Molecular Biology Reports (2024) 51:378
https://doi.org/10.1007/s11033-024-09323-w

ORIGINAL ARTICLE

Genetic diversity, spatial connectivity, and population structure of Asian silurid catfish *Wallago attu* (Bloch and Schneider, 1801) in the Ganga River System: insights from mitochondrial DNA analysis

Ajit Kumar[1] · Naveen Negi[1] · Neha Yadav[1] · Ruchi Badola[1] · Syed Ainul Hussain[1] · Sandeep Kumar Gupta[1]

Received: 7 December 2023 / Accepted: 6 February 2024

Abstract

Background The Ganga River System (GRS) is a biodiversity hotspot, its ecological richness is shaped by a complex geological history. In this study, we examined the genetic diversity, spatial connectivity, and population structure of the Asian Silurid catfish, *Wallago attu*, across seven tributaries of the GRS.

Methods and results We employed three mitochondrial DNA (mtDNA) regions: cytochrome c oxidase subunit *I* (COX*I*), cytochrome *b* (Cyt *b*), and control region (CR). Our comprehensive dataset encompassed 2420 bp of mtDNA, derived from 176 *W. attu* individuals across 19 sampling sites within the seven rivers of GRS. Our findings revealed high gene diversity (Hd:0.99) within *W. attu* populations. Analysis of Molecular Variance (AMOVA) highlighted that maximum genetic variations were attributed within the populations, and the observed genetic differentiation among the seven populations of *W. attu* ranged from low to moderate. Network analysis uncovered the presence of three distinct genetic clades, showing no specific association with seven studied rivers. Bayesian skyline plots provided insights into the demographic history of *W. attu*, suggesting a recent population expansion estimated to have occurred approximately 0.04 million years ago (mya) during the Pleistocene epoch.

Conclusions These results significantly enhance our understanding of the genetic diversity and spatial connectivity of *W. attu*, serving as a vital foundation for developing informed conservation strategies and the sustainable management of this economically valuable resource within the Ganga River System.

Keywords *Wallago attu* · Mitochondrial DNA · Genetic diversity · Population genetic structure · Ganga river system

Introduction

The Ganga River System (GRS) is the world's most populous river basin, which extends over an area of 1,086,000 km^2 and also known as important aquatic biodiversity hotspots. This extensive basin is intricately interwoven with a vast network of tributaries, threatened by anthropogenic pollution and degradation, significantly impacting aquatic biodiversity and their habitats [1]. It also faces the challenge of accommodating the highest human density. Consequently, these ecological disruptions hinder the conservation efforts to preserve the river's delicate ecosystem [2]. To comprehensively assess the implications of the contemporary environmental threats and losses for biodiversity management, harnessing the link between genetics and associated habitat in populations of focal species within the tributaries is crucial. This approach allows for a more insightful evaluation of the impact on the microecosystem and aids in developing targeted conservation strategies. Consequently, the preservation and sustainable management of the freshwater ecosystem, rich biodiversity, and natural habitats within the GRS have emerged as significant concerns in India. This is particularly pertinent under the ambit of the National Mission for Clean Ganga Project, which is undertaking comprehensive measures aimed at reinstating the ecological health of the GRS.

The Asian Silurid catfish, *Wallago attu* is widely distributed in the Indian subcontinent and several Southeast Asian

✉ Sandeep Kumar Gupta
skg@wii.gov.in

[1] Wildlife Institute of India, Dehra Dun, P.O. Box # 18, Chandrabani, Uttarakhand 248001, India

Published online: 01 March 2024

countries including Thailand, Laos, Cambodia, Vietnam, Malaysia, and Indonesia [3]. *W. attu* is a fast-growing catfish that belongs to the Siluridae family and can reach a maximum length of 2 m, weighing more than 45 kg and long-live around 10 years [4, 5]. Observations in the Brahmaputra basin suggest confined breeding of *W. attu* but prolonged till the end of monsoon from June to September [6]. Moreover, Prasad and Desai [7] reported that the absolute fecundity of *W. attu* ranged from 16,565 to 29,883 in the Bhadar reservoir of Gujarat, India. The species exhibited year-round spawning activity, with a peak observed in July and August. This Silurid catfish has been one of the most economically important freshwater resource and a dietary staple in southeast Asia since ancient times because of its large body size and high nutritive value [4]. Since the past few decades, *W. attu* has experienced a sharp decline throughout its range due to overfishing, habitat fragmentation, dam construction and habitat alteration [8, 9], and hence this species is categorized as Vulnerable in the IUCN Redlist.

