Current Zoology 60 (4): 460–471, 2014

Genealogy and phylogeography of Cyprinid fish *Labeo rohita* (Hamilton, 1822) inferred from ATPase 6 and 8 mitochondrial DNA gene analysis

Rupesh K. LUHARIYA[1], Kuldeep K. LAL[1*], Rajeev K. SINGH[1], Vindhya MOHINDRA[1], Arti GUPTA[1], Prachi MASIH[1], Arvind K. DWIVEDI[1], Rakhi DAS[1], U. K. CHAUHAN[2], J. K. JENA[1]

[1] National Bureau of Fish Genetic Resources (ICAR), Lucknow- 226 002, UP, India

[2] School of Environmental Sciences, APS University, Rewa – 486 003, MP, India

Abstract ATPase 6/8 gene (842 bp) of mitochondrial DNA was sequenced in *Labeo rohita* samples (n = 253) collected from nine rivers belonging to four river basins; Indus, Ganges, Brahmaputra and Mahanadi. Analysis revealed 44 haplotypes with high haplotype diversity (Hd) 0.694 and low nucleotide diversity (π) 0.001. The within population variation was larger (83.44%) than among population differences (16.56%). The mean F_{ST} value (0.166; $P < 0.05$) for overall populations revealed moderate level of genetic structuring in the wild *L. rohita* populations. The haplotype network presented a single clade for wild *L. rohita* population, from different rivers. Negative values for Fu's index (F_S), mismatch distribution analysis indicated period of expansion in *L. rohita* population. The time after recent expansion was estimated for each population, between 0.042 to 0.167 mya. The pattern of Isolation by Distance (IBD) was not significant (r = -0.113, $P < 0.287$), when all the sampling locations were compared (Mantel test), however, when an outlier (Indus, Brahmaputra and Mahanadi) was removed from the whole population set, a clear positive correlation between pairwise F_{ST} and geographic distance (Km) was seen. The analysis of data demonstrated that ATPase6/8 gene polymorphism is a potential marker to understand genetic population structure of wild *L. rohita* existing in different rivers. The study identified population substructure in wild *L. rohita* with common ancestral origin [*Current Zoology* 60 (4): 460–471, 2014].

Keywords *Labeo rohita*, ATPase6/8, mtDNA, Polymorphism, Genetic Divergence, India

Labeo rohita (Hamilton, 1822), commonly known as rohu, (Order- Cypriniformes, Family- Cyprinidae) is the most important among the three Indian major carps used in Indian carp polyculture systems. This species is the natural inhabitant of Indo-Gangetic riverine system spread across northern and central India, and the rivers of Pakistan, Bangladesh, Nepal and Myanmar (Reddy, 1999). It has been successfully transplanted out of its natural range within India and parts of Asia as well as Europe. Aquaculture production of IMC was estimated to 4% of the world production, where as in India the production of *L. rohita* was 9, 45, 233mt. *Labeo rohita*, assessed as Least Concern, is a widespread species cultured in captivity throughout India and adjacent countries (IUCN, 2010[1]), but like many other cultured cyprinids (*Catla catla, Cirrhinus mrigala*) wild population of *L. rohita* are at risk of loss of genetic diversity and variability due to extinction of genetically distinct wild stocks and mixing with the farmed accidental escapes or reservoir stock programmes (Reddy, 2005; FAO, 2006). Given the occurrence of this event, over a course of period and generations, the introgression may ultimately render the species lesser fit to adapt to the changing environmental conditions. Considering its importance in the culture system, emphasis has also been given to its genetic improvement through selective breeding in India. Natural genetic resources form the basis for selection of the founder stocks for the selection programmes. Therefore, the genetic diversity data could have a vital role in scientific planning of the breeding programs for conservation and effective management of its natural genetic diversity.

Mitochondrial DNA is a genetic material that exists outside the nucleus in eukaryotic cells, which is haploid genome and mostly inherited maternally. On account of its lower effective population size and rapid rate of evolution, various genes of mtDNA genome are being used for investigating different issues (Suneetha et al.,

Received July 26, 2013; accepted Oct. 25, 2013.

∗ Corresponding author. E-mail: kulvin100@yahoo.co.in

[1] IUCN Red List of Threatened Species. Version 2010.2. Downloaded in August 2010.

© 2014 *Current Zoology*

2000; Apostolidis et al., 2008; Wu et al., 2010; He et al., 2011; Mandal et al., 2011). Different parts of the mitochondrial gene are known to evolve at different rates (Meyer, 1993). ATPase 8 and ATPase 6 genes of mtDNA are generally variable in vertebrates (Zardoya et al., 1996). These genes have been consistently found to have high evolutionary rate (1.3% per million years) in fishes (Bermingham et al., 1997). ATPase 8 and ATPase 6 genes have been potentially used for analyzing both phylogeny as well as phylogeography in several fish species, (Chow and Ushiama, 2004; Hurwood et al., 2008; Dammannagoda et al., 2008; Vergara et al., 2009). ATPase 8 and ATPase 6 mtDNA gene in different cyprinids have been used as genetic markers for monitoring variation in progeny of crosses (Xin et al., 2004). A recent study by Yan, (Yan et al., 2009) also suggested that ATPase 6/8 gene is a valuable genetic marker to track genealogies and variations in progenies of hybrids. Besides, it has also been used in studying genetic analysis based population structure investigation. No published information is yet available on the population structure analysis of *L. rohita* using ATPase region of mitochondrial genome, however, data on various mitochondrial genes for evaluation of phylogenetic implications and molecular identification for few other fishes across taxonomic orders such as, Acipenceriformes (Fontana et al., 2007); Squaliformes (Murray et al 2008); Salmoniformes (Oleinik et al., 2007; Bouza et al., 2008) are available. In a previous investigation on *L. rohita*, low to moderate genetic differentiation in natural population was assessed using mitochondrial cytochrome *b* region (Luhariya et al., 2011). Generally, the fragmented populations of freshwater species are expected to have high levels of genetic differentiation (Habib et al., 2012; Ward et al., 1994; Habib et al., 2010). Such extent of differentiation is the net consequence of several interactive evolutionary forces such as restricted or lack of gene flow and other random genetic drift, mutations that acted on the population during the time lapsed after fragmentation (Hartle and Clark, 1997). Therefore, exploration with alternate polymorphic mtDNA marker was deemed necessary to establish suitability for application in the genetic stock identification programme for *L. rohita* in its vast native distribution range.

This paper is contribution to the demographic history and phylogeographic knowledge based on the polymorphism in mitochondrial ATPase 6/8 gene and analyses distribution pattern of genetic variation in large sample representative of populations of *L. rohita* in India. These results are used to derive information on genetic stock structure of the wild *L. rohita.* The importance of the results in conservation and management wild genetic resources of *L. rohita* are discussed.

