barilius barna*	Mishra et al. 2010
HM042180	*Barilius barna*	Mishra et al. 2010
HM042181	*Barilius barna*	Mishra et al. 2010
EU417797	*Barilius barna*	Lakra et al. 2008
EU417798	*Barilius barna*	Lakra et al. 2008
EU417799	*Barilius barna*	Lakra et al. 2008
JN965190	*Barilius barna*	Kalyankar et al. 2011

Preparation of sequence dataset and phylogenetic analyses: Twenty raw sequence reads, each PCR product sequenced at least twice, were curated in BioEdit ver. 7.0.9 (Hall, 1999) for character uncertainty and then assembled in Geneious R7 ver 7.0.6 (trial) (Rozen and Skaletsky, 2000) using pro features and retrieved the final sequences as Fasta files. Multiple sequence alignment was implemented in CLUSTALX ver 2.0.3 (Thompson et al., 1997) using high gap opening (50) and gap extension (50) penalties. The best model of DNA substitution selected was GTR or General Time Reversal (Tavare, 1986) plus I+Γ based on the Akaike Information Criterion (AIC) implemented in jModeltest v0.1.1 (Guindon and Gascuel, 2003; Posada, 2008). Publicly available twenty-eight *B. barna* COI partial CDS corresponding to isolates submitted from northern India were retrieved from GenBank before August 21, 2017. The isolate names and their accession numbers are presented in Table 1. The final dataset contained 655 positions which included twenty-eight 606 to 655 nucleotides long GenBank sequences and twenty 605 to 606 nucleotide long *B. barna* COI sequences from the Teesta River.

Phylogenetic analyses: The evolutionary relationship between forty-eight *B. barna* mitochondrial COI sequences was estimated using Bayesian coupled with Markov Chain Monte Carlo (BMCMC) methods of phylogenetic inference and Maximum Likelihood (ML). BMCMC searches (Huelsenbeck, 2001) were performed in MrBayes v3.1 (Huelsenbeck, 2001; Ronquist, 2003) using the following model parameters: base frequencies f(A), 0.2220; f(C), 0.2506; f(G), 0.2164; f(T), 0.3110; substitution rates r[CT], 3.3482; r[CG], 0.9238; r[AT], 2.1907; r[AG], 4.7431; r[AC], 1.1675; proportion of invariant sites pinv, 0.0780; and shape parameter of the gamma distribution α, 1.3620. Four Markov chains ($4x2x10^6$ cycles; ngen = 1000000) were run simultaneously, which were started from random trees and sampled every 1,000th cycle (samplefreq = 1000). TRACER v1.2 (Rambaut, 2003) was used to check whether stationarity in the fluctuating values of the likelihood and all the phylogenetic parameters had been reached. Each simulation was repeated four times (nchain = 4) starting from different random trees (startingtree = random). All sample points prior to reaching stationarity were discarded as *burn in* (burninfrac = 0.25). The posterior probabilities for individual clades obtained from separate analyses were compared for congruence and then combined and summarized in a

Table 2. Population genetic diversity parameters based on *Barilius barna* MtDNA COI partial coding sequences.

Popu lations	No. of variable sites (S)	Total no. of mutatio ns (Eta)	No. of Haplotype s (h)	Haplotype (gene) diversity (Hd±SD)	Nucleotide diversity (Pi±SD)	Average number of nucleotide differences (k)	Genetic differentiation, (F_{ST}) between GBb and BBb population	Gene flow (N_m) between GBb and BBb population
Bholarhat (BBb)	106	107	5	0.756 ±0.01678	0.06329 ±0.0000684	38.28889		
Gajoldoba (GBb)	102	102	8	0.933 ±0.00597	0.03607 ±0.0006035	36.05555	0.08434	2.71
Whole Population	108	109	13	0.926 ±0.00185	0.08996 ±0.0000487	54.42632		

*COI, Cytochrome oxidase subunit I; SD, standard deviation

majority rule consensus tree.