A study by Zafar et al. [10] indicates decline in genetic diversity and weak genetic structure within *W. attu* population from Jhelum River, Pakistan using five microsatellites' loci. A comprehensive understanding of the genetic background holds immense significance for conserving and managing threatened species. For this purpose, molecular markers have been extensively used in population genetic studies [11, 12]. The mtDNA is inherited maternally and has proven to be a useful molecular marker in studying fish populations due to the rapid base substitution rate and lack of recombination [13]. Furthermore, the varying evolution rate in different regions makes mtDNA suitable for research at various levels of conservation and population genetics studies [14–16]. In a recent study, the phylogenetic relationship among *W. attu* from the Indus River was investigated through the CO*I* gene, indicating a close genetic association [17]. Additionally, a study conducted by Kumar et al. [18] focused on examining the length-weight relationships and condition factor of *W. attu* from various rivers in India. To date, limited research has been conducted on the genetics of *W. attu* from GRS. The catchments of the GRS are significantly altered by human-induced modifications, which may lead to drastic changes in the living conditions of the aquatic organisms inhabiting these areas. Hence, it is crucial to understand the genetic characteristics of *W. attu* from tributaries of the GRS to cover a wide range of natural habitats threatened with anthropogenic alteration.

This study examines the genetic diversity level, spatial connectivity and demographic pattern of *W. attu*. We utilized three mtDNA regions, CO*I*, Cyt *b*, and CR, to analyze the genetic structure of *W. attu* populations from GRS. The information gathered through this research will provide fundamental knowledge necessary for managing the species' genetic resources and developing long-term conservation strategies.

Materials and methods

Sample collection and DNA extraction

We collected 176 samples of *W. attu* from 19 sites covering seven tributaries of the Ganga River System (GRS) namely Ganga (n = 102; collected from Seven sites L1 to L7); Ramganga (n = 5; L8); Sharada (n = 18; two sites L9-L10); Sarayu/Ghaghra (n = 7; two sites: L11-L12); Rapti (n = 7; L13); Gandak (n = 12; two sites: L14-L15) and Kosi (n = 19; four sites L16-L19) (Fig. 1) during the period of 2020–2023. A comprehensive list of the sampling sites, sample numbers, sample codes, and corresponding GPS coordinates and accession numbers are provided in Supplementary Table ST1. All samples were collected with the help of local and commercial fishermen using small-mesh cast nets. A small piece of the caudal fin was carefully clipped from each sample and stored in 2 ml screw cap tubes containing 95% ethanol. Species identification was conducted through morphological analysis, following the standard reference by Talwar and Jhingran [4].

For the collection of biological samples, the Institutional Animal Ethics Committee (Letter No. WII/IAEC/2017-18) approved the sampling protocols and has no objection to carrying out the research. Genomic DNA (gDNA) was extracted using the DNeasy Blood Tissue Kit (QIAGEN, Germany) in a final elution volume of 100 µl. The concentration and quality of the extracted gDNA were evaluated using 0.8% agarose gel electrophoresis. The samples were collected as part of a sponsored project under Biodiversity Conservation and Ganga Rejuvenation project Phase I and Phase II funded by National Mission for Clean Ganga (NMCG), Ministry of Jal Shakti, and Government of India.

mtDNA amplification and sequencing

Polymerase Chain Reaction (PCR) was carried out on all 176 samples to amplify three mtDNA regions, Cyt *b* using primers L14724 and H15915 [19], partial CO*I* using primers Fish F1 and Fish R1 [20], and partial CR using primers Fish D-loopF and Fish D-loopR [21]. Details of primers are available in Supplementary Table ST2. PCR was performed in a final volume of 10 µl reaction, containing PCR buffer, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 0.25 mM of both forward and reverse primer, 5 U of Taq polymerase and 1 µl of the template DNA (approximately 50-80ng). PCR thermal conditions comprised an initial denaturation at 95 C for 5 min, followed by 35 cycles of 95 C for 35 s, annealing

