Molecular Biology Reports (2020) 47:8579–8592
https://doi.org/10.1007/s11033-020-05901-w

ORIGINAL ARTICLE

Mitochondrial DNA markers reveal genetic connectivity among populations of Osteoglossiform fish *Chitala chitala*

Nimisha Dutta[1,2] · Rajeev Kumar Singh[1] · Abhinav Pathak[1] · Vindhya Mohindra[1] · Sangeeta Mandal[1] · Gurjeet Kaur[2] · Kuldeep Kumar Lal[1]

Received: 19 January 2020 / Accepted: 7 October 2020 / Published online: 20 October 2020

Abstract

Genetic diversity and population structure in Indian featherback fish, *Chitala chitala* (Hamilton, 1822) was investigated by combined analyses of two full mitochondrial genes, *ATPase 6/8* and *Cytochrome b*. A total of 403 individuals, collected from 14 rivers yielded 61 haplotypes. Hierarchical partitioning analysis identified 19.01% variance '*among*' and 80.99% variance '*within groups and populations*'. The mean coefficient of genetic differentiation (F_{ST}) was observed to be significant 0.26 ($p < 0.05$). Mantel tests rejected the hypothesis that genetic and geographic distances were correlated. The patterns of genetic differentiation, AMOVA and principal coordinate analyses indicated that natural populations were sub-structured and comprised of four genetic stocks of *C. chitala* in Indian rivers. The results also supported the higher resolution potential of concatenated gene sequences. The knowledge of genetic variation and divergence, from this study, can be utilized for its scientific conservation and management in the wild.

Keywords Mitochondrial genes · *Cytochrome b* · *ATPase 6/8* · Osteoglossiform · Genetic variability

Introduction

The Order Osteoglossiformes represents an ancestral teleost lineage, whose fossil records have been retrieved from the deposits, belonging to late Jurassic or early Cretaceous period [1]. The family Notopteridae comprises of ten species' belonging to four genera, distributed in the freshwaters of Africa and South Asia. The Indian featherback fish, *Chitala chitala* (Hamilton, 1822) [2], is widely distributed in Asian countries including India, Bangladesh, Sri Lanka, Pakistan, Nepal [3]. According to Taverne [4] the species exhibited wider distribution during prehistoric period, than today.

In India, *C. chitala* constitutes a substantial component of the inland fisheries. It has high demand as ornamental trade, food, sport and is considered a potential candidate species for aquaculture [5]. The increased sedimentation, environmental degradation, changes in river courses and overexploitation [6] have caused decline in its capture fisheries, due to which the species has been categorized as *'Near Threatened'* by the International Union for Conservation of Nature [7]. Several studies have been conducted in *C. chitala,* which include development of captive breeding protocols, larval rearing and age-growth relationship [8]. In an investigation, morphometric and meristics tools were employed to identify stocks of *C. chitala* in Indian rivers [9].

However, the efficient management requires the knowledge of natural genetic diversity and population structure [10]. Molecular data can provide critical inputs for determining genetic divergence and connectivity [11]. Mitochondrial DNA markers are proven tools to effectively characterize populations and capture demographic signatures due to their unique characteristics, such as high copy number, faster mutation rate and low effective population size [12]. Mitochondrial genes, particularly *ATPase 6/8* and *Cytochrome b (Cytb)* have been widely used to reveal intraspecific genetic variability and differentiation in fishes belonging to several

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11033-020-05901-w) contains supplementary material, which is available to authorized users.

✉ Rajeev Kumar Singh
rajeevsingh1@yahoo.com

[1] National Bureau of Fish Genetic Resources, Lucknow, India

[2] Amity Institute of Biotechnology, Amity University, Lucknow Campus, Lucknow, India

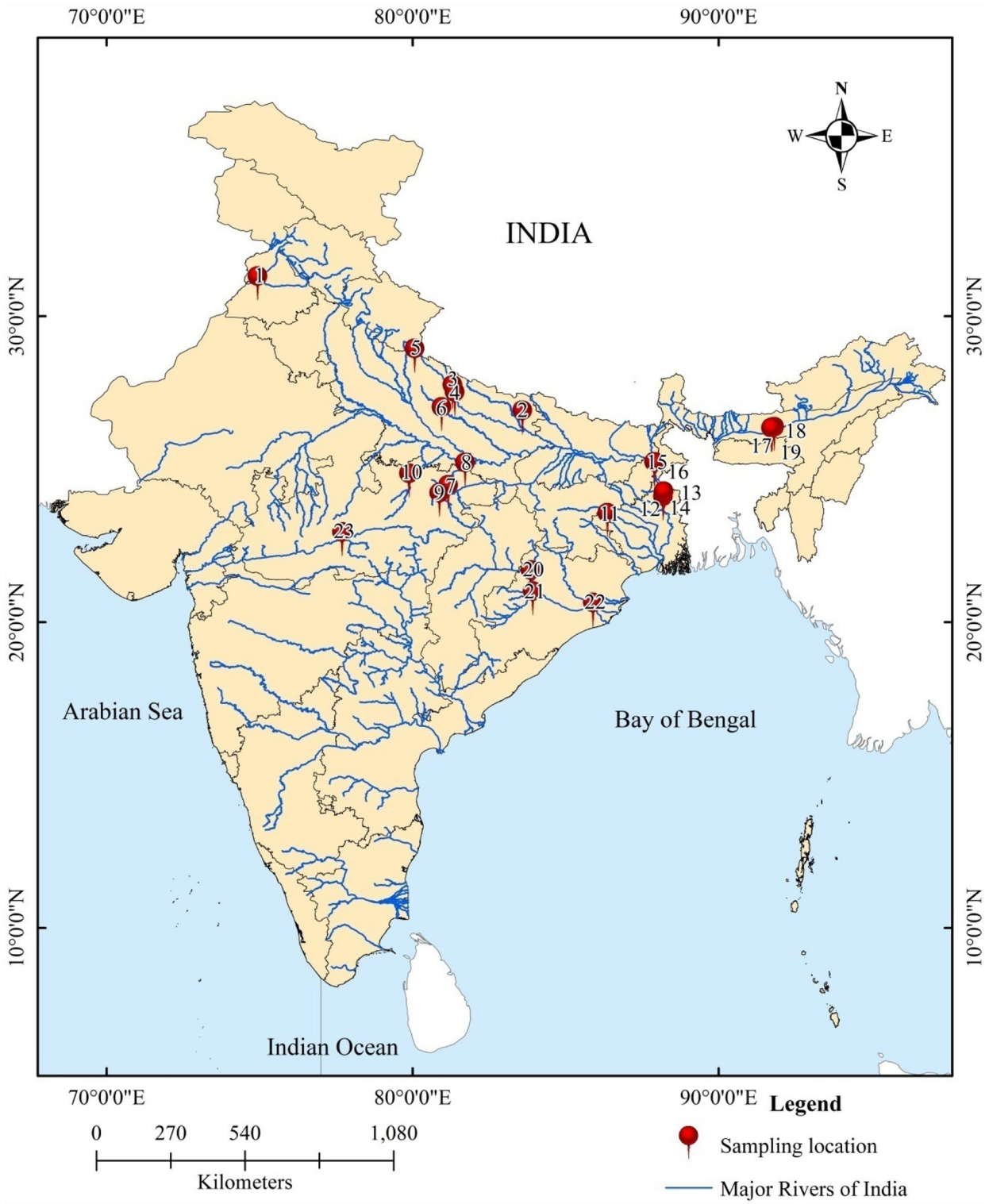

Fig. 1 Sampling locations of *C. chitala* from Indian rivers covering wide geographical distribution range. (**1.**Satluj River (Harike Pattan, Punjab); **2.**Gandak River (Valmiki Nagar, Bihar); **3.**Ghaghra River (Mehmoodabad, UP); **4.**Chauka River (UP); **5.**Sharda River (Sharda Barrage, Lakhimpur Kheri, UP) **6.**Gomti River (Lucknow, UP); **7.**Tons River (Rewa, MP); **8.**Tons River (Chakghat, UP); **9.**Son River (Bansagar, Beohari, Madhya Pradesh); **10.**Ken River (Rangua, MP); **11.**Hooghly River (Nabadeep, WB); **12.**Hooghly River (Beldanga, WB); **13.**Hooghly River (Satui, WB); **14.**HooghlyRiver (Behrampur, WB); **15.**Padma River (Farakka, WB); **16.**Padma River (Manikchak, WB); **17.**Brahmaputra River (Kolongpar, Assam); **18.**Brahmaputra River (Uzan Bazaar, Assam); **19.**Brahmaputra River (Guwahati, Assam); **20.**Mahanadi River (Hirakud Dam, Odisha); **21.**Mahanadi River (Sonepur, Odisha); **22.**Mahanadi River (Jobra Barrage, Odisha); **23.**Narmada River (Hoshangabad, MP)

taxonomic orders, such as Osteoglossiformes, Perciformes and Cypriniformes, etc. In *C. chitala,* previous studies had indicated the genetic variability, inferred from two different mitochondrial regions, partial *Cytb* and D-loop [5]. Molecular studies, based on *ATPase 6/8* genes, revealed genetic divergence in other notopterid *Notopterus notopterus* [13]. Till now, most genetic studies have used single and partial mitochondrial genes for population level analyses. However, recently, there has been a growing inclination towards using longer gene sequences that can provide sufficient resolution and robust computational support.