ML searches (Felsenstein, 1981) of sequence data set were performed in Mega v6 (Tamura et al., 2013) based on software suggested model (K2+G) parameters: equal base frequencies; substitution rates r(AT), 0.032; r(AC), 0.032; r(AG), 0.187; r(TA), 0.032; r(TC), 0.187; r(TG), 0.032; r(CA), 0.032; r(CT), 0.187; r(CG), 0.032; r(GA), 0.187; r(GT), 0.032; r(GC), 0.032; and shape parameter of the gamma distribution α, 0.3845. ML heuristic search was implemented using 100 bootstrapping option by a slow but accurate Subtree Pruning and Regrafting (SPR Level 5) method with starting Bionj trees. All the trees were built as midpoint-rooted consensus trees using Figtree v1.3.1 (http://tree.bio.ed.ac.uk /software/figtree/).

Haplotype network analysis: Intraspecific gene genealogies were inferred using haplotype network construction method, implemented in freely available software package Network ver 5 (www.fluxus-engineering.com/sharenet.htm). The algorithm for constructing the minimum spanning trees (MSTs) from a matrix of pairwise distances (absolute number of differences) among haplotypes (Prim, 1957; Rohlf, 1973) has been modified to include all possible MSTs within a single graph as the minimum spanning network (MSN) (Excoffier and Smouse, 1994). All 48 taxa were aligned by CLUSTALW ver. 2 (Thompson et al., 1994) and were concatenated to yield a total length of 605 bp. In the median-joining network approach (Bandelt et al., 1999), all MSTs are first combined within a single network (MSN) following an algorithm analogous to that proposed by Excoffier and Smouse (1994). Then, using the parsimony criterion, inferred intermediate haplotypes were added to the network to reduce overall tree length. Using the parsimony criterion, inferred intermediate haplotypes were added to the network to reduce overall tree length. In addition, by setting a parameter nucleotide character weight = 10 and epsilon (ε) value = 0, less parsimonious pathways were excluded in the inferred network (Bandelt et al., 1999).

Results

Diversity and population structure: We have found a total number of variable sites were 106 and 102 in Bholarhat (BBb) and Gajoldoba (GBb), respectively, and 108 when two populations were considered together (Table 2). Total numbers of mutations were 107 and 102 in BBb and GBb populations, respectively (Table 2). Total 13 haplotypes were found in the Teesta river population, where 5 were for BBb population and 8 for GBb (Table 2). The haplotype diversity and nucleotide diversity of BBb population were 0.756±0.01678SD and 0.06329 ±0.0000684SD, respectively; and of GBb population were 0.933±0.00597SD and 0.03607±0.0006035SD, respectively (Table 2). The genetic differentiation (F_{ST}) and gene flow (N_m) between GBb and BBb populations were 0.08434 and 2.71, respectively (Table 2). The observed values of Tajima's D, Fu and Li's D and Fu and Li's F analyses of Bholarhat population were 0.10843 (Not significant; P>0.10), 1.61711 (significant; P<0.02) and 1.40063 (Not significant; P>0.10), respectively; and in Gajoldoba population were 1.95643 (significant; P<0.05), -2.35348 (significant; P<0.02) and -2.54752 (significant; P<0.02), respectively.