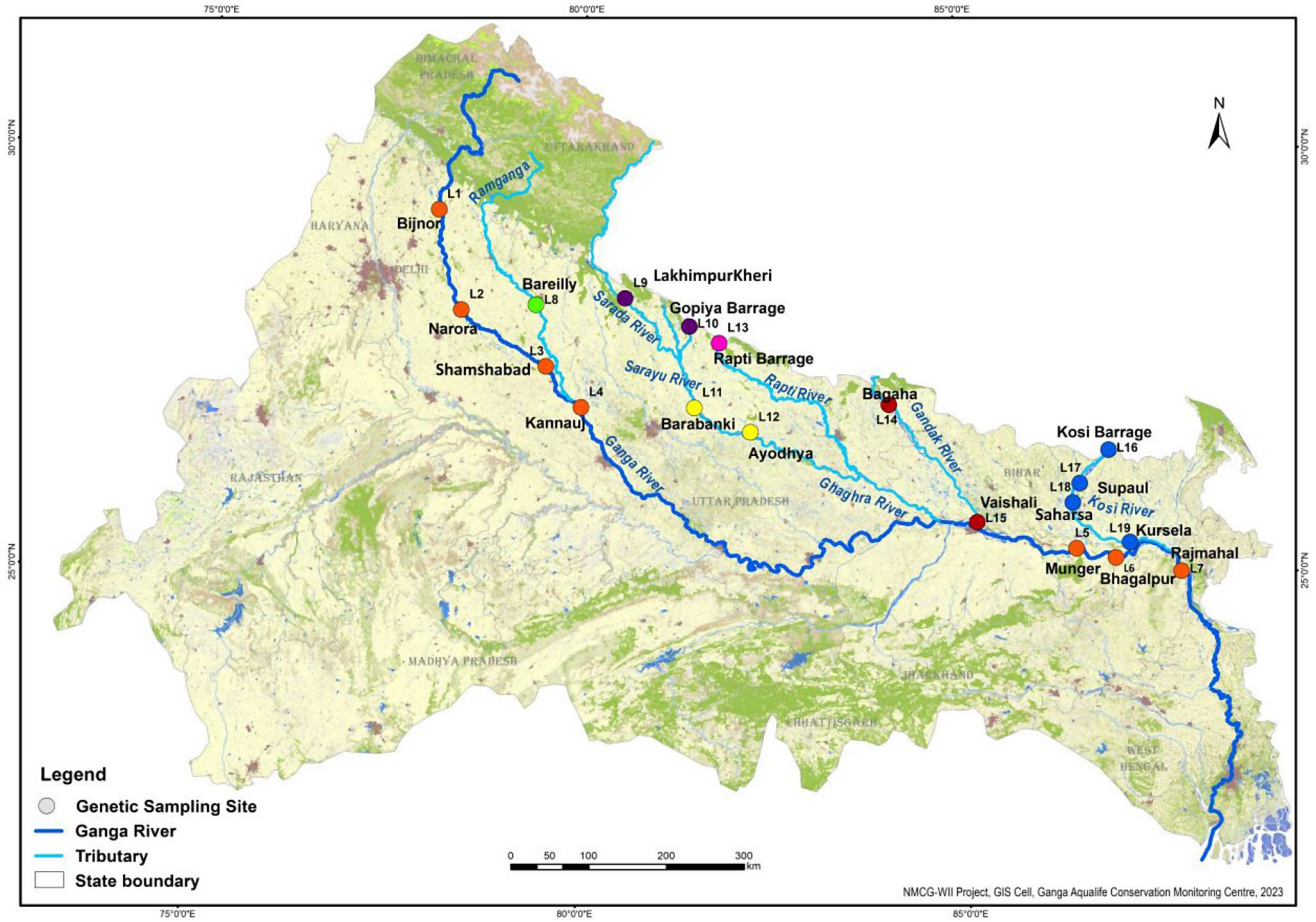

Fig. 1 Sampling sites of *Wallago attu* in the Ganga river system. Map was created using ArcGIS Desktop: Release 10.6.1 Redlands, CA: Environmental systems Research Institute

temperature at 56˚C for 45 s, and extension at 72˚C for 75 s. The final extension was at 72˚C for 10 min. Positive controls were incorporated to monitor PCR reaction efficacy and consistency. The amplified PCR products were checked on a 2% agarose gel stained with ethidium bromide and visualized under UV light. To eliminate any residual primer, the positive PCR products were treated with Exonuclease I (EXO-I) and shrimp alkaline phosphatase (SAP) for 15 min each at 37˚C and 80˚C, respectively. Subsequently, the cleaned PCR products were directly sequenced using the BigDye Terminator Kit (v3.1) and analyzed on an ABI 3730XL Applied Biosystems Genetic Analyzer. Both forward and reverse sequences were obtained for all products.

Data analysis

The sequences of Cyt *b*, CO*I* and CR were derived from the forward and reverse directions and SEQUENCHER version 4.9 (Gene Codes Corporation, Ann Arbor, MI, USA) was used to edit the sequences. All generated sequences were aligned in the CLUSTAL X alignment program [22], and the sequences were manually examined. DnaSP 5.0 [23] was used to analyze the haplotype (h), haplotype diversity (Hd), nucleotide diversity (π), and polymorphic sites (s) from concatenate mtDNA data. Median-joining (MJ) network was constructed to visualized the spatial distribution of haplotypes using the PopART software [24]. To determine whether the *W. attu* populations shows sign of spatial range expansion or a stationary population history, Tajima's D [25] and Fu's Fs [26] neutrality test was performed in DnaSP 5.0 [23]. Mismatch analysis was carried to examine the trends in spatial demography history using the population growth-decline model in DnaSP 5.0 [23]. ARLEQUIN v3.5 program [27], was used to determine the sum of squared deviations (SSD), the raggedness index (r) under the growth-decline model, genetic differentiation (FST estimates) and Analysis of Molecular Variance (AMOVA). Phylogenetic analyses was conducted in BEAST ver 1.7 [28]. The analysis was performed with MCMC chains for 10 million generations, sampled every 100 generations, and using a burn-in of 5000 generations was used. The credibility of the results was assessed using Tracer v1.6. The first 10%

per run was discarded as burn-in. Maximum credibility trees were obtained with TreeAnnotator (implemented in BEAST ver 1.7 Package). The final phylogenetic tree was visualized in FigTree v.1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). The best substitution model was determined using MrModeltest Version 3.7 [29] based on the Akaike Information Criterion (AIC). The chosen model was the Hasegawa-Kishino-Yano model (HKY) with gamma distribution for (G). The effective population size over time was determine using coalescent Bayesian skyline plots (BSP) generated in BEAST ver 1.7 [28] and visualized in Tracer version 1.7 [30]; all effective sample size (ESS) were above 200. The analysis was performed using Hasegawa–Kishino–Yano (HKY) model substitution model and the strict molecular clock model was used. The analysis was run with sampling every 10,000 generations and 10% of first generations were discarded (burn-in).