In the present study, we analyzed a concatenated data set of two full mitochondrial genes (*ATPase 6/8 and Cytb*) of *C. chitala* to investigate patterns of genetic variability, differentiation and population structure. The demographic history and genetic constraints were also assessed. The baseline data, generated from this study, will be a useful resource for the conservation and management of *Chitala* fisheries.

Materials and methods

The tissue samples of *C. chitala* individuals (n = 403) were collected from 14 rivers belonging to five river basins (Indus, Ganges, Brahmaputra, Mahanadi, and Narmada) (Fig. 1). Sampling strategy covered a wide natural range of distribution. The details of collection, GPS coordinates and year of sampling are provided in Table 1.

Genomic DNA extraction, amplification, and sequencing

The total genomic DNA was extracted from tissue/fin-clip/blood (ethanol preserved) using the modified phenol–chloroform method [14]. DNA quality was checked on 0.8% agarose gel stained with Nucleic Acid Safe Dye (G Biosciences) and visualized in gel documentation system (UVP Imaging System, Cambridge, UK). The purity and concentration was examined on DeNovix® DS-11 spectrophotometer. The DNA concentration was adjusted to 50 ng/µl and stored at 4^0C.

Two complete mitochondrial genes, *Cytb* (1139 bp) and *ATPase 6/8* (842 bp) were amplified using primer pairs *L14724*: 5′-CGA GAT CTG AAA AAC CAT CGT TG-3′; *H15915*: 5′-AAC TGC AGT CAT CTC CGG TTT ACA AGA A-3′ [15] and *ATP8.2L8331*: 5′-AAA GCR TYR GCC TTT TAA GC-3′; *CO3.2H9236*: 5′- GTT AGT GGT CAK GGG CTT GGR TC-3′ [16], respectively. Gene amplification was performed in a final reaction volume of 25 µl, which contained 1X reaction buffer (10 mM Tris, 50 mM KCl, 0.01% gelatin, pH 9.0), 200 µM of each dNTP, 1.5 mM $MgCl_2$, 5 pmol of each primer, 3U Taq polymerase (Genei, India) and approximately 50 ng of template DNA. PCR conditions included: denaturation at 94^0C for 5 min, 30 cycles of 94^0C for 30 s, 55^0C for 1 min and 72^0C for 1 min 30 s; final extension at 72^0C for 10 min. The amplicons were purified and sequenced bidirectionally on ABI automated sequencer (Applied Biosystems).

Genetic analysis

The sequences were edited and aligned using BioEdit sequence alignment editor version 7.2 [17]. We conducted a partition homogeneity test in PAUP* 4.0 [18] with heuristic search and 100 replicates (optimality criterion = parsimony) on all 1981 nucleotide characters. This was done to examine whether these two genes (*ATPase 6/8* and *Cytb)* could be combined in a unique data matrix. The results showed no incongruence between the two gene data sets, and therefore could be analyzed jointly. MEGA X [19] was used to estimate gene characteristics such as, nucleotide composition, conserved/variable sites and transition/transversion ratio. Number of haplotypes, haplotype diversity (*Hd*, probability that two randomly chosen haplotypes are different in a sample) and nucleotide diversity (π, the mean no. of nucleotide differences between two randomly chosen DNA sequences) were estimated using DnaSP v 5.0 [20]. Median-joining network [21] was drawn using Network v5.0 software (Fluxus Technology, Ltd.) to determine the linkages among haplotypes.

Historical patterns of population dynamics were estimated through mismatch distributions and neutrality tests (Tajima's D and Fu's Fs) implemented in Arlequin 3.5. This analysis demonstrates whether the population had undergone expansion in the past. The parameters of demographic expansion, θ_0 and θ_1, and τ (mode of mismatch distribution) were calculated using the $\tau = 2\mu t$ equation. The raggedness index (r) and the sum of squared deviations (SSD) for each population were calculated according to Schneider and Excoffier [22]. These parameters quantify mismatch distribution pattern, where non-significant result indicates population expansion. Gene flow (Nm), was estimated, for each population, using Arlequin 3.5 [23] with 1000 replicates. It is a product of the average size (N) and proportion of immigrants (m) in each population. The levels of gene flow indicate the extent to which each local population may act as an independent evolutionary unit.

The hierarchical partitioning of genetic variation was assessed using analysis of molecular variance (AMOVA). The samples were grouped according to their geographic relevance as well as river basins (10,000 permutations) for testing levels of significance. Mantel test [24] was performed in XL-STAT ver. 2010 (Addinsoft, France), with 10,000 permutations, to determine the extent of isolation by distance (IBD). This analysis was done using two matrices, (pair-wise F_{ST}) and geographical distances (km) between populations.

Table 1 Description of collection localities, number of samples (N) and time of collection of *C. chitala*

River	Site	Coordinates (Longitude/Latitude)	Time of collection	N
Satluj River	Harike Pattan, Punjab	74.94, 31.14	May. 2000	04
			Sep. 2000	24
Gandak River	Valmiki Nagar, Bihar	83.59, 26.75	Nov. 2017	08
GhaghraRiver	Mehmoodabad, UP	81.31, 27.58	Dec. 2000	16
			April. 2001	18
			March. 2002	09
Chauka River	Rampur,	81.38, 27.35	Nov. 2001	06
Sharda River	Sharda Barrage, Lakhimpur Kheri, UP	80.07, 28.77	Nov. 2003	05
Gomti River	Lucknow, UP	80.95, 26.86	Oct. 2016	19
			June. 2017	05
Tons River	Rewa, Madhya Pradesh	81.17, 24.31	March. 2006	05
	Chakghat, UP	81.72, 25.04	May. 2009	02
Son River	Bansagar, Beohari, MP	80.88, 24.055	May. 2011	22
			March.2012	07
			May. 2012	02
			May. 2015	09
			May. 2016	23
			April. 2017	03
Ken River	Rangua,MP	79.89, 24.69	Sep. 2015	14
			June. 2016	02
			July. 2017	04
Hooghly River	Nabadeep, WB	86.37, 23.4,	Jan. 2001	13
			April. 2016	10
	Beldanga, WB	88.19, 23.94	April. 2016	02
	Satui, WB	88.19, 23.95	May. 2001	02
	Behrampur, WB	88.21, 24.1	May. 2001	02
			June. 2001	03
			July. 2002	01
Padma River (Ganga)	Farakka, WB	87.90, 25.06	June. 2001	02
			May. 2002	08
			July. 2003	14
			May 2006	05
			April. 2016	03
			Nov. 2016	06
	Manikchak, WB	89.33, 23.93	April. 2016	02
Brahmaputra River	Kolongpar, Assam	91.82, 26.21	Jan. 2000	47
			Feb. 2005	09
	Uzan Bazaar, Assam	91.75, 26.19	Nov. 2005	16
	Guwahati, Assam	91.72, 26.18	Dec. 2007	03
Mahanadi River	Hirakud Dam, Odisha	83.86, 21.53	May. 2004	18
	Sonepur, Odisha	83.93, 20.84	May. 2004	05
	Jobra Barrage, Odisha	85.9, 20.85	April. 2005	14
Narmada River	Hoshangabad, MP	77.70, 22.77	April. 2003	11
Total				403

Population structure was estimated using two clustering approaches, Principal Coordinate Analysis (PCoA), implemented in GenAlex 6 [25] and Bayesian assignment method through BAPS version 6.0 [26]. PCoA plot was

prepared based on the covariance of the genetic distance matrix. In BAPS, samples were grouped according to river basins to estimate the pattern of clustering in concatenated mitochondrial genes. We employed a mixture model and tested K from 1 to 14 (maximum no. of clusters) with 1000 replicates per K. The number of reference individuals was set to 200.