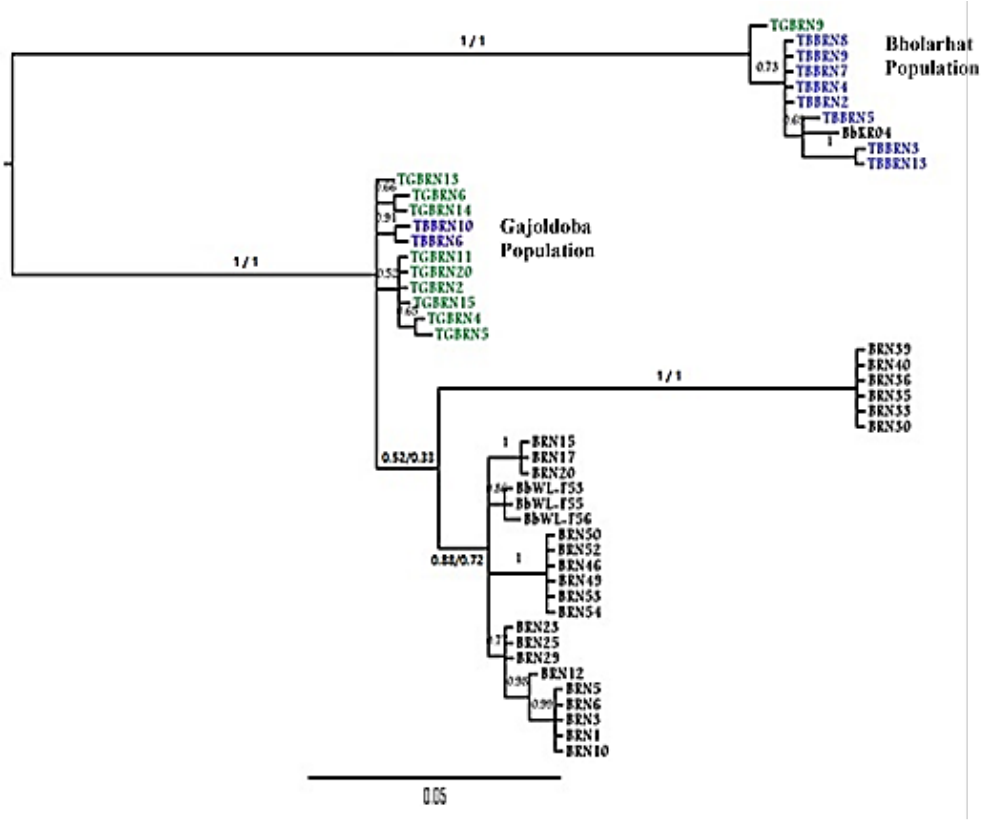
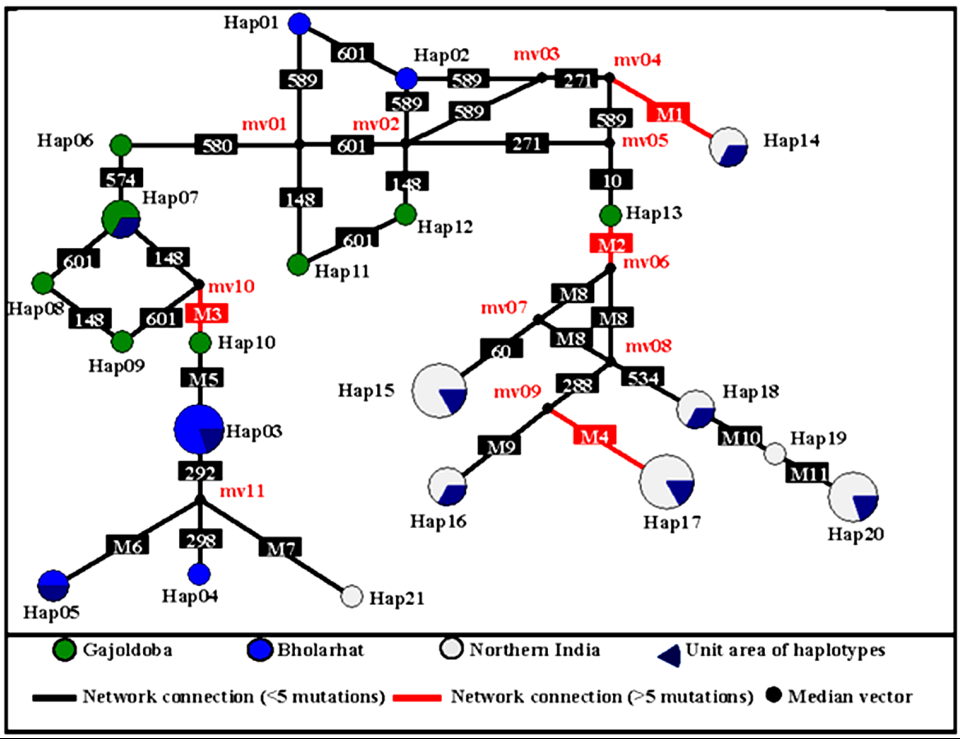


Figure 3. BMCMC and ML phylogenetic analyses showing the evolutionary relationship of forty-eight *Barilius barna* mitochondrial COI inferred by using the GTR and Kimura 2 models, respectively. The posterior probability (BMCMC) and bootstrap values (ML) in which the associated taxa clustered together are shown next to the major branches (Gajoldoba isolates are in green and Bholarhat isolates are in blue. Posterior probabilities (pP) (1.0) and bootstrap (100% represented as 1.0) values are represented as pP/Bootstrap).