Results

Sequence variations, haplotype distribution and spatial connectivity

The three mtDNA regions consisting of CO*I*: 632 bp, Cyt *b*: 1054 bp and CR: 734 bp from 176 samples were generated. Sequence variations and gene diversity were estimated using the concatenated dataset CO*I* + Cyt*b* + CR of 2420 bp. In the whole dataset, eighty polymorphic sites were identified, which represent 3.30%, including 74 parsimony informative sites and 6 singleton variable sites. The estimated nucleotide frequencies indicated AT bias among the generated region. Among all the concatenated sequences, 103 haplotypes were identified. Among 103 haplotypes, 62 were identified from Ganga (Hap 1-Hap 62), 2 from Ramganga (Hap 63- Hap 64), 12 Sharda (Hap 61, Hap 65- Hap 75), 10 Sarayu/Ghaghra (Hap1, Hap 68, Hap 76-Hap 83), 6 Rapti (Hap 84-Hap 89), 6 Gandak (Hap 61, Hap 62, Hap90-Hap93), and 10 Kosi (Hap 94 - Hap 103). Hap1 shared with Ganga and Sarayu/Ghaghra whereas, Hap 61 was found to share with Ganga, Gandak and Sharda River, Hap 62 shared with Ganga and Gandak and Hap68 shared with Sharda and Sarayu/Ghaghra. Interestingly, no haplotype sharing was observed in Ramganga, Kosi and Rapti and also each population had its own independent haplotypes (Supplementary Table ST3). The median-joining network of the 103 haplotypes represented the distribution pattern of haplotypes among the *W. attu* populations (Fig. 2). The results obtained from Median-joining networks indicate that the haplotypes from the seven populations do not exhibit a clear river wise genetic structure. Moreover, MJ analysis reveals the existence of three genetic clades within the GRS. Clade I was found in all seven rivers with wide distribution. Clade II was found in Ganga, Sharda, Sarayu/Ghaghra, Rapti and Kosi River, whereas Clade III was confined to Ganga, Gandak and Sarayu/Ghaghra river. Despite the three genetic groups, the network did not display any evident star burst-like topological structures or core haplotypic distributions. These findings offer significant insights into the genetic relationships and spatial connectivity among the populations in the region, indicating complex interactions and potentially unique evolutionary processes of *W. attu* in the GRS. In phylogenetic analysis, the result was consistent with MJ network and three Clades were formed (Supplementary Fig. SF1). The Clade III which consists of individuals from Ganga, Gandak and Sarayu/Ghaghra river formed basal clade, indicating that the sequences of Clade III are much diverse than the individuals of Clade I and Clade II.

Genetic diversity and genetic differentiation

The Hd and π of the Ganga, Ramganga, Sharda, Sarayu/Ghaghra, Rapti, Gandak and Kosi river were 0.984 and 0.0063; 0.60 and 0.0002; 0.961 and 0.005; 0.962 and 0.0062; 0.952 and 0.003; 0.848 and 0.0039; 0.860 and 0.0046 respectively (Table 1). The overall Hd and π value was 0.947 ± 0.006 and 0.0062, respectively in *Wallago attu*. Pairwise Fst comparison of populations showed significant genetic difference between the population's (Fig. 3). Moreover, low F_{ST} was observed between Ganga and Sarayu/Ghaghra (0.047), where it was high between Ramganga and Rapti (0.735). AMOVA of all seven river populations were determined to identified the presence of population genetic structure and found that 84.70% variation was attributed within the populations and 10.11% variation among populations (Table 2).