1 Material and Methods

1.1 Sample collection

A total of 253 tissue samples of *L. rohita* were collected from commercial riverine catches from four different river basins, include Indus, Ganges, Brahmaputra and East Coast river system. Total nine rivers were chosen for the study and the samples collected from eleven different collection sites, located in different geographical areas in India i.e. Satluj, Brahmaputra, Mahanadi, Rapti, Son, Chambal, Tons, Chauka and Bhagirathi (Table 1). Sampling sites were selected to document genetic variation across a wide geographical distribution

Table 1 Details of sampling locations and sample size of *Labeo rohita* collected for the study

River Basin	River	Location/ Collection sites	Longitude & Latitude	*n*
Indus	Satluj	Harike patan, Punjab	31°13′N, 75°12′E	24
Brahmaputra	Brahmaputra	Kalangpar, Assam	26°11′N, 91°47′E	23
Mahanadi	Mahanadi	Cuttack, Orissa	20°27′N, 85°52′E	22
Ganges	Rapti	Gorakhpur, U.P.	26°13′N, 83°10′E	21
	Son	Beohari, M.P.	24°45′N, 81°85′E	25
	Tons Chakghat	Chakghat, M.P.	25°06′N, 81°45′E	22
	Tons Rewa	Rewa, M.P.	24°32′N, 81°18′E	21
	Chauka	Mahmudabad U.P.	27°21′N, 81°23′E	25
	Bhagirathi	Nabadeep, West Bengal	23°24′N, 88°23′E	24
Yamuna	Chambal GWL	Gwalior, M.P.	26°13′N, 78°10′E	21
	Chambal Kota	Kota, Rajasthan	25°11′N, 75°50′E	25
Total				253

n: number of samples studied

range (31°13′N, 75°12′E to 21°11′N, 91°47′E). The river Satluj belongs to the Indus basin and river Brahmaputra which originates from South- western Tibet and flows southwest through the valley of Assam; river Mahanadi which is an independent river originating from central plateau in India and draining into the Bay of Bengal; both Mahanadi and Son originates near Amarkantak mountain ranges in central India while river Bhagirathi is the lower stretch of Ganga. The rivers Chambal, Son, Tons, Chauka, Rapti, and Bhagirathi are distant tributaries of Ganges (ECAFE, 1966). Weights of specimens ranged from 1.2 to 9.5 kg. Sampling procedures were performed at actual site of collection. The blood was extracted through caudal puncture and fixed in 95% ethanol in 1:5 (blood: ethanol) ratio.

1.2 DNA extraction and PCR amplification

Total Genomic DNA was extracted from blood using the Phenol-Chloroform method, protocol modified by Ruzzante et al. (1996). ATPase8 and ATPase6 gene fragment was amplified using universal primers ATP8.2 L8331 and COIII.2H9236 (Sivasundar et al., 2001) in 50 μl reaction, reaction volumes containing 1X reaction buffer (10 mM Tris, 50 mM KCl, 0.01% gelatine, pH 9.0), 1.5 mM $MgCl_2$, 200 μl of each dNTP's, 3U Taq polymerase (Genei, India), 5 pmoles of each primer and approximately 30–50 ng of template DNA. The amplification consisted of 30 cycle with an initial denaturation at 94°C for 300 secs, denaturation at 94°C for 30 secs; annealing at 55°C for 60 secs and extension at 72°C for 90 secs per cycle and final extension at 72°C for 600 secs. The amplicons were purified and sequenced bidirectionally on ABI sequencer using machine protocol.

1.3 Genetic diversity analysis

Amplified ATPase6/8 genes were sequenced in both the directions to check the validity of the sequence data. All DNA sequences were aligned using ClustalW and were further analysed for determining parameters of population genetic variation. MEGA 4.1 (Tamura et al., 2007) was used to estimate parameters of genetic variation parameters. Haplotype diversity and nucleotide diversity were estimated using DnaSP 4.5 (Rozas et al., 2003). Sequence composition, molecular diversity indices, genetic differentiation and FST values were calculated using Arlequin 3.11 (Excoffier et al., 2005). A minimum-spanning haplotype network was estimated using the TCS program (Clement et al., 2000), which implements the statistical parsimony method of Templeton (Templeton et al., 1995).

To test the patterns of isolation by distance (IBD) and effects of the reservoir on genetic distance among populations by using Mantel test (Mantel, 1967; Slatkin, 1993) with 10, 000 permutations in XL-STAT implemented via the software XLSTAT 2010. The genetic distance matrix used was pairwise F_{ST} between populations while geographical distances in Kilometer (Km) between populations were based on stream segment and reservoir distances.

1.4 Neutrality and demographic history

Both mismatch analysis and neutrality tests were performed using the software Arlequin 3.11 (Excoffier et al., 2005). The mismatch analysis in Arlequin 3.11 (Excoffier et al., 2005) included a raggedness index to determine goodness of fit to a unimodal distribution. Two widely used statistical tests were employed: Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997), Tajima's D test compares two estimators of the mutation parameter θ, Watterson's estimator θs and Tajima's estimator θπ, significant D values can be estimated due to factors such as selection, population expansion, and bottleneck (Tajima, 1989). As *S* depends more on the present population size and k on the size of the original population, a history of population growth can inflate S significantly compared with k and generate a negative value of Tajima's D (Tajima, 1989). Fu's Fs test is constructed based on selective neutrality using the probability of the number of alleles in a sample. Fu (1997) found that the Fs are sensitive to population demographic expansions, which generally lead to large negative Fs values. Historic demographic expansions were also investigated by examination of frequency distributions of pairwise differences between sequences (mismatch distribution), which was based on three parameters: θ_0, θ_1 (before and after the population growth) and τ (time since expansion expressed in units of mutational time (Rogers and Harpending, 1992). The mismatch distribution is usually multimodal in samples drawn from populations at demographic equilibrium, but it is usually unimodal in populations following a recent population demographic expansion and range expansion (Rogers and Harpending, 1992; Slatkin and Hudson, 1991; Excoffier, 2004). The parameters of the demographic expansion τ, θ_0 and θ_1 are estimated by a generalized non-linear least-square approach, and confidence intervals of the parameters are computed using a parametric bootstrap approach (Schneider and Excoffier, 1999). The values of τ were transformed to estimate of real time since expansion with the equation $\tau = 2ut$, where τ is the mode of the mismatch distribution, u is the mutation rate of the sequence considering that $u = 2\ \mu k$ (μ is the mutation rate

per nucleotide and k is the number of nucleotides in the sequence analysed). Thereafter, t in years was calculated as time after expansion using generation time of 2.94 years for *L. rohita* (Froese and Pauly, 2013[2]). A substitution rate of 1.3% per million years (Myr) was used for ATPase which is reported as the mean rate for vertebrate mtDNA (Bermingham et al., 1997).

2 Results

2.1 Sequence composition

Total 842 bp of ATPase 8 and ATPase 6 mitochondrial gene was sequenced in 253 individuals from eleven different populations belonging to nine different rivers (four different river basins) to determine the genetic variability in wild *L. rohita*. ATPase 8 region spanned from 1–165 bp of the sequence and ATPase 6 from 159–842 bp. An overlapping region (7 bp) between two genes was found from 159–165 bp. The two regions have been analysed together for determining variation in *L. rohita* in this study. We identified 36 variable positions with 44 haplotypes including 28 parsimony informative sites (Table 2). The average frequencies of four nucleotides for all the samples of *L. rohita* are A: 32.00%; T: 26.70%; C: 29.00%, G: 12.30%; Nucleotide sequences of ATPase6/8 were A+T rich (58.70%) with transition to transversion ratio (Ts: Tv) was 5.324. GenBank accession numbers of the 44 haplotypes observed for mtDNA ATPase 6/8 sequences are from KF365258-KF365301.