BMCMC and ML analyses: The dichotomy of the two main lineages was strongly supported by high posterior probabilities (p*P*) (1.0) and bootstrap (100% represented as 1.0) in the BMCMC and ML midpoint rooted phylogenetic trees, respectively (Fig. 3). The phylogenetic trees suggested that eight TBBRN (Bholarhat population), one TGBRN (Gajoldoba population) and one north Indian isolate (KR04) are closely related to each other than to other *B. barna* COI sequences. While nine TGBRN (Gajoldoba population) and two TBBRN (Bholarhat population) formed a highly supported sister clade along with all the other north Indian *B. barna* sequences in BMCMC analysis (Fig. 3). Almost identical topology and phylogenetic relationships were produced in the ML analysis as well as found in BMCMC analyses implemented in Mega ver. 6 (Fig. 3).

Network analysis of haplotypes: Total 48 taxa were used to construct the haplotype network (Fig. 4). Haplotype network of *B. barna* COI were in agreement with the phylogenetic reconstruction (Fig. 3). The haplotype network showed that the segregation of mitochondrial haplotypes was in accordance with the locality (Figs. 3, 4). A total of 21 haplotypes were identified in the *B. barna*, distributed in three main geographic groups. Among them five (Hap 01-05) haplotypes were distributed in Bholarhat and eight (Hap06-13) were distributed in Gajoldoba (Fig. 4). Remaining eight (Hap14-21) haplotypes were distributed in the northern region of India. A unique haplotype (Hap21 i.e., KR04) connects with the Hap 04 and Hap 05 haplotypes in Bholarhat population. There was no sharing of haplotypes within these three groups. The network shows variations in nucleotide character positions 601 (Hap7-Hap8, Hap9-mv10 and Hap1-Hap2, Hap11-Hap12, mv1-mv2), 148 (Hap7-mv10, Hap8-Hap9, mv1-Hap11, mv2-Hap12), 589 (Hap1-mv1, Hap2-mv2, mv2-mv3, mv4-mv5) and 271 (mv3-mv 4, mv2-mv5) which indicate parallel mutations or homoplasy (Fig. 4).



Number of Haplotypes	Name of Haplotypes	Taxon names	Source
1	Hap 01	TBBRN10	(Teesta River) Bholarhat
2	Hap 02	TBBRN6	
3	Hap 03	TBBRN2, TBBRN4, TBBRN7, TBBRN8, TBBRN9	
4	Hap 04	TBBRN5	
5	Hap 05	TBBRN3, TBBRN13	
6	Hap 06	TGBRN15	(Teesta River) Gajoldoba
7	Hap 07	TGBRN2, TGBRN11, TGBRN20	
8	Hap 08	TGBRN4	
9	Hap 09	TGBRN5	
10	Hap 10	TGBRN9	
11	Hap 11	TGBRN6	
12	Hap 12	TGBRN14	
13	Hap 13	TGBRN13	
14	Hap 14	WL-F53, WL-F55, WL-F56	Northern India
15	Hap 15	BRN46, BRN49, BRN50, BRN52, BRN53, BRN54	
16	Hap 16	BRN15, BRN17, BRN20	
17	Hap 17	BRN30, BRN33, BRN35, BRN36, BRN39, BRN40	
18	Hap 18	BRN23, BRN25, BRN29	
19	Hap19	BRN12	
20	Hap 20	BRN1, BRN3, BRN5, BRN6, BRN10	
21	Hap 21	KR04	

Figure 4. Median-joining network for the COI haplotypes in *Barilius barna* populations from Teesta river of North Bengal, India and northern region of India. Each circle represents a unique haplotype, with the area being proportional to the frequency of the haplotypes. The network showed the mutational changes. The black and red boxes indicate the nucleotide position of mutational changes. M1-4=contain >5 number of mutations, M5=583,586,589; M6=37, 220, 268, 313, 454; M7=34, 35, 133; M8=214, 336; M9=474, 558; M10=306, 474; M11=72, 132.