Temporal stability of genetic variation

Demonstrating temporal stability is a powerful way of confirming the reliability of observed spatial genetic patterns. In this investigation, we used historical samples of *C. chitala* from three rivers, Son, Hooghly and Padma. Because of substantial intervening time between tissue sampling from these rivers, temporal stability of *C. chitala* populations was studied, so that these samples could be analyzed together with contemporary data. The samples from Son were collected during 2011–2012 (n = 31) and 2015–2017 (n = 35), while Padma (n = 29) samples during 2001–2003, 2006 and 2016 (n = 11), with almost a gap of two to five generations. The collections of Hooghly samples had two sampling points, 2001–2002 (21) and 2016 (12). The pair-wise F_{ST} values were calculated for each population, collected at different time intervals. The results showed low and insignificant genetic differentiation between different sampling points (for each river), therefore the samples were merged river-wise, for further data analysis. Hereafter, the samples from Son1 & 2, Hooghly1 & 2 and Padma1 & 2 were referred to as Son, Hooghly and Padma, respectively (Suppl Table 1).

Results

Genetic variability

Sequence analyses of *ATPase 6/8* (842 bp) and *Cytb* (1139 bp) yielded 23 and 30 haplotypes, respectively. *ATPase 8* spanned from 1 to 168 bp, whereas *ATPase 6* from159—842 bp, with overlapping region of 10 bp. *ATPase 6/8* genes were characterized by 22 variable sites, 13 parsimony informative and 9 singleton sites, whereas *Cytb* had 26 variable sites, 17 parsimony informative and 9 singletons. The partition homogeneity test indicated no incongruence between the two genes and therefore, we concatenated the sequences (5′*ATPase 8-ATPase 6-Cytb*-3′) as a data set for further analysis. A total of 61 haplotypes were identified for concatenated mitochondrial DNA sequences. The results of individual genes, as well as the concatenated sequence (*ATPase 6/8* + *Cytb*), were concordant. The average frequency of nucleotides were A: 30.7%, T: 27.1%, G: 12.6%, C: 29.6%, with a transition to transversion ratio (Ts:Tv) of 6.479. The nucleotide composition of both individual genes as well as the concatenated displayed bias towards A + T content. The haplotype diversity *Hd* ranged from 0.000 (Mahanadi) to 0.933 (Chauka).

Table 2 Intra-population haplotype (*h*) and nucleotide (π) diversities for *ATPase 6/8, cytochrome b* gene and their concatenated sequence. This shows moderate to high haplotype diversity and low to moderate nucleotide diversity

River	*ATPase 6/8*			*Cytb*			Concatenated		
	Hd	Variance of *Hd* with standard deviation	π ± SD	*Hd*	Variance of *Hd* with standard deviation	π ± SD	*Hd*	Variance of *Hd* with standard deviation	π ± SD
Satluj	0.632	0.007 ± 0.083	0.0011 ± 0.0009	0.479	0.011 ± 0.107	0.0008 ± 0.0007	0.767	0.006 ± 0.075	0.0009 ± 0.0006
Gandak	0.607	0.027 ± 0.164	0.0008 ± 0.0008	0.571	0.009 ± 0.094	0.0020 ± 0.0014	0.750	0.019 ± 0.139	0.0015 ± 0.0010
Ghaghra	0.625	0.003 ± 0.050	0.0009 ± 0.0007	0.652	0.005 ± 0.070	0.0019 ± 0.0012	0.860	0.001 ± 0.034	0.0015 ± 0.0009
Chauka	0.733	0.024 ± 0.155	0.0011 ± 0.0010	**0.800**	0.029 ± 0.172	**0.0028 ± 0.0019**	**0.933**	0.015 ± 0.122	0.0020 ± 0.0014
Sharda	**0.900**	0.026 ± 0.161	**0.0014 ± 0.0013**	0.600	0.031 ± 0.175	0.0021 ± 0.0016	0.900	0.026 ± 0.161	0.0018 ± 0.0013
Gomti	0.159	0.009 ± 0.094	0.0002 ± 0.0003	0.391	0.008 ± 0.091	0.0014 ± 0.0009	0.507	0.009 ± 0.093	**0.0088 ± 0.0006**
Tons	0.286	0.039 ± 0.196	0.0003 ± 0.0005	0.667	0.026 ± 0.160	0.0023 ± 0.0016	0.714	0.033 ± 0.181	0.0014 ± 0.0010
Son	0.434	0.005 ± 0.071	0.0006 ± 0.0006	0.736	0.002 ± 0.043	0.0022 ± 0.0013	0.842	0.001 ± 0.032	0.0015 ± 0.0009
Ken	0.442	0.008 ± 0.087	0.0005 ± 0.0005	0.521	0.002 ± 0.042	0.0014 ± 0.0010	0.679	0.003 ± 0.052	0.0010 ± 0.0007
Hooghly	0.371	0.010 ± 0.100	0.0005 ± 0.0005	0.801	0.003 ± 0.051	0.0024 ± 0.0015	0.858	0.002 ± 0.043	0.0016 ± 0.0010
Padma	0.523	0.007 ± 0.086	0.0007 ± 0.0007	0.699	0.002 ± 0.046	0.0021 ± 0.0013	0.856	0.001 ± 0.036	0.0015 ± 0.0009
Brahmaputra	0.503	0.004 ± 0.066	0.0007 ± 0.0006	0.391	0.005 ± 0.068	0.0011 ± 0.0008	0.698	0.003 ± 0.056	0.0010 ± 0.0006
Mahanadi	**0.000**	0.000 ± 0.000	**0.0000 ± 0.000**	**0.000**	0.000 ± 0.000	**0.0000 ± 0.0000**	**0.000**	0.000 ± 0.000	**0.0000 ± 0.0000**
Narmada	0.691	0.016 ± 0.128	0.0013 ± 0.0010	0.473	0.026 ± 0.162	0.0004 ± 0.0005	0.727	0.021 ± 0.144	0.0008 ± 0.0006

SD- Standard Deviation; The highest and lowest *Hd* and π values are presented in bold

Mean nucleotide diversity (π) of 0.002 was observed, with a range of 0.000 (Mahanadi) to 0.002 (Chauka) (Table 2).

Phylogenetic relationships among haplotypes

The phylogeographic inferences were performed using the median-joining network of concatenated data set (Fig. 2). Haplotypes H3, H10 and H12 were the most common, out of total 61 haplotypes. Haplotype H3 was shared between all populations except Gomti and Tons, Haplotype H10 (in populations, except Chauka, Ken, Mahanadi, and Narmada) while H12 was represented by 11 populations except for Satluj, Sharda and Mahanadi. Population-specific haplotypes were found in Satluj (5), Ghaghra (6), Sharda (1), Son (11), Hooghly (3), Padma (3), Brahmaputra (9) and Narmada (4). Samples from Son exhibited the maximum number of haplotypes (19), followed by Brahmaputra (18) and Ghaghra (15). River Satluj was represented by 10 haplotypes, Narmada (6), whereas Mahanadi was represented by only one haplotype. A total of 42 haplotypes were population-specific, while 19 were shared.

Demographic history and neutrality tests

Cytb gene did not show any significant negative values for Tajima and Fu's Fs, whereas *ATPase 6/8* showed significant negative values for Fu's Fs and Tajima's D in Padma and Brahmaputra. Significant negative values of Fu's Fs were observed in two rivers, Sharda and Son. Pairwise mismatch distribution, Tajima's D and Fu's Fs tests were performed for combined mitochondrial genes for each population. The parameters used to estimate the pattern of population dynamics are shown in Table 3. The mean value of Tajima's D and Fu's Fs were 0.103 and -2.068, respectively. Based on combined (1.98 kb) gene analysis, the genetic constraints for *C. chitala* were evident in rivers Satluj, Narmada and five rivers (Ghaghra, Son, Hooghly, Padma, and Brahmaputra) of Ganga–Brahmaputra basin. The mismatch distribution plot, for all populations together, displayed a multimodal pattern, revealing demographic equilibrium or a stable population [27] (Fig. 3). Based on the sum of squared deviations, all populations were non-significant except Gomti and Brahmaputra. The raggedness indices were also calculated under the demographic expansion model for each population, wherein all populations were non-significant (Table 3). The rivers belonging to Ganga basin, when analyzed separately, presented negative significant Fu's Fs for all rivers, except Sharda.