2.2 Nucleotide and Haplotype Diversity

The average nucleotide diversity (pi) for the all samples from eleven populations was found to be 0.001 while haplotype diversity (Hd) was found 0.694 (variance 0.0012±0.034). Haplotype diversity was found to be high ranging from 0.471 (Rapti) to 0.857 (Son) and the nucleotide diversity was low, ranged between river 0.0007 (Satluj, Mahanadi, Rapti) to 0.0028 (Son) (Table 3).

2.3 Genetic relationship among haplotypes

Haplotype network based on nucleotide divergences among the haplotypes detected in this study indicated that the most of the haplotypes were closely related, with the most common haplotype (h01) as the center of radiation (Fig. 1). Haplotype h01 was the dominant and the most common haplotype of these eleven populations. Total shared haplotypes was three; haplotype h03 was shared between river Satluj and Brahmaputra, haplotype h14 was shared between river Son and Tons Rewa and the haplotype h32 was shared between three different rivers: Tons Chakghat, Tons Rewa and Bhagirathi. Haplotype network also depicted the formation of a single clade and all the haplotypes originated from the haplotype h01 either directly or through subsequent mutations (Fig. 1). Haplotypes from the other populations like Mahanadi, Rapti, Chauka, Chambal Gwalior and Chambal Kota were exclusive and they exhibited population specific haplotypes. River Son significantly diverged from the other populations because within Son, diversity is maximum with maximum number of haplotypes which was 10, followed by river Brahmaputra and Chauka both, with 5 haplotypes each. River Chambal Gwalior, Chambal Kota, Tons Chakghat and Bhagirathi had 4 haplotypes in each. River Satluj, Rapti and Tons Rewa had 3 haplotypes each and 2 haplotypes observed in river Mahanadi (Table 2).

2.4 Genetic differentiation

Analysis of Molecular Variance (AMOVA) of all the eleven populations (Satluj, Brahmaputra, Mahanadi, Rapti, Son, Chambal Gwl, Chambal Kota, Tons Chakghat, Tons Rewa, Chauka and Bhagirathi) revealed that out of total variation, only 16.56% was contributed variation among population and 83.44% was contributed variation to within populations and the F_{ST} value was found to be significant 0.166 (Table 4). Population pair wise F_{ST} values ranged from (0.000) to (0.199) (Table 5). The mean diversity for the entire population was 0.002 and the coefficient of differentiation for all eleven populations was 0.156. The mean distance within groups for all population was ranged (0.000) (Satluj and Tons Rewa) to (0.003) Son.

2.5 Isolation by Distance and Mantel test

The pattern of isolation by distance (IBD) was not supported when all the sampling locations were compared by Mantel test, the IBD observed was non-significant r = -0.113, P < 0.287 (Fig. 2A). IBD was also not found among Indus, Brahmaputra and Mahanadi river basin but it was found significant within Ganga basin, r = 0.409, P < 0.042 (Fig. 2B), Indus and Ganga basin r = 0.406, P < 0.001 (Fig. 2C), Brahmaputra and Ganga basin r = 0.406, P < 0.001 (Fig. 2D), Mahanadi and Ganga basin r = 0.406, P < 0.001 (Fig. 2E). When an outlier (Indus, Brahmaputra and Mahanadi) was removed from the whole population set, a clear positive correlation between pairwise F_{ST} and geographic distance was seen (Table 6).

2.6 Demographic history and Neutrality

Pair-wise mismatch distribution and results of Taji-ma's D-test and Fu's Fs tests were performed for AT-

[2] Froese R, Pauly D, 2013. FishBase. Available from http://www.fishbase.org, on 16 October 2013.

Table 2 Relative haplotype frequencies between eleven populations of *Labeo rohita*

Haplotype	Satluj (24)	Brahmaputra (23)	Mahanadi (22)	Rapti (21)	Son (25)	Chambal Gwalior (21)	Chambal Kota (25)	Tons Chakghat (22)	Tons Rewa (21)	Chauka (25)	Bhagirathi (24)	GenBank Accession Numbers ↓
h01	0.624	0.522	0.636	0.714	0.36	0.524	0.48	0.636	0.524	0.56	0.5	KF365258
h02	0.292	0	0	0	0	0	0	0	0	0	0	KF365259
h03	**0.042**	**0.13**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	KF365260
h04	0.042	0	0	0	0	0	0	0	0	0	0	KF365261
h05	0	0.087	0	0	0	0	0	0	0	0	0	KF365262
h06	0	0.13	0	0	0	0	0	0	0	0	0	KF365263
h07	0	0.087	0	0	0	0	0	0	0	0	0	KF365264
h08	0	0.043	0	0	0	0	0	0	0	0	0	KF365265
h09	0	0	0.182	0	0	0	0	0	0	0	0	KF365266
h10	0	0	0.182	0	0	0	0	0	0	0	0	KF365267
h11	0	0	0	0.19	0	0	0	0	0	0	0	KF365268
h12	0	0	0	0.048	0	0	0	0	0	0	0	KF365269
h13	0	0	0	0.048	0	0	0	0	0	0	0	KF365270
h14	**0**	**0**	**0**	**0**	**0.08**	**0**	**0**	**0**	**0.048**	**0**	**0**	KF365271
h15	0	0	0	0	0.08	0	0	0	0	0	0	KF365272
h16	0	0	0	0	0.12	0	0	0	0	0	0	KF365273
h17	0	0	0	0	0.08	0	0	0	0	0	0	KF365274
h18	0	0	0	0	0.04	0	0	0	0	0	0	KF365275
h19	0	0	0	0	0.04	0	0	0	0	0	0	KF365276
h20	0	0	0	0	0.04	0	0	0	0	0	0	KF365277
h21	0	0	0	0	0.04	0	0	0	0	0	0	KF365278
h22	0	0	0	0	0.08	0	0	0	0	0	0	KF365279
h23	0	0	0	0	0.04	0	0	0	0	0	0	KF365280
h24	0	0	0	0	0	0.237	0	0	0	0	0	KF365281
h25	0	0	0	0	0	0.048	0	0	0	0	0	KF365282
h26	0	0	0	0	0	0.143	0	0	0	0	0	KF365283
h27	0	0	0	0	0	0.048	0	0	0	0	0	KF365284
h28	0	0	0	0	0	0	0.16	0	0	0	0	KF365285
h29	0	0	0	0	0	0	0.2	0	0	0	0	KF365286
h30	0	0	0	0	0	0	0.04	0	0	0	0	KF365287