Discussions

The results revealed low level of genetic diversity in the dwindling wild population of *B. barna* from two different locations of the Teesta River of sub-Himalayan northern West Bengal, India. The mitochondrial DNA COI gene sequences are highly polymorphic in nature, as reported in many genetic studies on geographically isolated populations of different organisms. Similar study had carried out on snakehead fish from Thailand, a total 33 haplotypes among 60 individuals was found and; haplotype and nucleotide diversity ranged from 0.75-1.0 and 0.00442-0.2672, respectively (Boonkusol et al., 2016). In the present study, we found that the total numbers of haplotypes were 13 for 20 individuals of *B. barna* distributed in two different populations of Teesta River of North Bengal, India. These reveal that the genetic diversity (number of haplotypes, haplotype diversity and nucleotide diversity) is low in *B. barna* population as evident from the present mitochondrial COI study. The present findings are also in accordance with our previous study on *B. barna* of the same river, where we have found that the overall genetic diversity (polymorphism, Nei's genetic diversity and Shannon's information index) was low as revealed by RAPD and ISSR marker (Paul et al., 2018).

The study carried out by Low et al. (2014) categorized the levels of genetic differentiation as F_{ST}>0.25 (great differentiation), 0.15 to 0.25 (moderate differentiation), and F_{ST}<0.05 (negligible differentiation). The levels of gene flow can be categorized as Nm>1 (high gene flow), 0.25 to 0.99 (intermediate gene flow), and Nm<0.25 (low gene flow). The genetic differentiation (F_{ST}) observed in our study was 0.08434 which is comparatively low than the study carried by Boonkusol et al. (2016) on snakehead fish from Thailand (F_{ST} ranged from 0.27891 to 0.40627). The low level of genetic divergence between the study regions may be attributed to the migratory behaviour and gene flow (Nm=2.71), as evident in our present study. These findings were also supported by our previous study on *B. barna* where we have detected sufficient level gene flow between Gajoldoba and Bholarhat population (Paul et al., 2018). Anthropogenic could be the possible reasons behind the dwindling population structure of this species in the study region.

The Bholarhat population gave positive values for Tajima's *D*, Fu and Li's *D* and Fu and Li's *F* test, which indicated that the population showed balancing selection and sudden population contraction; whereas, the Gajoldoba population gave negative values for all the test, which indicated a selective sweep and a population expansion after a recent bottleneck (Tajima, 1989; Fu and Li, 1993). The study of Thirumaraiselvi and Thangaraj (2015) on six *Eleutheronema tetradactylum* populations from south Asian countries also found negative values for the tests which indicated sudden population expansion of the *E. tetradactylum* population.

In our study, we have found a total 13 haplotypes from twenty different individuals of *B. barna* from Teesta River of North Bengal India and 9 haplotypes from 28 individuals from Northern India. The haplotypes network showed possible ancestral connections within the network along with differentiation of haplotypes with respect to its connecting haplotypes. Our study also revealed that some of the haplotypes showed shared parallel mutation at specific nucleotides or characters. This homoplasic condition decreases the genetic diversity within the population (Wake, 1991; Wake et al., 2011) and this observation is also in accordance with the haplotype diversity in our study.

Our mitochondrial cytochrome oxidase I based BMCMC and ML trees showed the phylogenetic relatedness between the two populations of the Teesta river system of the Dooars region of north-eastern West Bengal, India. The clubbing of TGBRN9 (Gajodoba) with the Bholarhat population (blue taxa) and that of TBBRN6 and TBBRN10 (Bholarhat) with the Gajoldoba population (green taxa) also indicate a significant level of gene flow between the sites. With the unavailability of any data from the region we resorted to comparing our data with that of other *B. barna* COI sequences from the Uttar Pradesh state of northern India. Our result indicated that the Gajoldoba isolates are genetically closer, than that of

the Bholarhat isolates, to the other published sequences.

Acknowledgement

The work was supported by a University Research Grant (2015-16) awarded to the corresponding author.

References

Avise J.C. (1994). Molecular markers, natural history and evolution. Chapman and Hall. New York. 511 p.

Bandelt H.J., Forster P., Rohl A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution, 16: 37-48.

Billington N. (2003). Mitochondrial DNA. In: E.M. Hallerman (Ed.) Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, MD. pp: 59-100.

Boonkusol D., Tongbai W. (2016). Genetic variation of striped snakehead fish (*Channa striata*) in river basin of central Thailand inferred from mtDNA COI gene sequence analysis. Journal of Biological Sciences, 16: 37-43.