Population demography

The neutrality test (Tajima's D and Fu's Fs test), and mismatch distributions analysis were conducted to infer the demographic history of *W. attu*. Multimodal mismatch distributions and non-significant Tajima's D and Fu's Fs values (except Ganga, significant Fu's Fs) indicated that *W. attu* populations have remained relatively stable over time (Fig. 4A; Table 1). The demographic analysis was also supported by the generalized least square procedure and the raggedness index of the distribution in studied rivers (Table 1). The Bayesian skyline plot was performed to understand the past demography history of studied species. Bayesian skyline plot (BSP) analyses supported the hypothesis of a relatively recent population expansion of *W. attu* in GRS. BSP revealed that the population size had no pronounced demographic changes for a long time, before it experienced

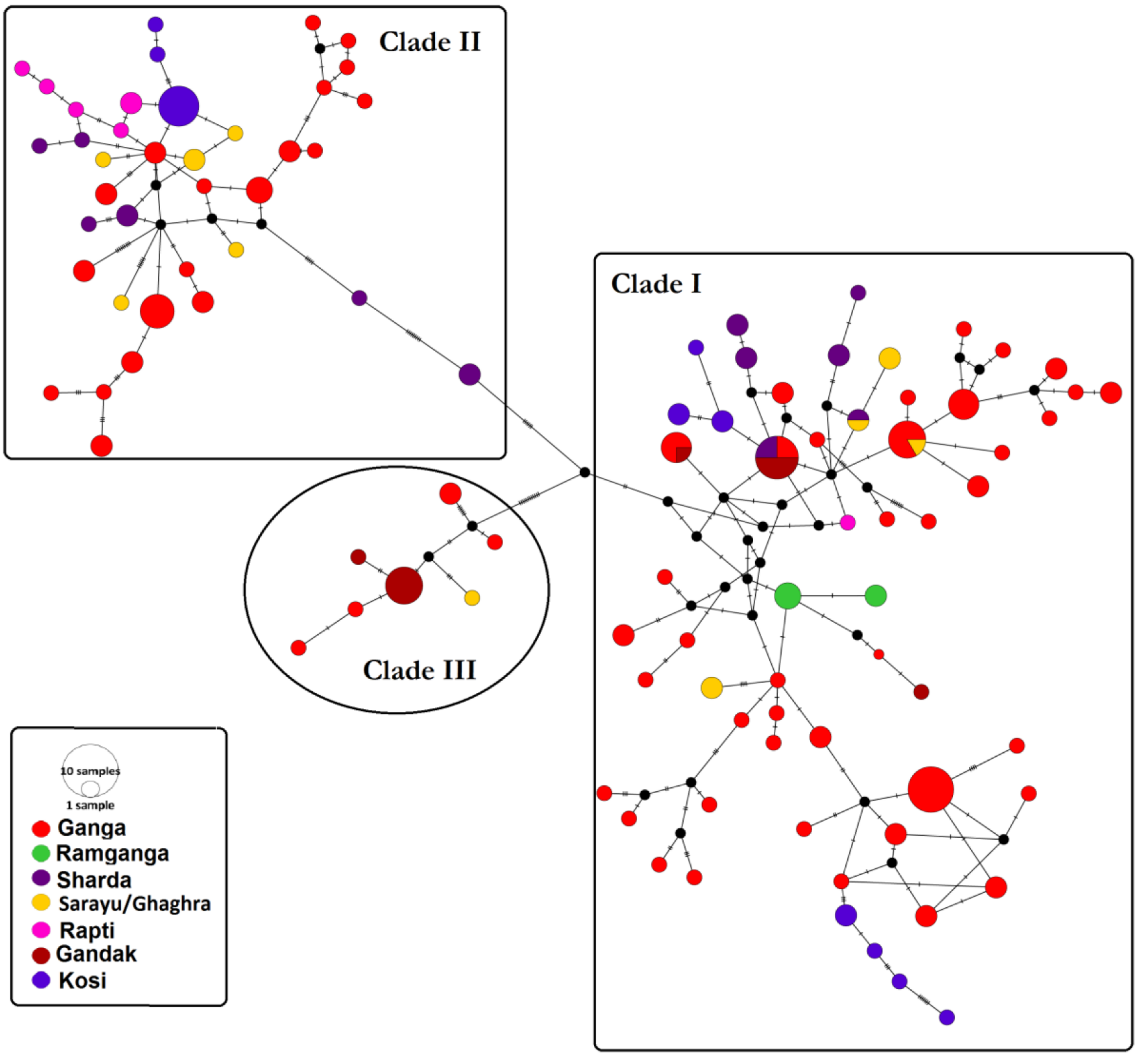

Fig. 2 The median-joining network based on the concatenate dataset of three mtDNA regions CO*I*, Cyt *b* and CR of *Wallago attu*. The number of lines (bar) between nodes represents the mutation sites, and the size of the circle represents the number of individuals in each haplotype

Table 1 Summary of genetic diversity and neutrality tests of demographic patterns in *Wallago attu*

River	N	S	H	Hd	π	Fu's F_S (P)	Tajima's D (P)	SSD (P)	Rg (P)
Ganga	102	70	62	0.984 ± 0.005	0.0063	-23.73*	0.462	0.011	0.0038
Ramganga	5	1	2	0.600 ± 0.175	0.0002	0.626	1.224	0.054	0.400
Sharda	18	29	12	0.961 ± 0.026	0.005	0.049	2.066	0.012	0.013
Sarayu/Ghaghra	13	44	10	0.962 ± 0.041	0.0062	0.186	0.290	0.026	0.024
Rapti	7	23	6	0.952 ± 0.096	0.003	-0.308	-1.00	0.170*	0.052
Gandak	12	23	6	0.848 ± 0.074	0.0039	2.953	1.078	0.081	0.168*
Kosi	19	23	10	0.860 ± 0.070	0.0046	1.606	2.700	0.033	0.072
Overall	176	80	103	0.99 ± 0.002	0.0062	-2.66	0.973	0.055	0.105