To be continue

Continued Table 2

Haplotype	Satluj (24)	Brahmaputra (23)	Mahanadi (22)	Rapti (21)	Son (25)	Chambal Gwalior (21)	Chambal Kota (25)	Tons Chakghat (22)	Tons Rewa (21)	Chauka (25)	Bhagirathi (24)	GenBank Accession Numbers
h31	0	0	0	0	0	0	0.12	0	0	0	0	KF365288
h32	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0.091**	**0.285**	**0**	**0.042**	KF365289
h33	0	0	0	0	0	0	0	0.045	0	0	0	KF365290
h34	0	0	0	0	0	0	0	0.182	0	0	0	KF365291
h35	0	0	0	0	0	0	0	0.046	0	0	0	KF365292
h36	0	0	0	0	0	0	0	0	0.143	0	0	KF365293
h37	0	0	0	0	0	0	0	0	0	0.16	0	KF365294
h38	0	0	0	0	0	0	0	0	0	0.04	0	KF365295
h39	0	0	0	0	0	0	0	0	0	0.16	0	KF365296
h40	0	0	0	0	0	0	0	0	0	0.04	0	KF365297
h41	0	0	0	0	0	0	0	0	0	0.04	0	KF365298
h42	0	0	0	0	0	0	0	0	0	0	0.292	KF365299
h43	0	0	0	0	0	0	0	0	0	0	0.124	KF365300
h44	0	0	0	0	0	0	0	0	0	0	0.042	KF365301

Table 3 Intra-population haplotype diversities (*h*) and Nucleotide diversities (π) for ATPase6/8 region from eleven different populations of *Labeo rohita*

Population	Haplotype Diversity (*h*)	Variance of *h* with standard deviation	Nucleotide Diversity (π)
Satluj	0.543	0.007±0.085	0.0007
Brahmaputra	0.708	0.008±0.090	0.0012
Mahanadi	0.554	0.009±0.097	0.0007
Rapti	0.471	0.013±0.116	0.0007
Son	0.857	0.003±0.059	0.0028
Chambal Gwl	0.676	0.007±0.85	0.0010
Chambal Kota	0.717	0.005±0.070	0.0013
Tons Chakghat	0.576	0.011±0.108	0.0008
Tons Rewa	0.652	0.006± 0.077	0.0009
Chauka	0.657	0.008±0.090	0.0012
Bhagirathi	0.674	0.005±0.071	0.0013

Pase 6/8 region on each population are given in Fig.3. The parameters of the sudden expansion and goodness of fit test to the model are given as in Table 7. All histograms presented unimodal curves (P<0.05) characteristics that populations have passed through a recent demographic expansion. Populations from Satluj, Brahmaputra, Chambal Kota, Tons Chakghat, Tons Rewa and Bhagirathi exhibited moderate to high negative values for both Tajima's D-statistic and Fu's Fs. The populations from Brahmaputra, Son and Tons Chakghat showed significant values (P<0.05). The time after recent expansion was estimated for each population, between 0.042 to 0.167 mya (Table 7).

3 Discussion

Variation in ATPase6/8 gene (842 bp) sequences of mtDNA and resulting haplotypes revealed genetic structuring in wild *L. rohita* population, with overall moderate level of genetic differentiation between the subpopulations but high variation within the subpopulation. This indicates that gene flow that can offset the genetic differentiation among subpopulation is possibly happening due to direct or indirect continuity across the rivers and their tributaries of Indo-Gangetic plain. Substantial haplotype diversity with a total of 44 haplotypes was recorded in *L. rohita* populations and number of population specific haplotypes ranged from 2 (river Mahanadi) to 9 (river Son).

3.1 Population differentiation and genetic structure

Present study revealed the monophyletic origin of *L. rohita* population present in the nine rivers representing four independent river basins and derived from a com-

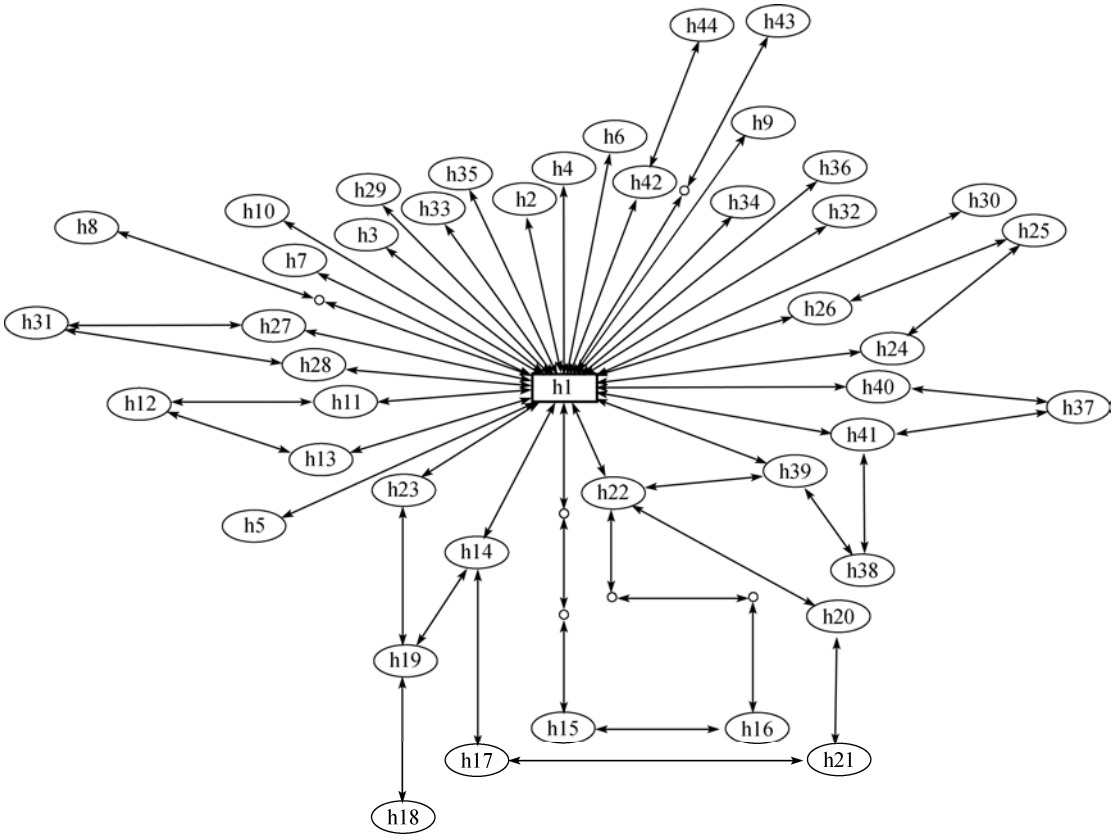

Fig. 1 Haplotype network obtained between eleven different populations of *Labeo rohita*

Table 4 Two hierarchical AMOVA of ATPase6/8 sequence varaition in *Labeo rohita* samples. The Fixation Index and the *P*-values was calculated with variance components using ARLEQUIN 3.11 at 1000 random permutations of the data matrix

Source of variation	Sum of squares	Variance components	% of variation	Fixation Index	*P* value
Among populations	27.416	0.097 Va	16.56%	F_{ST} : 0.166	<0.0001
Within populations	119.288	0.492 Vb	83.44%		
Total	146.704	0.590			

Table 5 Population pair wise F_{ST} (below diagonal), population specific F_{ST} (at diagonal) and P values (above diagonal) between *Labeo rohita* samples collected from different riverine locations