Genetic differentiation among populations

The overall coefficient of genetic differentiation (F_{ST}) was 0.263 for combined data set. The pair-wise F_{ST} ranged from 0.000 (Sharda-Son) to 0.866 (Satluj-Mahanadi) (Table 4).

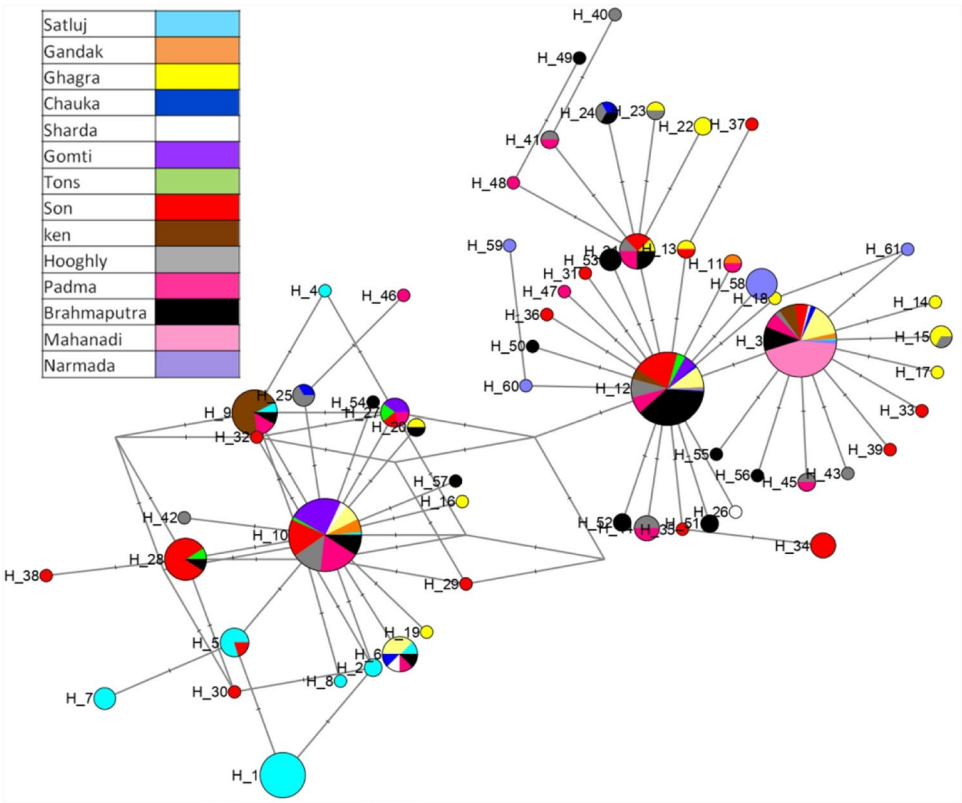

Fig. 2 Median joining network of 61 concatenated haplotypes of *C. chitala* (Network V 5.0), with star-like topologies, indicating population expansion. (The size of nodes are proportional to the number of individuals. Perpendicular tick marks on the lines joining haplotypes represent the number of nucleotide substitutions)

Table 3 Summary of diversity, neutrality and expansion time estimated by *ATPase 6/8*, *Cytb* and concatenated sequences for *C. chitala*

Population	Satluj	Gandak	Ghaghra	Chauka	Sharda	Gomti	Tons	Son	Ken	Hooghly	Padma	Brahmaputra	Mahanadi	Narmada
Cyt B														
H	5	2	10	4	2	2	3	11	2	11	6	9	1	3
Ts	6	4	11	7	4	4	5	9	3	13	8	8	0	1
Tv	0	0	1	0	0	0	0	0	0	0	0	1	0	1
S	6	4	12	7	4	4	5	9	3	13	8	9	0	2
Tajima's D	−1.090	2.101	−0.687	0.128	1.641	1.261	1.288	0.870	2.266	−0.435	0.737	−0.770	0.000	−0.778
Fu's Fs	−0.805	3.933	−1.956	0.314	3.022	4.651	1.853	−1.182	4.362	−2.348	1.407	−2.251	0.000	−0.659
T	0.000	5.070	5.488	5.379	4.986	3.000	5.420	5.656	4.061	5.393	4.625	3.000	0.000	0.645
T	0.000	0.139	0.151	0.148	0.137	0.082	0.149	0.155	0.111	0.148	0.127	0.082	0.000	0.018
θ_0	0.000	0.002	0.000	0.004	0.000	0.000	0.002	0.002	0.000	0.000	0.281	0.000	0.000	0.000
θ_1	107	4.495	2.455	6.287	6.201	0.455	5.042	3.785	2.258	4.627	3.352	0.538	0.000	999
SSD	0.306	0.329	0.046	0.043	0.341	0.205	0.141	0.045	0.266	0.029	0.077	0.062	0.000	0.015
R	0.105	**0.837***	0.122	0.138	0.880	0.677	0.283	0.084	**0.772***	0.062	0.166	0.339	0.000	0.170
ATPase 6/8														
H	5	3	6	3	4	2	2	6	2	4	8	9	1	4
Ts	3	1	3	1	2	0	0	3	0	2	6	7	0	2
Tv	1	1	2	1	1	1	1	2	1	1	1	1	0	1
S	4	2	5	2	3	1	1	5	1	3	7	8	0	3
Tajima's D	−0.286	−0.448	−0.877	−0.050	−1.048	−0.681	−1.006	−1.184	1.026	−0.918	**−1.712***	**−1.617***	0.000	0.101
Fu's Fs	−0.938	−0.478	−2.087	−0.427	**−1.938***	−0.249	−0.095	**−2.909***	1.169	−1.410	**−5.547***	**−6.146***	0.000	−0.627
T	1.273	0.877	0.910	1.113	1.500	2.965	2.982	0.574	0.650	0.469	0.721	0.693	0.000	1.473
T	0.073	0.050	0.052	0.064	0.086	0.169	0.170	0.033	0.037	0.027	0.041	0.040	0.000	0.084
θ_0	0.004	0.000	0.002	0.000	0.000	0.450	0.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000
θ_1	6.309	999	999	999	999	0.450	3.600	999	999	999	999	999	0.000	15.162
SSD	0.007	0.038	0.024	0.064	0.088	0.026	**0.249***	0.003	0.012	0.000	0.007	0.007	0.000	0.019
R	0.085	0.241	**0.180***	0.347	0.430	0.489	0.265	0.136	0.209	0.167	0.120	0.124	0.000	0.134
Concatenated														
H	10	4	15	5	4	3	4	19	3	14	14	18	1	6
Ts	9	5	14	8	6	4	5	12	3	15	14	15	0	3
Tv	1	1	3	1	1	1	1	2	1	1	1	2	0	2
S	10	6	17	9	7	5	6	14	4	16	15	17	0	5
Tajima's D	−0.865	1.310	−0.841	0.088	0.498	0.835	0.847	0.074	2.254	−0.614	−0.481	−1.339	0.000	−0.322
Fu's Fs	**−3.662***	0.976	**−4.973***	−0.746	−0.036	2.996	0.562	**−6.864***	3.226	**−4.462***	**−4.043***	**−9.699***	0.000	**−2.228***
T	1.590	5.660	6.230	6.479	6.107	0.000	5.662	5.211	4.498	5.188	4.953	0.143	0.000	2.453
θ_0	0.555	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002	0.002	0.000	0.000	0.000
θ_1	5.098	8.862	4.644	12.715	14.026	999	7.356	5.513	3.666	6.515	5.920	999	0.000	4.312
SSD	0.003	0.130	0.025	0.043	0.121	**0.394***	0.120	0.015	0.076	0.012	0.024	**0.384***	0.000	0.023

Table 3 (continued)