Sample size (N), polymorphic sites (S), number of haplotypes (H), haplotype diversity (Hd), nucleotide diversity (π), the sum of squared deviations (SSD) and Harpending's raggedness index (Rg) and *P* is the probability value.*$P < 0.05$

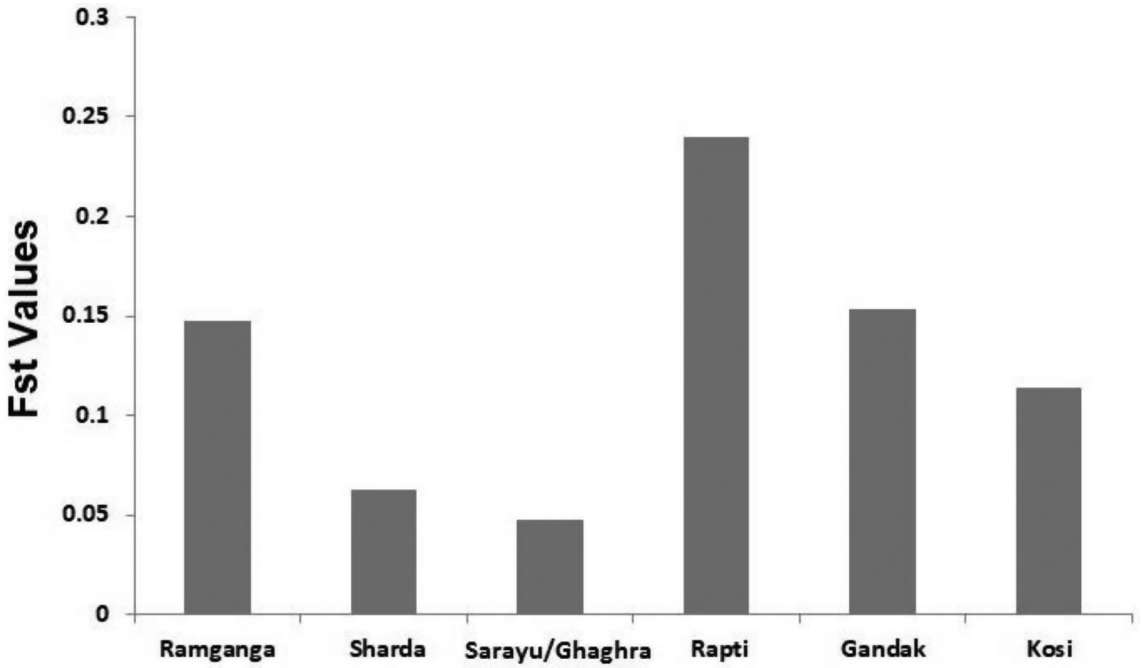

Fig. 3 Genetic differentiation among the *Wallago attu* populations in the Ganga river system. X-axis represent the F_{ST} values relative to the Ganga River based on the concatenate dataset of three mtDNA regions CO*I*, Cyt *b* and CR.

Table 2 AMOVA analyses of mtDNA sequences for seven populations of *Wallago attu* from Ganga river system

Source of variation	df	Sum of squares	Variance components	Percentage of variation	Fixation index	p-value
Among groups	5	148.78	0.81432	10.11	F_{CT}= 0.101	0.099
Among populations Within groups	1	16.47	0.41841	5.19	F_{SC}= 0.057	0.048
Within populations	169	1153.41	6.82	84.70	F_{ST}=*0.152*	$P<0.001$
Total	175	1318.670	8.05			

an apparent population expansion at approximately 0.04 Ma (Fig. 4B).

Discussion

One of the largest rivers ecosystems in the Asia, the GRS is characterized by high species richness and has been identified as one of the important centers of aquatic biodiversity [31]. The GRS has been identified as a biodiversity hotspot, requiring immediate protection to preserve its rich variety of aquatic species. The region faces numerous anthropogenic threats, making conservation efforts urgent and crucial to safeguard its unique biodiversity. Hence, understanding the population and connection with response to gene flow and population demography of exploited species is of utmost importance in ensuring sustainable fisheries management [32]. The present study provides the first detailed insights into the genetic characteristics of *W. attu* from GRS.