Population	Satluj	Brahmaputra	Mahanadi	Rapti	Son	Chambal Gwl	Chambal Kota	Tons Chakghat	Tons Rewa	Chauka	Bhagirathi
Satluj	0.180	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Brahmaputra	0.110*	0.166	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mahanadi	0.173*	0.094*	0.179	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Rapti	0.181*	0.096*	0.149*	0.181	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Son	0.196*	0.159*	0.179*	0.179*	0.114	0.000	0.000	0.000	0.000	0.000	0.000
Chambal Gwl	0.198*	0.125*	0.172*	0.178*	0.188*	0.171	0.000	0.000	0.000	0.000	0.000
Chambal Kota	0.185*	0.124*	0.161*	0.165*	0.192*	0.173*	0.163	0.000	0.000	0.000	0.000
Tons Chakghat	0.144*	0.069*	0.113*	0.118*	0.168*	0.147*	0.140*	0.178	0.036	0.000	0.000
Tons Rewa	0.192*	0.117*	0.164*	0.171*	0.173*	0.186*	0.175*	0.082*	0.173	0.000	0.000
Chauka	0.189*	0.127*	0.165*	0.168*	0.168*	0.184*	0.180*	0.144*	0.179*	0.163	0.000
Bhagirathi	0.191*	0.130*	0.167*	0.171*	0.192*	0.186*	0.181*	0.140*	0.163*	0.183*	0.162

* $P<0.05$. Bold values indicate population specific F_{ST}

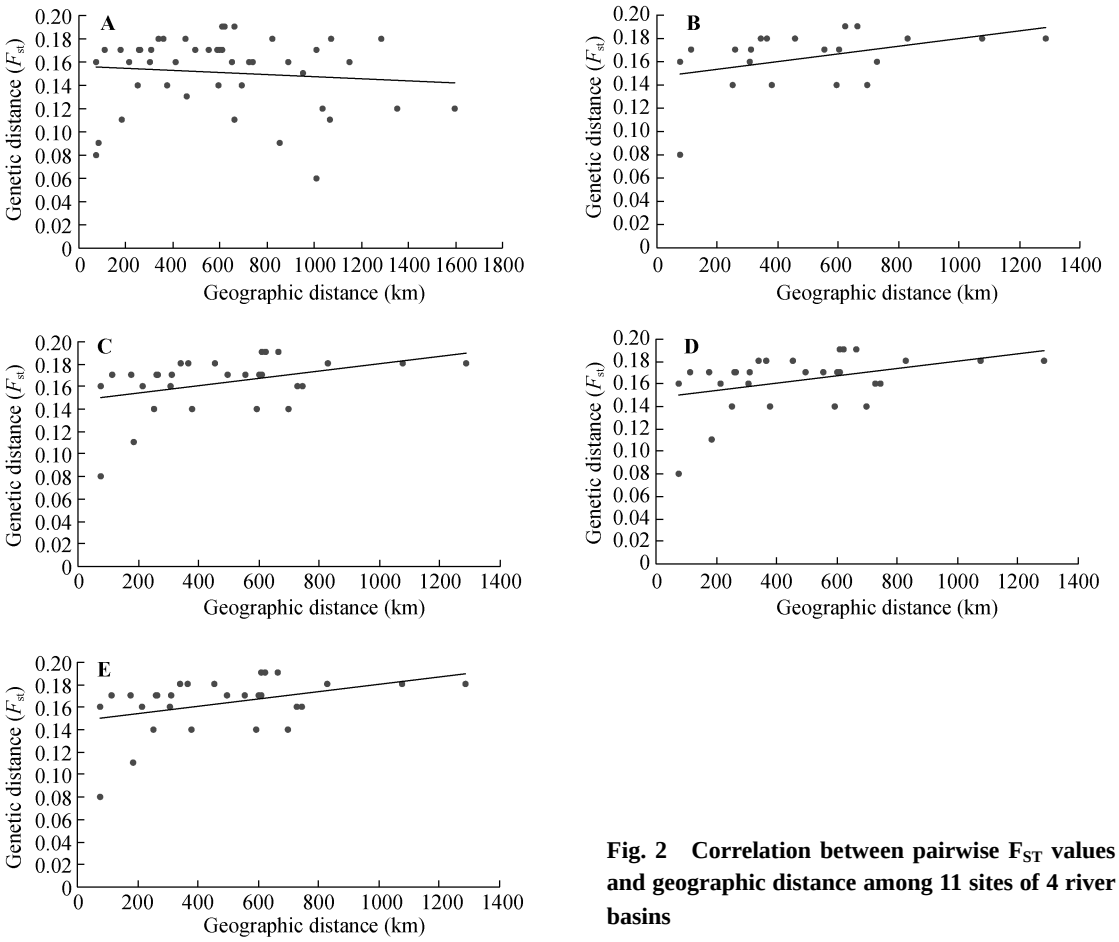

Fig. 2 Correlation between pairwise F_{ST} values and geographic distance among 11 sites of 4 river basins

Table 6 The relationship between Isolation by Distance (IBD) obtained from geographical distance (km) and genetic distance (F_{ST}) between four river basins for *Labeo rohita*. IBD was performed by removing one basin at a time and the significance of the subsequent relationship was determined using Mantel test at 10000 permutations

Populations excluded	R^2	P
Satluj and Brahmaputra	0.162	0.001
Brahmaputra and Mahanadi	0.164	0.001
Satluj and Mahanadi	0.162	0.001
Satluj, Brahmaputra and Mahanadi	0.167	0.042
Mean	0.163	0.011
All populations	0.012	0.287

R^2: Correlation coefficient

mon monomorphic ancestry. The observed pattern and distribution of genetic variation was supported by haplotype network and phylogeographic analysis. Haplotype networks reconstruct the genealogical history of haplotypic variation and illustrate the evolutionary relationship among unique haplotypes. Under coalescent principles, the most common haplotype and that occupy central position in a network are assumed ancestral, while tip haplotypes are considered younger, more recently derived types (Templeton and Sing, 1993; Crandall, 1996). Hence, the haplotype h01, the most dominant in all the populations, must have been ancestral and precursor to other haplotypes including the population exclusive haplotypes. Out of total 43 haplotypes that originated from h01, 28 (65%) haplotypes have mutated directly from h01 and among these 14 (50%) haplotypes are yet to have any subsequent mutation. Such genealogical relationship indicates that *L. rohita* populations in different localities have recently diverged from each other. Population from river Son was significantly different from the other populations as it has the maximum number of haplotypes supports the observation from cytochrome *b* gene analysis of *L. rohita* (Luhariya et al., 2011). Three shared haplotypes, h03, h14 and h32 originated directly or indirectly from the common ancestor through subsequent mutation. Haplotype h03 was shared between Satluj and Brahmaputra, h14 was shared between Son and Tons Rewa and h32 was shared between three rivers; Tons at Chakghat and Rewa and Bhagirathi.