Population	Satluj	Gandak	Ghaghra	Chauka	Sharda	Gomti	Tons	Son	Ken	Hooghly	Padma	Brahmaputra	Mahanadi	Narmada
R	0.021	0.313	0.051	0.093	0.270	0.371	0.259	0.028	0.191	0.030	0.048	0.053	0.000	0.073
Mismatch	Multimod	Multimod	Multimod	Multimod	Multimod	Multimod	Multimod	Multimod	Multimod	Multimod	Multimod	Multimod	Multimod	Multimod

H No. of haplotypes, Ts No. of transitions, Tv No of transversions, S Segregating sites, τ Time in number of generation, T Time after expansio, θ_0 Pre-expansion population size, θ_1 Post-expansion population size, SSD Sum of Squared Deviation, r Raggedness index, *Significant value are presented in bold ($p < 0.05$). The 999 value of θ_1 depicted a pure expansion

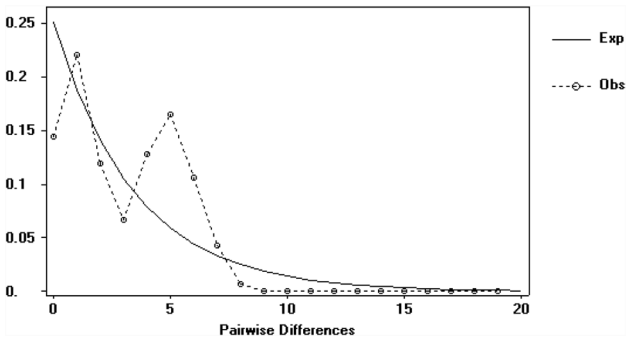

Fig. 3 Mismatch graph (DnaSP v5.0) illustrating multimodal pattern for overall locations based on pairwise sequence differences. The overall pattern displays demographic equilibrium or stable population (The solid line and dotted line represent the curves for expected and observed values)

The mean F_{ST} of Ganga basin reduced remarkably (0.074) when analyzed separately. The pair-wise population genetic differentiation in concatenated genes was similar to both the individual genes, when analyzed separately (Suppl Table 2).

The Principal Coordinate Analysis (PCoA), based on the pair-wise genetic distances between populations, resulted in existence of three distinct clusters, and a mixed cluster of rivers belonging to Ganga basin (Fig. 4). The eigenvalue for PC1 was found to be significant (1.116) and contributed 74.62%, while PC2 was found to be non-significant, contributing 10.31%. The cumulative total variation by PC1 and PC2 was 84.94%. Basin-wise, principal component scores, (PC1, PC2) for Satluj (0.691, 0.134), Mahanadi (−0.484, 0.301) and Narmada (− 0.403, − 0.150) suggested that they are genetically distinct, while these scores were comparable for rivers of Ganga basin (Gandak, Ghaghra, Chauka, Sharda, Gomti, Tons, Son, Ken, Hooghly and Padma) and Brahmaputra. These rivers formed a mixed cluster, distant from three other basins (Suppl Table 3).

Bayesian clustering partitions the individuals into genetically distinct units without prior group assignment. In this study, BAPS 6.0 was used to ascertain the number as well as the putative mixing within the basins. Based on log likelihood estimates, K = 4 was found to be the best partition (Fig. 5). Cluster I, comprised haplotypes from the Ganges; Cluster II, haplotypes from Ganges and Narmada; Cluster III, haplotypes from the Ganges, Mahanadi, and Satluj; Cluster IV, haplotypes from Ganges and Satluj. The representation of Ganges samples in all the clusters was observed. The lack of river basin wise clustering was obvious from the results. However, the Mahanadi population was genetically differentiated and therefore the most distinct (Fig. 6).

A three way AMOVA was performed, considering four groups namely, Satluj, Ganga, Mahanadi and Narmada. The hierarchical analysis of molecular variance demonstrated that 19.01% of the total genetic variance was due to

Table 4 Pairwise F_{ST} (below diagonal), population-specific F_{ST} (at diagonal) and gene flow (Nm values) (above diagonal) of *C. chitala* samples from fourteen rivers studied through concatenated sequences indicated genetic differentiation among rivers

Population	Satluj	Gandak	Ghaghra	Chauka	Sharda	Gomti	Tons	Son	Ken	Hooghly	Padma	Brahmaputra	Mahanadi	Narmada
Satluj	0.255	0.698	0.453	0.488	0.912	0.971	0.591	0.668	0.410	0.608	0.663	0.291	0.077	0.194
Gandak	**0.417**	0.245	**19.194**	**inf**	**inf**	**8.010**	**inf**	**inf**	**14.633**	**inf**	**inf**	**2.755**	0.157	0.941
Ghaghra	**0.524**	0.025	0.242	**inf**	**5.972**	**1.238**	**inf**	**10.334**	**11.935**	**34.508**	**19.313**	**18.427**	**1.221**	**2.577**
Chauka	**0.506**	0.000	0.000	0.235	**inf**	**1.650**	**inf**	**inf**	**809.707**	**inf**	**inf**	**17.749**	0.256	**1.916**
Sharda	**0.354**	0.000	0.077	0.000	0.241	**inf**	**inf**	**inf**	**4.920**	**inf**	**inf**	**1.549**	0.085	0.689
Gomti	**0.340**	0.059	**0.288**	**0.233**	0.000	0.257	**3.822**	**2.633**	**1.165**	**2.360**	**3.118**	0.684	0.110	0.335
Tons	**0.458**	0.000	0.000	0.000	0.000	0.116	0.247	**inf**	**65.880**	**inf**	**inf**	**6.101**	0.140	**1.184**
Son	**0.428**	0.000	**0.046**	0.000	0.000	**0.160**	0.000	0.240	**8.522**	**82.987**	**88.515**	**5.344**	0.702	**1.654**
Ken	**0.549**	0.033	0.040	0.001	0.092	**0.300**	0.008	**0.056**	0.254	**13.077**	**19.182**	**4.031**	0.323	**1.105**
Hooghly	**0.451**	0.000	0.014	0.000	0.000	**0.175**	0.000	0.006	0.037	0.238	**inf**	**7.732**	0.608	**1.891**
Padma	**0.430**	0.000	0.025	0.000	0.000	**0.139**	0.000	0.006	0.025	0.000	0.242	**4.999**	0.599	**1.516**
Brahmaputra	**0.632**	0.154	0.026	0.028	**0.244**	**0.422**	0.076	**0.086**	**0.110**	**0.061**	**0.091**	0.255	0.782	**2.551**
Mahanadi	**0.866**	**0.762**	**0.291**	**0.661**	**0.855**	**0.820**	**0.781**	**0.416**	**0.607**	**0.452**	**0.455**	**0.390**	0.279	0.171
Narmada	**0.720**	**0.347**	**0.162**	**0.207**	**0.420**	**0.598**	**0.297**	**0.232**	**0.311**	**0.209**	**0.248**	**0.164**	**0.746**	0.260

Significant pairwise F_{ST} values ($p < 0.05$) and $Nm > 1$ are in bold

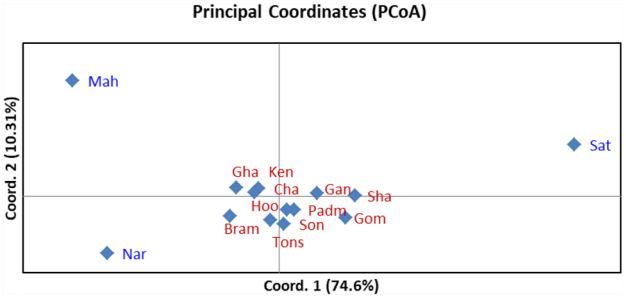

Fig. 4 Principal coordinate analysis (PCoA) (GenAlex 6) using pairwise distances between populations of *C. chitala.* The pattern showed existence of three distinct clusters of Satluj, Mahanadi and Narmada and a mixed cluster of rivers belonging to Ganga basin and Brahmaputra

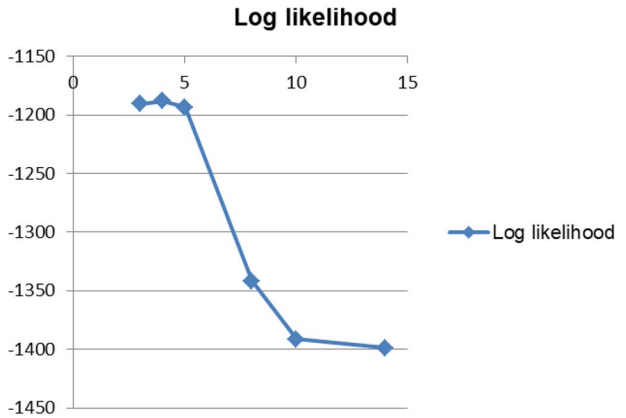

Fig. 5 Plot of K (x axis) against log likelihood value (y axis) for BAPS in *C. chitala*

differences among groups ($F_{CT} = 0.190$, $p < 0.05$), 73.71% was due to differences within populations ($F_{ST} = 0.263$, $p < 0.05$) and 7.27% was attributed to variation among populations within groups ($F_{SC} = 0.090$, $p < 0.05$) (Table 5).