Genetic diversity and spatial connectivity

Examining the mtDNA variations has unveiled high mtDNA gene diversity within the populations of *W. attu* across seven distinct rivers. *W. attu*, a widely distributed and frequently observed fish species in GRS, the large population size and high fecundity rate may account for the high haplotype diversity observed in this region. Haplotype sharing was observed in various river combinations: Ganga and Sarayu/Ghaghra, Ganga, Gandak, and Sharda, and Ganga, Gandak, Sharda, and Sarayu. This phenomenon underscores the natural genetic interchange, as exemplified by the connection between Sharda, an upper segment of the Ghaghra River, also known as Sarayu River, merging with the Ganga in Saran, Bihar. Similarly, the convergence of Ganga and Gandak occurs in Hajipur, Bihar, contributing to genetic sharing. The sharing of haplotypes and significant variations in *W. attu* populations, could be attributed to the diverse environmental conditions and distinct selective pressures across different river systems [33]. Moreover, the rapid expansion of the population likely facilitated the accumulation of

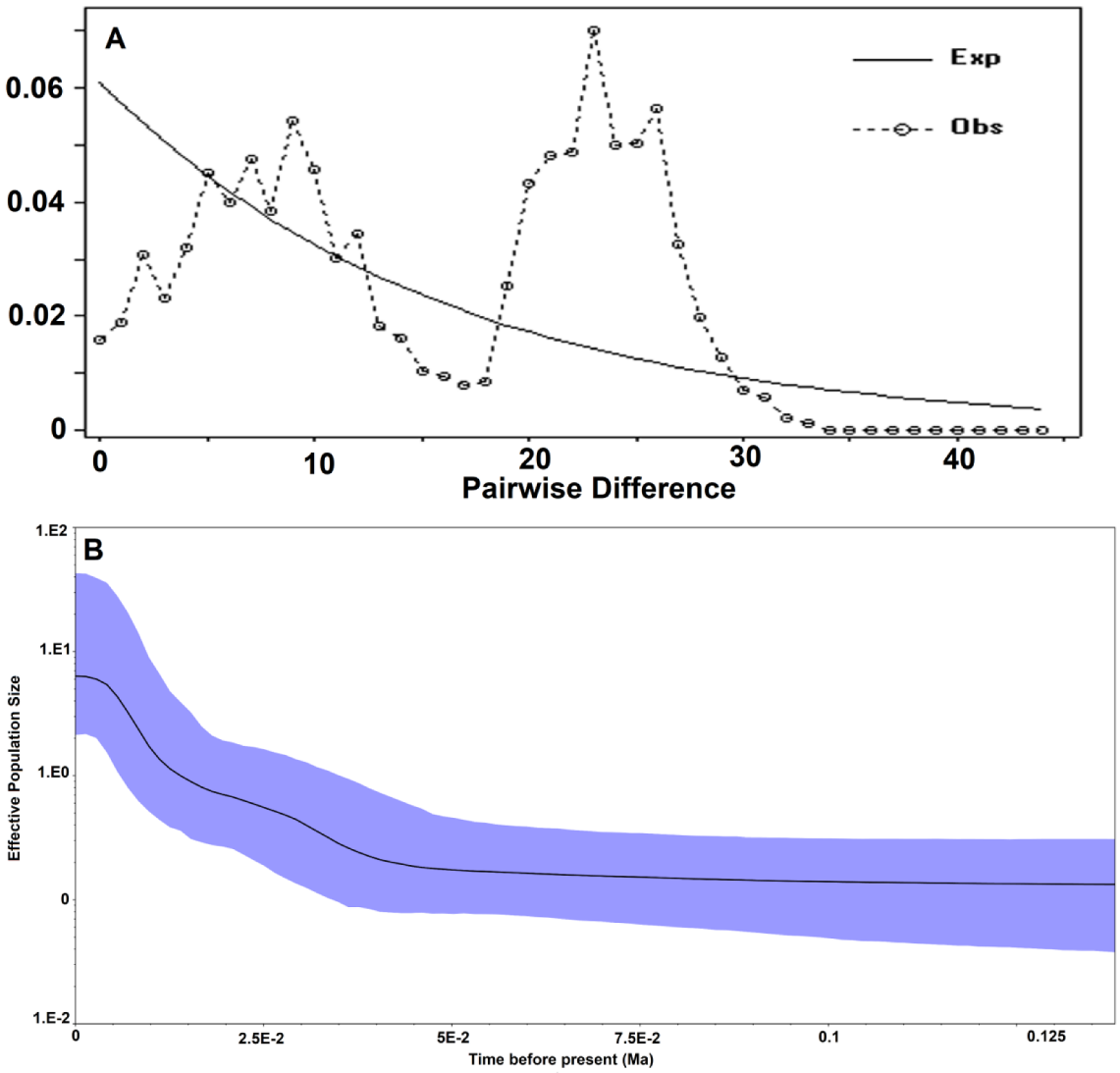

Fig. 4 **A** Mismatch distribution graph for overall populations of *Wallago attu* in the Ganga river basin. The X and Y axis show the number of pairwise differences and the frequency of the pairwise comparisons, respectively. The observed frequencies are represented by dotted line. The frequency expected under the hypothesis of constant population model is depicted by solid line. **B** Bayesian Skyline plot illustrating changes in effective population size (Ne) over time (x-axis) using the entire dataset. The y-axis (logarithmic scale) represents population size, estimated as Ne × τ (Ne = effective population size; τ = generation time). The shaded region around the solid median estimates indicates the 95% highest posterior density estimate of the historic effective population size

mutations, with the period of growth providing ample time for the development of haplotype diversity [34].