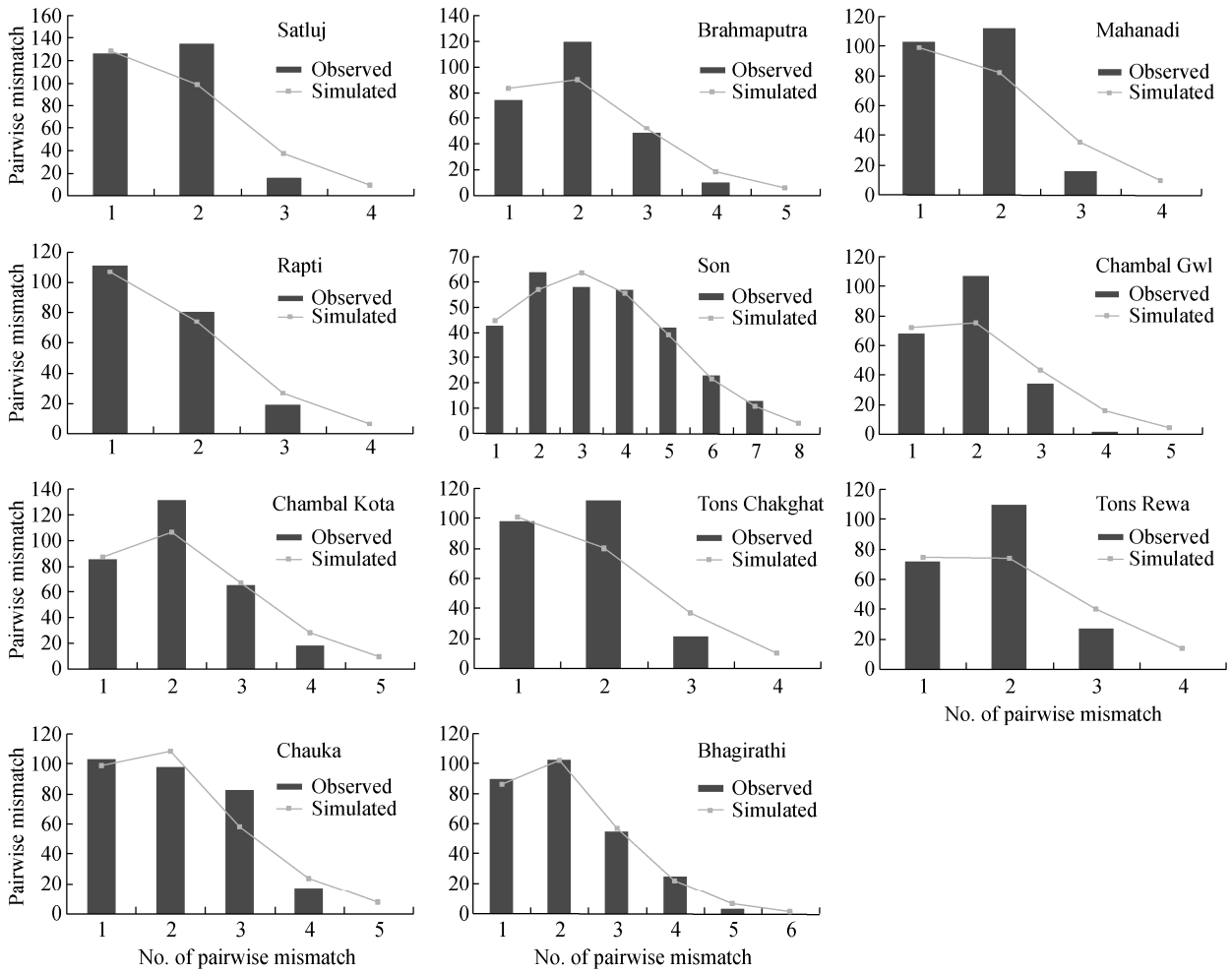

Fig. 3 The observed pairwise differences (bars) and the simulated mismatch distributions (line) under the sudden expansion model of ATPase 6/8 haplotypes in *L. rohita*

Table 7 Summary of diversity, neutrality and expansion time estimates for *L. rohita*, number of haplotypes (H), segregating sites (S), Tajima's *D* and Fu's *Fs*, corresponding *P* values, time in number of generation (τ), time after expansion (t mya) and mismatch distribution parameter estimates (θ_0, pre expansion population size; θ_1 post expansion population size), sum of squared deviation (SSD), Raggedness index (r)

Population	H	*S*	Tajima's D		Fu's *Fs*		Time after expansion event		Mismatch distribution		Goodness of- fit test	
			D	*P**	*Fs*	*P**	τ	t	θ_0	θ_1	SSD	(r)
Satluj	3	3	-0.643	0.306	-0.995	0.170	0.764	0.051	0.000	99999	0.024	0.193
Brahmaputra	5	6	-1.214	0.037	-2.094	0.043	1.121	0.074	0.000	99999	0.014	0.137
Mahanadi	2	2	0.308	0.723	0.285	0.483	0.789	0.052	0.000	99999	0.021	0.179
Rapti	3	2	0.025	0.622	-1.258	0.095	0.639	0.042	0.007	99999	0.002	0.114
Son	10	7	0.867	0.818	-4.056	0.018	2.527	0.167	0.465	7.82	0.001	0.014
Chambal Gwl	4	3	0.043	0.602	-1.480	0.093	1.021	0.068	0.000	99999	0.025	0.180
Chambal Kota	4	4	-0.015	0.514	-0.703	0.307	1.179	0.078	0.002	99999	0.007	0.102
Chakghat Tons	4	4	-1.102	0.042	-2.060	0.020	0.826	0.055	0.005	2.453	0.019	0.167
Tons Rewa	3	3	-0.167	0.419	-0.521	0.301	0.984	0.065	0.000	99999	0.028	0.203
Chauka	6	3	0.789	0.813	-1.768	0.113	1.156	0.077	0.000	21.77	0.000	0.054
Bhagirathi	4	5	-0.545	0.332	-0.670	0.316	1.044	0.069	0.077	99999	0.000	0.050

AMOVA analysis also supported the observed differentiation of haplotypes in *L. rohita* and indicated moderate but significant genetic structuring of the populations. The results reveal that 16.56% of the total variation, was attributed to among population differences and 83.44% was due to within population variation. The significant heterogeneity, evident between all the population pairs appears to be due to the differences in the observed population exclusive haplotypes. The results from ATPase 6/8 genes more or less support the pattern of variation detected in *L. rohita* using cytochrome *b* region of mtDNA (Luhariya et al., 2011).

3.2 Demographic history and population expansion

Our data on *L. rohita* support the pattern of rapid population expansion as evident from negative values of Fu's (Fs) index (Fu, 1997) in almost all the populations except Rapti. These values were more negative and significant in three populations Brahmaputra, Tons chakghat and Son, perhaps indicating stronger expansion signals in these three populations as compared to other populations. The population expansion is also supported by star-like genealogies (Slatkin and Hudson, 1991) in the haplotype network of *L. rohita* and unimodal mismatch distribution patterns for all the populations (Rogers and Harpending, 1992).

Low nucleotide diversity, evident in *L. rohita* data, is characteristic of populations with shallow genetic structure because of rapid lineage sorting between small founder populations (Grant and Bowen 1998). The criteria suggested by Grant and Bowen (1998 indicate that the population of *L. rohita* experienced population bottleneck or founder event. Subsequently, fragmented populations might have undergone sudden expansion and formation of new haplotypes, characterized with low frequencies. The time after expansion for the *L. rohita* populations (0.042 to 0.167 mya) correspond to late Pleistocene periods, when many drainage rearrangements have been taken place (Valdiya, 2002). Range expansion of these carps possibly happened during Eocene till late Pleistocene through westward flowing Indo-Brahma river (Daniel, 2001; Mandal et al., 2011; Chauhan et al., 2007) or with the formation of Indo-Gangetic plains and associated drainages (Valdiya, 2002) in early Pleistocene (1.7 to 1.5 mya). During these period, the river systems such as Son, Tons and Chambal, historically older than formation of Ganga basin, are likely to have become accessible to the migrating Indo-Malayan fish fauna including carps The dismemberment of the river into separate Indus, Ganga and Brahmaputra basins and subsequent drainage rearrangements are likely to have influenced the fragmentation, expansion and consequently differentiation in *L. rohita* population.