Results of Isolation by Distance (IBD) were not statistically significant when all populations were considered together, showing thereby that there existed no correlation between genetic and geographical distance matrices ($R^2 = 0.008$, $r = 0.179$; $p = 0.121$).

Population genetic connectivity

Gene flow estimates (number of migrants, Nm) showed moderate to high values (Table 4). According to Nei (1987), Nm values (> 1) contributed positively against genetic differentiation among populations. In present study, higher gene flow was observed among the rivers of Ganga basin (Gandak, Ghaghra, Chauka, Sharda, Gomti, Tons, Son, Ken, Hooghly, Padma). The Brahmaputra showed moderate to higher gene flow for Gangetic rivers.

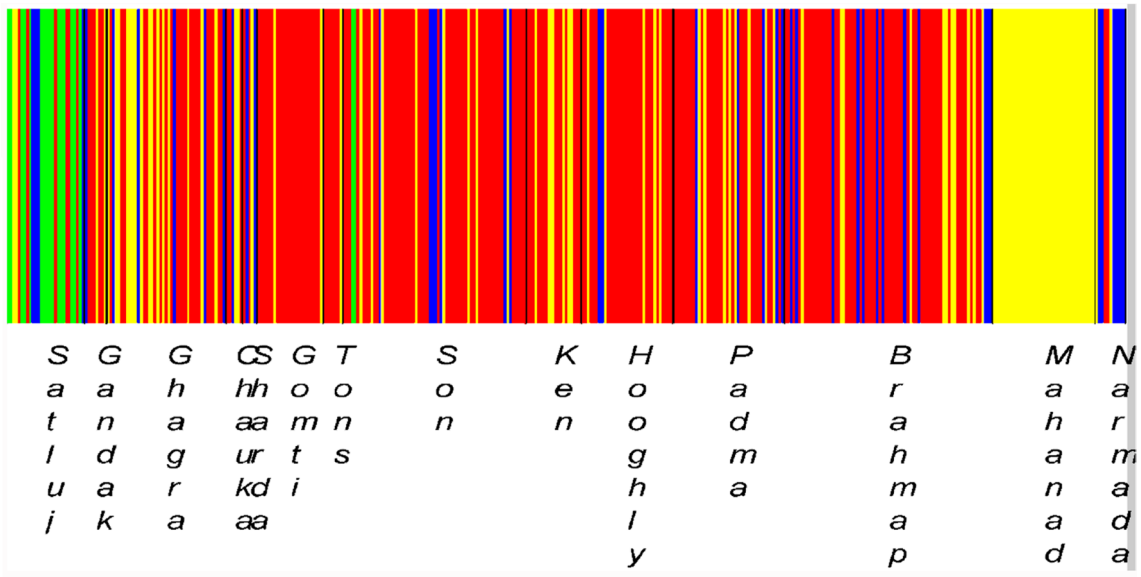

Fig. 6 Histogram of the assignment test using BAPS6.0

It was observed that rivers Satluj, Mahanadi and Narmada presented low to moderate levels of gene flow, however with pronounced genetic differentiation.

Discussion

Information of genetic variability and population structure can be utilized in identifying distinct management units which is needed for effective conservation of fisheries resources. This study provides useful inputs and delineates stock boundaries of *C. chitala.* A combined dataset (*ATPase 6/8* + *Cytb*) revealed high haplotype diversity and low nucleotide diversity across all riverine populations, except for river Mahanadi. The haplotype and nucleotide (%) diversity levels are considered high if the value is greater than 0.5 [28]. This pattern, particularly (high *Hd* and low π) might be attributed to demographic expansion that occurred after a reduction in effective population size, retaining new mutations [29]. Similar trend in molecular diversities are concordant to other notopterids, *Chitala lopis* [30], *Chitala chitala* [5] and *N. notopterus* [13, 31]. Overall, the pattern of haplotype sharing among the river basins indicated genetic structuring in *C. chitala*. In Mahanadi population, a single haplotype was fixed in all individuals. This might indicate a possibility

Table 5 AMOVA analysis displaying hierarchical partitioning of genetic variance based on *ATPase 6/8*, *Cytb* and concatenated sequences from fourteen riverine localities of *C. chitala*

Source of variation	*ATPase 6/8 gene*			*Cytochrome b gene*			Concatenated		
	Percentage of variation	Fixation Index	P value*	Percentage of variation	Fixation Index	P value	Percentage of variation	Fixation Index	P value
Among groups	26.23	**F_{CT}: 0.262**	0.022 ± 0.005	16.51	**F_{CT}: 0.165**	0.119 ± 0.011	19.01	**F_{CT}: 0.190**	0.042 ± 0.003
Among populations	2.36	**F_{SC}: 0.032**	0.066 ± 0.007	9.08	**F_{SC}: 0.109**	0.003 ± 0.002	7.27	**F_{SC}: 0.090**	0.004 ± 0.001
Within groups									
Within populations	71.41	**F_{ST}: 0.286**	0.000 ± 0.000	74.41	**F_{ST}: 0.256**	0.000 ± 0.000	73.71	**F_{ST}: 0.263**	0.000 ± 0.000

*Significant values are presented in bold ($p < 0.05$)

of small effective population size and inbreeding. Similar results were observed in another notopterid, *N. notopterus*, where the lack of genetic connectivity in Mahanadi and Indo-Gangetic Rivers was noticed [13]. It is likely that the fragmentation of Mahanadi populations during the early Pleistocene [32] might be responsible for present loss of haplotype sharing. In our earlier study, the partial sequence of *Cytb* (307 bp) revealed only one haplotype in Satluj samples [5]. However, in the present investigation, the concatenated gene set yielded five population-specific haplotypes, of which two haplotypes were found to be ancestral, giving rise to other related haplotypes. This supports the potential of longer gene sequences for genetic analysis [33]. However, higher genetic variability in river Satluj might also be due to the comparatively larger sample size, used in this study. Small sample size might affect the frequency and distribution of variation as well as amount of uncertainty in the estimates and resolution. In present study, though the sample sizes from four rivers Gandak, Chauka, Sharda and Tons were small, the patterns of haplotype sharing, frequency and clustering were consistent with other rivers of the basin.

At population level, genealogical relationships among haplotypes are preferred to phylogenetic methods, because the network accommodates both, existent ancestral sequence as well as alternative connections [21]. In present study, haplotype network presented three haplogroups displaying star-like topologies, indicating genetic bottleneck and population expansion [34]. These results are important in view of adaptive potential and probability of population persistence [35]. The pattern of molecular diversity levels show signatures of population expansion in Satluj, Ghaghra, Son, Hooghly, Padma, Brahmaputra and Narmada.

Neutrality tests are considered sensitive measures of demographic events, where significant negative values indicate signatures of population expansion [36]. The results of Fu's Fs showed concordance with diversity patterns and explained the events of population expansion in these rivers. The population expansion was also confirmed by non significant raggedness index. Tajima's D values were negative all the seven populations, indicating the excess of rare nucleotide site variants. Moreover, on analyzing Gangetic rivers separately, significantly negative Fu's Fs were observed in all populations except Sharda, which may be due to small sample size, analyzed in present study. It is likely that founder effect may be responsible for genetic constraints in these populations.

The F_{ST} indicates the genetic differentiation between populations. In general, a higher value (>0.25) is correlated with a high genetic differentiation between populations [37], which is consistent with our study ($F_{ST}=0.26$) and with other freshwater species like *N. notopterus* [13]. In *C. chitala* individuals from river Satluj, Mahanadi, and Narmada showed significant differentiation from other rivers. These three basins presented significant pair-wise $F_{ST,}$ suggesting limited gene flow. The degree of gene flow presented high level of genetic connectivity among rivers of Ganga basin; and these populations maintained genetic homogeneity, probably due to ancestral mixing. However, in another congener *Chitala lopis,* the low level genetic differentiation and corresponding higher gene flow were observed [30]. Sodsuk and Sodsuk (2000) assayed 28 allozyme loci for *C. lopis*, but no locus exhibited polymorphism [38].