Du X., Chen Z., Deng Y., Wang Q. (2009). Comparative analysis of genetic diversity and population structure of *Sipunculus nudus* as revealed by mitochondrial COI sequences. Biochemical Genetics, 47: 884-891.

Excoffier L., Smouse P.E. (1994). Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. Genetics, 136: 343-359.

Felsenstein J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution, 17: 368-376.

Folmer O., Black M., Hoeh W., Lutz R., Vrijenhoek R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology, 3: 294-299.

Fu Y.X., Li W.H. (1993). Statistical tests of neutrality of mutations. Genetics, 133: 693-709.

Galtier N., Nabholz B., Glemin M.S., Hurst G.D.D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Molecular Ecology, 18: 4541-4550.

Govindaraju D.R. (1989). Variation in gene flow level among predominantly self-pollinated plants. Journal of Evolutionary Biology, 2: 173-181.

Guindon S., Gascuel O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology, 52: 696-704.

Hall T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium, Series 41: 95-98.

Huelsenbeck J.P., Ronquist F., Nielsen R., Bollback J.P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. Science, 294: 2310-2314.

Kimura M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution, 16: 111-120.

Librado P., Rozas J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25: 1451-1452.

Low V.N., Adler P.H., Takaoka H., Yacob Z., Lim P.E., Tan T.K., Lim Y.A.L., Chen C.D., Rashid Y.N., Azirun M.S. (2014). Mitochondrial DNA markers reveal high genetic diversity but low genetic differentiation in the Black fly *Simulium tani* Takaoka and Davies along an elevational gradient in Malaysia. PLOS One, 9: e100152.

Molur S., Walker S. (1998). Conservation Assessment and Management Plan for freshwater fishes of India. Zoo Outreach Organization, Conservation Breeding Specialist Group. Coimbatore, India. 5 p.

Paul A., Mukhopadhyay T., Bhattacharjee S. (2018). Genetic characterization of *Barilius barna* (Hamilton, 1822) in the Teesta river of sub-Himalayan West Bengal, India, through RAPD and ISSR fingerprinting. Proceedings of the Zoological Society, 7: 203-212.

Posada D. (2008). jModel Test: phylogenetic model averaging. Molecular Biology and Evolution, 25: 1253-1256.

Prim R.C. (1957). Shortest connection networks and some generalizations. Bell System Technical Journal, 36: 1389-1401.

Rambaut A., Drummond A.J. (2003). Tracer: MCMC trace analysis tool, 2nd. University of Oxford, Oxford, UK. Available from: http: //evolve.zoo.ox.ac.uk. Retrieved 10/13/2017.

Rohlf F.J. (1973). Algorithm 76. Hierarchical clustering using the minimum spanning tree. The Computer Journal, 16: 93-95.

Ronquist F., Huelsenbeck J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19: 1572-1574.

Rozen S., Skaletsky H. (2000). Bioinformatics methods and protocols. In: S. Krawetz, S. Misener (Eds.). Methods in Molecular Biology. Humana Press, Totowa, NJ. 365 p.

Shaw G.E., Shebbeare E.O. (1937). The fishes of North Bengal. Journal of the Royal Asiatic Society of Bengal Science, 3: 1-137.

Tajima F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123: 585-95.

Talwar P.K., Jhingran A.G. (1991). Inland fishes of India and adjacent countries. Oxford and IBH Company Private Limited. New Delhi, India. 344 p.

Tamura K., Stecher G., Peterson D., Filipski A., Kumar S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution, 30: 2725-2729.

Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins G.G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research, 25: 4876-4882.

Thompson J.D., Higgins D.G., Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22: 4673-4680.

Wake D.B. (1991). Homoplasy: the result of natural selection, or evidence of design limitations? American Naturalist, 138 (3): 543-567.

Wake D.B., Wake M.H., Specht C.D. (2011). Homoplasy: from detecting pattern to determining process and mechanism of evolution. Science, 331: 1032–1035.

Ward R.D., Zemlak T.S., Innes B.H., Last P.R., Hebert P.D.N. (2005). DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B, 360: 1847-1857.

Wright S. (1978). Evolution and the genetics of populations, variability within and among natural populations. University of Chicago Press, Chicago, USA. 465 p.