As a population becomes more isolated, the potential for stochastic processes to create different haplotypes in the population becomes more likely and analyzed as isolation by distance (Slatkin, 1993). Barluenga and Mayer (2005) suggested two alternative hypothesis to explain the genetic differences from different locations of the rivers (i) river was colonized by a large heterogeneous lineage of fish that differentiated in several rather isolated groups in diverse areas of the river; or (ii) the river was colonized by multiple genetically diverse lineages of fish from different refugia that posteriorly homogenized in the river. In case of *L. rohita* correlation was not obtained when all sampling locations were compared by Mantel test, the IBD was non-significant (r= -0.113, P < 0.287), but when an outlier (Indus, Brahmaputra and Mahanadi) was removed from the whole population set, a clear positive correlation between pairwise F_{ST} and geographic distance was seen. Similar observation was used to assess isolation by distance in coral reef fish (Purcell et al., 2012). In *L. rohita* population, the isolation by distance is evident within the subpopulations in tributaries and rivers of Ganga basin, which forms a continuous system. Genetic relationship, sharing of common haplotypes and high within population variation suggest that the first of the two hypotheses explain the population differentiation in *L. rohita* population.

The small genetic divergence observed within the river Tons is unexpected, however, such differences have been reported in other co-generic carps, *L. calbasu* (Singh et al., 2012) and *L. dero* (Chaturvedi et al., 2011) populations from the same river. This indicate restricted gene flow which is possible due to physical barriers such as large (430 ft) fall at Chachai, (24°47′31″N, 81°18′10″E), separating upstream the Bihad river and downstream Tons river or even the topology of the river (Chaturvedi et al., 2011; Singh et al., 2012).

Reddy (2005) raised the question, if Indian major carps are native to river Mahanadi or represent a naturalized population. Significant frequency (0.636) of ancestral haplotype (h01) in Mahanadi river samples, moderate genetic differentiation levels and comparable to that found in other rivers of Indo-Gangetic plains, leaves no ambiguity that these populations shared a common ancestry and fails to prove these population have fragmented since Pleistocene. The river Mahanadi is flowing southward since Pleistocene and colonized

with Indo-Malayan fishes (including the carps) during migration of fishes from Assam Himalayas to Peninsular India (Silas, 1952). Therefore, genetic evidence suggest that *L. rohita* in river Mahanadi is a naturalized population, a result of a more recent introduction from Indo-Gangetic origin, concordant to the findings for another Indian major carp *Cirrhinus mrigala* (Chauhan et al., 2007).

In summary, there is clear evidence of distinct genetically structure in *L. rohita* population in its native distribution range. The subpopulations harboring different rivers might have fragmented from common ancestry and consequently expanded as evolutionary significant unit. The results recommend caution in culture based capture fishery through stocking farmed fish in rivers and reservoir in Indo-Gangetic and Brahmaputra basin in south Asia, the native distribution range of *L. rohita*. The recent initiatives such as establishment of brood bank in India can utilize the information from this study to plan acquisition of genetically distinct subpopulations. The population genetics data thus generated will have useful application in planning breeding programme for this important aquaculture species and for evolving conservation and management strategies of the wild populations. Further studies involving nuclear markers such as microsatellite DNA and with biological descriptors can enhance insights into genetic stocks of *L. rohita* across its native distribution.

Acknowledgements The present work is the part of the project entitled," Outreach Activity on Fish Genetic Stocks," funded by Indian Council of Agricultural Research, Department of Agricultural Research and Education, Government of India, New Delhi. The authors are thankful to Director, National Bureau of Fish Genetic Resources (NBFGR), Lucknow for facilitating the work.

References

Apostolidis AP, Loukovitis D, Tsigenopoulos CS, 2008. Genetic characterization of brown trout *Salmo trutta* populations from the Southern Balkans using mtDNA sequencing and RFLP analysis. Hydrobiologia 600: 169–176.

Barluenga M, Mayer A, 2005. Old fish in a young lake: Stone loach (Pisces: *Barbatula barbatula*) populations in Lake Constance are genetically isolation by distance. Mol. Ecol. 14: 1229–1239.

Bermingham ES, McCafferty A, Martin A, 1997. Fish biogeography and molecular clocks: perspectives from the Panamian Isthmus. In: Kocher T, Stepien C Ed. Mol. Syst. Fish. New York: Academic Press, 112–126.

Bouza C, Vilas R, Castro J, 2008. Mitochondrial Haplotype variability of brown trout populations from north-western Iberian Peninsula, a secondary contact area between lineages. Conserv. Genet. 9: 917–920.

Chaturvedi A, Mohindra V, Singh RK, Lal KK, Punia P et al., 2011. Population genetic structure and phylogeography of cyprinid fish *Labeo dero* (Hamilton, 1822) inferred from allozyme and microsatellite DNA marker analysis. Mol. Biol. Rep. 38: 3513–3529.

Chauhan T, Lal KK, Mohindra V, Singh RK, Punia P et al., 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: 135–149.

Chow S, Ushiama H. 2004. Global population structure of albacore *Thunnus alalunga* inferred by RFLP analysis of the mitochondrial ATPase gene. Mar. Biol. 123(1): 39–45.

Clement M, Posada D, Crandall KA, 2000. TCS: A computer program to estimate gene genealogies. Mol. Ecol. 9: 1657–1659.

Crandall KA, 1996. Multiple interspecies transmissions of human and simian T-cell leukemia/lymphoma virus type I sequences. Mol. Biol. Evol. 13: 115–131.

Dammannagoda ST, Hurwood DA, Mather PB, 2008. Evidence for fine geographical scale heterogeneity in gene frequencies in yellowfin tuna *Thunnus albacares* from the north Indian Ocean around Sri Lanka. Fisheries Research 90: 147–157.

Daniels RJR, 2001. Endemic fishes of the Western Ghats and the Satpura hypothesis. Current Science 81(3): 240–244

ECAFE, 1966. A Compendium of Major International Rivers in the ECAFE Region. Water Resources Series, United Nation publication no. 29. New York: United Nations.

Excoffier L, 2004 Analytical methods in phylogeography and genetic structure. Mol. Ecol. 13: 727.

Excoffier L, Laval G, Schneider S, 2005. Arlequin (version 3.0): An Integrated Software Package for Population Genetics Data Analysis. Evol. Bioinform. Online. 1: 47–50.

FAO, 2006. Fishstat Plus: Universal Software for Fish Statistics Time Series, Version 2.3 2000. FAO, Fisheries Department, Fisheries Information, Data and Satistical Unit. http://www.fao.org/fi/statist/FISOFT/FISHPLUS.asp.

Fontana F, Conterio F, Gandolfi G, Tagliavini J, Rosenthal H et al., 2007. Mitochondrial DNA sequences of 6 Sturgeon species and Phylogenetic relationships within Acipenseridae. J. Appl. Ichthyol. 15: 17–22.