Barby et al. [39] illustrated the biogeography of notopteridae, wherein the arrival of Chitala species in Asian continent had been estimated to be about 55 Mya. This corresponded Eocene epoch, the estimated entry period of several other spacies including carps and catfishes [40]. The contemporary geological events were considered responsible for formation of Indus and Ganga, with its Himalayan tributaries. Satluj samples showed presence of distinct haplotypes, however limited connectivity might have allowed sharing of haplotypes with Gangetic rivers. Our analysis supported this hypothesis and Satluj populations were distinct from others. Chauhan et al. [41] have described the connectivity between Satluj and rivers of Ganga system via erstwhile river Ghaggar. Not enough variance and differentiation were observed for component rivers of Ganga system, in current study. All the rivers (Gandak, Ghaghra, Chauka, Sharda, Gomti, Tons, Son, Ken, Hooghly and Padma) clustered together, which is indicative of common gene pool in Gangetic system. The Ganges and the Brahmaputra river system flow through very different regions for most of their lengths and join upstream at the Bay of Bengal. This may be the reason for the genetic homogeneity between tributaries of the Ganges and Brahmaputra. Similar results of genetic proximity between Ganga and Brahmaputra have been reported in *C. chitala* [5] and another freshwater species *Sperata seenghala* [42].

River Narmada, is among the longest rivers in India, which originates from Amarkantak (M.P., India). The occurrence of physical barriers seems to contribute towards genetic fragmentation and dispersal capabilities of *C. chitala*. Earlier, Khedkar et al. [43] described the genetic impoverishment of two species (*Catla catla* and *Mastacembelus armatus*) owing to asymmetrical dispersal in river Narmada. Moreover, a study conducted on alluvial plain sedimentation for river Narmada envisaged that it had been a much bigger channel during late Pleistocene with higher discharge than present [44]. This might also contribute to genetic fragmentation of resources in river Narmada.

The Mahanadi river flows eastwards and drains into Bay of Bengal. Historically, the tributaries of Narmada river, *i.e.* Mahanadi and Godavari used to flow westward, but due to rise in plateau and tilt of peninsula during the middle Pleistocene, they changed their course [45] which resulted

in rivers, as seen today, flowing as independent river systems, southward through Eastern Ghats [32]. Such phylogeographic events affect long term histories in populations like Mahanadi. This became explicit, when Mahanadi samples were excluded from analysis. Out of a total of 26.28% genetic differentiation, over 6% was caused due to Mahanadi samples. It is likely that the population might be experiencing selection pressure. Moreover, this warrants for larger sample sizes, analyzed with multiple marker systems.

In a study, Martinez et al. [46] correlated the impact of habitat type with the diversity levels in fishes. Lower diversity levels in River Gomti might be attributed to multiple factors related to habitat. This river has been reported to have poor vegetation cover in the banks, siltation and low water velocity. Present findings are also supported by a recent investigation, wherein the negative allometric growth with poor condition factor of *C. chitala* in river Gomti, was reported [47]. Moreover, Verma and Dalela [48] pointed out the decline in featherbacks and carps due to the rise in pollution in river Gomti caused by domestic and industrial wastes. Similar decline in genetic diversity due to habitat modification has been observed in *C. chitala* in Mekong river [49]. River Brahmaputra is characterised by large and variable flow, enhanced rate of sediment discharge and distinct river morphology. Additionally, reduction in efficiency of energy utilization from primary energy to fish, was reported [50]. The primary consumers (both herbivore and detritivore) contributed only 17.1% of total energy. Moreover, anthropogenic activities such as dam construction might have caused alterations in fish breeding grounds resulting in the loss of genetic diversity.

The analysis of molecular variance also confirmed the sub-structuring in *C. chitala* poupulations. Similar patterns of genetic differentiation and variance partitioning had been obtained using *Cytb* and control region [5]. Bayesian clustering produced the highest log likelihood values for four clusters, however failed to resolve river basin wise. The Mantel tests in *C. chitala* populations rejected the hypothesis that genetic differentiation and geographic distance were mutually correlated. This lack of interdependence confirmed that the genetic differentiation among populations was not determined by the effect of distance between populations. The resulting genetic divergence in *C. chitala* from different river basins appeared to be an outcome of restricted gene flow between populations as well as phylogeographic events. These findings highlight that both geographic and environmental factors contributed towards mitochondrial variability patterns. Overall, the patterns of pairwise genetic differentiation, gene flow, PCoA and AMOVA indicated the existence of four management units namely, Ganges–Brahmaputra, Satluj, Mahanadi and Narmada. This baseline information will be useful in conservation and management of *C. chitala*, which has significant conservation and aquaculture value.

Conclusion

Understanding genetic structure of a species is among the key components, which is essential for planning management strategies. The resource managers should be aware with management units of important fish species. *C. chitala* is important for conservation and aquaculture viewpoint. This study unraveled high degree of genetic differentiation with evidences for presence of four natural genetic stocks. Demographic parameters indicate that there is need to assess a fine scale structuring through use of multilocus markers, particularly in Gangetic rivers. Two other rivers, Gomti and Brahmaputra, displayed lower diversities, while Mahanadi presented absence of alternate haplotype. These populations are important for conservation, and the resource managers need to adopt suitable measures, in this context. This may be achieved through cohesive approaches of policy makers with multiple stakeholders and researchers.

Acknowledgments This work has been carried out under the project scheme of the Indian Council for Agricultural Research, India entitled, "Outreach Activity on Fish Genetic Stocks Phase-II". The authors are thankful to Director, ICAR-National Bureau of Fish Genetic Resources, Lucknow, India for providing the facilities needed for executing this work. The authors are also grateful to Sh Pradipta Paul, Dr. L.K Tyagi, Dr. Santosh Kumar, Sh Kantharajan G, Dr. Achal Singh, and Sh Rajkumar for their help in sampling and geo-referenced riverine map preparations. Sh R.S. Sah and Sh Rajesh Kumar are thankfully acknowledged for their excellent technical assistance.

Compliance with ethical standard

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical approval All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. The methodologies were approved by the Institutional Animal Ethical Committee (IAEC), ICAR-NBFGR, Lucknow, India vide No. G/CPCSEA/IAEC/2015/2 dated 27 October, 2015.

Informed consent Informed consent was obtained from all the authors who were involved in this piece of work. The manuscript is submitted after forwarding through the prescribed procedures of the institute with the approval of the competent authority.

References

1. Hilton EJ, Lavoué S (2018) A review of the systematic biology of fossil and living bony-tongue fishes, Osteoglossomorpha (Actinopterygii: Teleostei). Neotropical Ichthyology 16(3):180031

2. Hamilton F (1822) An account of the fishes found in the river Ganges and its branches. Printed for A, Constable and company
3. Froese R, Pauly D (2003) Fish Base WWW Electronic Publication. World Fish. https://www.fishbase.org
4. Taverne L (1998) Les ostéoglossomorphes marins de l'Éocène du Monte Bolca (Italie): Monopteros Volta, 1796, Thrissopterus Heckel, 1856 et Foreyichthys Taverne, 1979. Considérations sur la phylogénie des téléostéens ostéoglossomorphes. Studi e ricerche sui giacimenti terziari di Bolca 7:67–158
5. Mandal A, Mohindra V, Singh RK, Punia P, Singh AK, Lal KK (2012) Mitochondrial DNA variation in natural populations of endangered Indian Feather-Back Fish. Chitala chitala. Molecular biology reports. 39(2):1765–1775
6. Sarkar UK, Deepak PK, Negi RS, Singh S, Kapoor D (2006) Captive breeding of endangered fish Chitala chitala (Hamilton-Buchanan) for species conservation and sustainable utilization. In Marine, Freshwater, and Wetlands Biodiversity Conservation (pp. 211–221). Springer, Dordrecht
7. Chaudhry S (2010) *Chitala chitala*. The IUCN Red List of Threatened Species 2010: e.T166510A6225101
8. Sarkar UK, Negi RS, Deepak PK, Lakra WS, Paul SK (2008) Biological parameters of the endangered fish *Chitala chitala* (Osteoglossiformes: Notopteridae) from some Indian rivers. Fish Res 90(1–3):170–177
9. Sarkar UK, Deepak PK, Negi RS (2009) Length–weight relationship of clown knifefish *Chitala chitala* (Hamilton 1822) from the River Ganga basin. India J App Ichthyol 25(2):232–233
10. Ward RD, Grewe PM (1994) Appraisal of molecular genetic techniques in fisheries. Rev Fish Biol Fisheries 4(3):300–325
11. Crandall KA, Posada D, Vasco D (1999, November) Effective population sizes: missing measures and missing concepts. In Animal Conservation forum (Vol. 2, No. 4, pp. 317–319). Cambridge University Press
12. Goios A, Alvarez L (2013) Mitochondrial Genetics/Evolution. In: Stanley M, Kelly H (eds) Brenner's Encyclopedia of Genetics, 2nd edn. Elsevier, London p, pp 439–440
13. Gupta A, Lal KK, Mohindra V, Singh RK, Punia P, Dwivedi AK, Gupta BK, Luhariya RK, Masih P, Mishra RM, Jena JK (2013) Genetic divergence in natural populations of bronze featherback, *Notopterus notopterus* (Osteoglossiformes: Notopteridae) from five Indian rivers, analyzed through mtDNA ATPase6/8 regions. Meta gene 1:50–57
14. Singh RK, Lal KK, Mohindra V, Punia P, Sah RS, Kumar R, Gupta A, Das R, Lakra WS, Ayyappan S (2012) Genetic diversity of Indian Major Carp, *Labeo calbasu* (Hamilton, 1822) populations inferred from microsatellite loci. Biochem Syst Ecol 44:307–316
15. Xiao W, Zhang Y, Liu H (2001) Molecular Systematics of Xenocyprinae (Teleostei: Cyprinidae): Taxonomy, Biogeography, and Coevolution of a Special Group Restricted in East Asia. Mol Phylogenet Evol 18(2):163–173
16. Sivasundar A, Bermingham E, Ortí G (2001) Population structure and biogeography of migratory freshwater fishes (Prochilodus: Characiformes) in major South American rivers. Mol Ecol 10(2):407–417
17. Hall TA (1999, January) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In Nucleic acids symposium series (Vol. 41, No. 41, pp. 95–98). [London]: Information Retrieval Ltd., c1979-c2000
18. Swofford D L (2001) Paup*: Phylogenetic analysis using parsimony (and other methods) 4.0. B5
19. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6):1547–1549
20. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11):1451–1452
21. Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16(1):37–48
22. Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 152(3):1079–1089
23. Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mole Ecol Res 10(3):564–567
24. Mantel N (1967) The detection of disease clustering and a generalized regression approach. Can Res 27(2):209–220
25. Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel Population genetic software for teaching and research. Mol Ecol Notes 6(1):288–295
26. Corander J, Marttinen P, Sirén J, Tang J (2013) BAPS: Bayesian analysis of population structure. Department of Mathematics and statistics, University of Helsinki, Finland, Manual, p 6
27. Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. Mol Biol Evol 20(1):76–86
28. Hobbs JPA, Van Herwerden L, Jerry DR, Jones GP, Munday PL (2013) High genetic diversity in geographically remote populations of endemic and widespread coral reef angelfishes (genus: Centropyge). Diversity 5(1):39–50
29. Avise JC, Neigel JE, Arnold J (1984) Demographic influences on mitochondrial DNA lineage survivorship in animal populations. J Mol Evol 20(2):99–105
30. Wibowo A, Affandi R, Soewardi K, Sudarto S (2017) Genetic Differentiation of the Kampar river's Giant Featherback (*Chitala lopis* Bleeker 1851) Base on Mitochondrial DNA Analysis. Indonesian Fisheries Research Journal 16(2):49–58
31. Takagi AP, Ishikawa S, Nao T, Song SL, Hort S, Thammavong K, Saphakdy B, Phomsouvanhm A, Nishida M, Kurokura H (2010) Genetic differentiation and distribution routes of the bronze featherback *Notopterus notopterus* (Osteoglossiformes: Notopteridae) in Indochina. Biol J Lin Soc 101(3):575–582
32. Menon AGK (1951) Further studies regarding Hora's Satpura hypothesis. The Role of Eastern Ghats in distribution of the Malayan flora and fauna to Peninsular India. Proceedings of the national Academy of Sciences. India. 17:475–497
33. Kim M, Morrison M, Yu Z (2011) Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. J Microbiol Methods 84(1):81–87
34. Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129(2):555–562
35. Frankham R, Ballou JD, Briscoe DA (2004) A primer of conservation genetics. Cambridge University Press, United Kingdom, p 220
36. Delrieu-Trottin E, Maynard J, Planes S (2014) Endemic and widespread coral reef fishes have similar mitochondrial genetic diversity. Proceedings of the Royal Society B: Biological Sciences 281(1797):20141068
37. Hartl DL, Clark AG, Clark AG (1997) Principles of population genetics, vol 116. Sinauer associates, Sunderland, MA
38. Sodsuk PK, Sodsuk, (2000) Gentic diversity of featherback fish in Thailand. Agriculture and Natural resources 34(2):227–239
39. Barby FF, Ráb P, Lavoué S, Ezaz T, Bertollo LAC, Kilian A, Maruyama SR, Aguiar de Oliveira E, Artoni RF, Santos MH, Ilesanmi Jegede O (2018) From chromosomes to genome: insights into the evolutionary relationships and biogeography of Old World knifefishes (Notopteridae; Osteoglossiformes). Genes 9(6):306

40. Daniels RR (2001) Endemic fishes of the Western Ghats and the Satpura hypothesis. Current Science-Bangalore 81(3):240–244
41. Chauhan T, Lal KK, Mohindra V, Singh RK, Punia P, Gopalakrishnan A, Sharma PC, Lakra WS (2007) Evaluating genetic differentiation in wild populations of the Indian major carp, *Cirrhinus mrigala* (Hamilton–Buchanan, 1882): evidence from allozyme and microsatellite markers. Aquaculture 269(1–4):135–149
42. Acharya AP, Pavan-Kumar A, Gireesh-Babu P, Joshi CG, Chaudhari A, Krishna G (2019) Population genetics of Indian giant river-catfish, *Sperata seenghala* (Sykes, 1839) using microsatellite markers. Aquat Living Resour 32:4
43. Khedkar GD, Jamdade R, Kalyankar A, Tiknaik A, Ron TB, Haymer D (2014) Genetic fragmentation in India's third longest river system, the Narmada. Springer Plus 3(1):385
44. Bhandari S, Maurya DM, Chamyal LS (2005) Late Pleistocene alluvial plain sedimentation in lower Narmada Valley, Western India: palaeoenvironmental implications. J Asian Earth Sci 24(4):433–444
45. Briggs JC (2003) The biogeographic and tectonic history of India. J Biogeogr 30(3):381–388
46. Martinez AS, Willoughby JR, Christie MR (2018) Genetic diversity in fishes is influenced by habitat type and life-history variation. Ecology and Evolution 8(23):12022–12031
47. Chandan R, Singh RK, Singh A, Paul P, Sah RS, Kumar R, Mohindra V, Lal KK, Jena JK (2020) Spatio-temporal variations in length-weight relationship and condition factor of two notopterids, Chitala chitala (Hamilton, 1822) and Notopterus notopterus (Pallas, 1769). Indian J Fish 67(2):120–124
48. Verma SR, Dalela RC (1975) Studies on the Pollution of the Kalinadi by Industrial Wastes near Mansurpur Part 2: Biological Index of Pollution and Biological Characteristics of the River. Actahydrochimica et hydrobiologica 3(3):259–274
49. Tran NN, Thanh VC, Hirata Y (2005) Genetic diversity and conservation of the threatened freshwater Clown Knifefish Notopterus chitala through artificial propagation in the Mekong Delta. Biosphere conservation: for nature, wildlife, and humans 6(2):61–66
50. Vass K, Das M, Tyagi R, Katiha P, Samanta S, Srivastava N, Bhattacharjya B, Suresh V, Pathak V, Chandra G, Debnath D (2011) Strategies for sustainable fisheries in the Indian part of the Ganga-Brahmaputra river basins. International Journal of Ecology and Environmental Sciences 37(4):157–218