Fu XY, 1997. Statistical test of neutrality of mutations against population growth, Hitchhiking and Background selection. Genetics 147: 915–925.

Grant WS, Bowen BW, 1998. Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation. J. Heredity 89: 415–426.

Guo XH, Liu SJ, Liu Y, 2004. Evidence for maternal inheritance of mitochondrial DNA in polyploidy fish of crosses by ATPase8 and ATPase6 genes. Acta Zool Sinica 50(3): 408–413.

Habib M, Lakra WS, Mohindra, Lal KK, Punia P et al., 2012. Assessment of ATPase 8 and 6 mtDNA sequences in genetic diversity studies of *Channa marulius* (Channidae: Perciformes). Proc. Natl. Acad. Sci. India Sect. B. Biol. Sci. DOI 10.1007/s40011–012–0061–x.

Habib M, Lakra WS, Mohindra V, Khare P, Barman AS et al., 2010. Evalution of Cytochrome *b* mtDNA sequences in genetic diversity studies of *Channa marulius* (Channidae: Perciformes)

Mol. Biol. Rep. 38: 841–846.

Hartle DL, Clark AG, 1997. Principles of Population Genetics. 4th edn. Sunderland, Massachusetts: Sinauer Associates, Inc..

He A, Luo Y, Yang H, Liu L, Li S, Wang C. 2011. Complete mitochondrial DNA sequences of the Nile tilapia *Oreochromis niloticus* and blue tilapia (Oreochromis aureus): Genome characterization and phylogeny applications. Mol. Biol. Rep. 38(3):2815–2821. Doi: 10.1007/s11033-010-0324-7.

Hurwood DA, Adamson EAS, Mather PB, 2008. Evidence for strong genetic structure in a regionally important, highly vagile cyprinid *Hemicorhynchus lobatus* in the Mekong river. Ecology of Freshwater Fish 17: 273–283.

Luhariya RK, Lal KK, Singh RK, Mohindra V, Punia P et al., 2011. Genetic divergence in wild population of *Labeo rohita* (Hamilton, 1822) from nine Indian rivers, analyzed through mtDNA cytochrome *b* region. Mol. Biol. Rep. 39 (4): 3659–3665.

Mandal A, Mohindra V, Singh RK, Punia P, Singh RK et al., 2011. Mitochondrial DNA variation in natural populations of endangered Indian feather-back fish *Chitala chitala*. Mol. Biol. Rep. 39 (2): 1765–1775.

Mantel NA, 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209–220.

Meyer A, 1993. Evolution of mitochondrial DNA in fishes. In: Hochachka PW, Mommsen TP ed. Biochemistry and Molecular Biology of Fishes: Molecular Biology Frontiers. Vol.2. Amsterdam: Elsevier Science Publishers, 1–38.

Murray WB, Wang JY, Yang SC, Stevens JD, Fisk A et al., 2008. Mitochondrial cytochrome *b* variation in sleeper sharks (Squaliformes: Somniosidae). Mar. Biol. 153: 1015–1022.

Oleinik AG, Skurikhina LA, Brykov VA, 2007. Divergence of *Salvelinus* sp from north eastern Asia based on mitochondrial DNA. Ecol. Freshw. 16: 87–98.

Purcell JFH, Cowen RK, Hughes CR, Williams DA, 2012. Weak genetic structure indicates strong dispersal limits: A tale of two coral reef fish. Proc. R. Soc. B 2006. 273: 1483–1490.

Reddy PVGK, 1999. Genetic Resources of Indian Major Carps. FAO Fisheries Technical Paper. No. 387. Rome: FAO.

Reddy PVGK, 2005. Carp Genetic Resources of India, Carp Genetic Resources for Aquaculture in Asia. Malaysia: World Fish Center Publishing Inc., 39–53.

Rogers AR, Harpending H, 1992. Population growth makes waves in the distribution of pairwise differences. Mol. Biol. Evol. 9: 552–559.

Rozas J, Sanchez-Delbarrio JC, Messeguer X, Rozas R, 2003. DnaSP, DNA polymorphism analysis by the coalescent and other methods. Bioinformatics 19: 2496–2497.

Ruzzante DE, Taggart CT, Cook D, 1996. Spatial and temporal variation in the genetic composition of a larval cod *Gadus morhua* aggregation: Cohort contribution and genetic stability. Can. J. Fish. Aquat. Sci. 53: 2695–2705.

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: 1079–1089.

Silas EG, 1952. Further studies regarding Hora's Satpura Hypothesis. 2: Taxonomic assessment and levels of evolutionary divergences of fishes with the so-called Malayan affinities in Peninsular India. Proc. Natl. Inst. Sci. India 18(5): 423–448.

Singh RK, Lal KK, Mohindra V, Punia P, Sah RS et al., 2012. Genetic diversity of Indian major carp *Labeo calbasu* (Hamilton, 1822) populations inferred from microsatellite loci. Biochem. Syst. Ecol. 44: 307–316.

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: 407–417.

Slatkin M, 1993. Isolation by distance in equilibrium and non-equilibrium population. Evolution 47: 264–279.

Slatkin M, Hudson R, 1991. Pairwise comparison of mitochondrial DNA sequence in stable and exponentially growing populations. Genetics 129: 555–562.

Suneetha KB, Dahle G, Naevdal G, 2000. Analysis of mitochondrial DNA sequences from two Maurolicus taxa: Evidence for separate species. J. Fish Biol. 57: 1605–1609

Tajima F, 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.

Tamura K, Dudley J, Nei M, Kumar S, 2007. MEGA 4: Molecular evolutionary genetics analysis (MEGA) software version 4. Mol. Biol. Evol. 24: 1596–1599.

Templeton AR, Phillips CA, 1995. Separating population structure from population history: A cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander *Ambystoma tigrinum*. Genetics 140: 767–782.

Templeton AR, Sing CF, 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. Genetics 134: 659–669.

Valdiya KS, 2002. Emergence and evolution of Himalaya: Reconstructing history in the light of recent studies. Progress in Physical Geography 26: 360–399.

Vergara-Chen C, Aguirre WE, González-Wangüemert M, Bermingham E, 2009. A mitochondrial DNA based phylogeny of weakfish species of the Cynoscion group (Pisces: Sciaenidae). Molecular Phylogenetics and Evolution 53: 602–607.

Ward RD, Woodwark M, Skinbinski DOF, 1994. A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. J. Fish Biol. 44: 213–232.

Wu X, Wang L, Chen S, Zan R, Xiao H et al., 2010. The complete mitochondrial genomes of two species from Sinocyclocheilus (Cypriniformes: Cyprinidae) and a phylogenetic analysis within Cyprininae. Mol. Biol. Rep. 37(5): 2163–2171.

Yan J, Guo X, Liu S, Xiao J, Liu Z et al., 2009. Maternal inheritance in polyploid fish inferred from mitochondrial ATPase genes analysis. Prog. in Nat. Sci. 19: 693–698.

Zardoya R, Vollmer DM, Craddock C, Streelman JT, Karl S et al., 1996. Evolutionary conservation of microsatellite flanking regions and their use in resolving the phylogeny of cichlid (Pisces: Perciformes). Proc. R. Soc. B. 263: 1589–1598.