Phylogenetic analysis revealed the presence of different genetic groups within the species, but they did not correspond to any region-specific genetic signatures. Additionally, network analysis also indicated the presence of three genetic clades in *W. attu*. It suggests that *W. attu* is currently undergoing diversification, increasing the likelihood that certain individuals within the population will possess variations better suited for their environment. Similar results were noted in different fish species, including *Schizopygopsis younghusbandi* [34], *Ancherythroculter nigrocauda* in the upper Yangtze River [35], *Tor putitora* from the Upper Ganga region in India [36], and the *Pethia* genus in Sri Lanka [37].

Genetic distance

Genetic distances between seven populations indicated that genetic differentiation ranged from low to high. As the distance between tributaries increases, so does the genetic differentiation among the populations. In general, river currents play an important role in dispersing fish, resulting in low genetic divergences concerning the Ganga river where many rivers merge to form the main river system. Moreover, AMOVA analysis showed that genetic variation in *W. attu*

mostly occurred within populations, while there was moderate differentiation among populations ($P < 0.001$). The level of genetic differentiation and AMOVA showed consistent results, perhaps attributable to high gene flow. Lack of barriers to dispersal, and strong dispersal capacity could facilitate genetic exchanges among groups across their distribution, being possible reasons for the low genetic differentiation among the GRS.

Historical demography

The historical demographic expansions of *W. attu* in the GRS were investigated using the neutrality tests Tajima's D and Fu's Fs. Significantly negative values of D and Fs can indicate either population expansion or purifying selection [38]. In this analysis, Tajima's D and Fu's Fs tests showed no statistically significant differences ($p > 0.05$), except for the Ganga population, which exhibited a notably large negative and significant Fu's Fs value. Consequently, historical population stability was observed in the studied populations, while the Ganga population demonstrated signs of demographic expansion. Additionally, interpreting the neutrality test results for the Ganga population, where a significant negative value was accompanied by a non-significantly positive Tajima's D value, requires caution. Thus, we conclude that all the examined *W. attu* populations within the GRS likely maintained equilibrium in the past. This result was further supported by BSP analysis, which indicated recent population growth following a prolonged period of historical stability in population size. BSP analysis revealed a pattern in which the population maintained a relatively stable size over an extended period before undergoing a significant expansion around 0.04 Ma, during the late Pleistocene era. Geological and climatic events in the past have undeniably played a pivotal role in the population dynamics of various fish species, including *Schizopygopsis younghusbandi* [34]. Our findings lend support to the hypothesis that the expansion and evolutionary trajectory of *W. attu* are closely linked to environmental and climatic shifts. Conversely, the Pleistocene epoch (0.01–1.9 Ma) was characterized by a series of substantial glacial-interglacial fluctuations, potentially exerting profound effects on the geographical distribution and abundance of organisms due to cycles of fluctuating water levels [39]. These recurrent glacial-interglacial changes within the GRS during the Pleistocene might have also played a role in influencing the expansion of *W. attu*.

Our findings imply that a higher genetic variation increases the likelihood of individuals within a population possessing advantageous traits for their environment. This adaptive advantage leads to the prolonged survival of these individuals and their descendants, eventually contributing to the diversification of the population into distinct genetic lineages over subsequent generations. Further exploration is warranted, particularly in other Indian river systems, to gain insights into the genetically diverse populations of this species.

Conclusion

Gaining insights into the genetic diversity, connectivity, and population structure can enhance both fishery management and conservation efforts aimed at preserving vulnerable freshwater fish species. As an economically important fish species, *W. attu* holds a widespread presence across the Indian subcontinent and other regions of Southeast Asia. This study's revealed high genetic diversity, and moderate genetic differentiation between populations, indicates the initiation of divergence within the *W. attu* found in the GRS. The outcomes of this research furnish comprehensive details about the population genetics and historical demographics of *W. attu*. Furthermore, these findings lay the groundwork for assessing germplasm resources and effective resource management for this species. The implications of these results extend towards refining our comprehension of population genetics, ultimately offering crucial insights for sustainable utilization, prudent fishery management, and enduring conservation strategies applicable to this species.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11033-024-09323-w.

Acknowledgements We extend our heartfelt gratitude to Shri G. Ashok Kumar, Director General (DG) of NMCG, along with Mr. Rajiv Ranjan Mishra and Mr. Upendra Prasad Singh, Former DGs of NMCG, and their dedicated team for their invaluable support and funding. We would also like to express our appreciation to the Chief Wildlife Warden of the Government of Uttar Pradesh, Bihar and Jharkhand for their assistance in providing research permits. We acknowledge the support and help provided by the Director, Wildlife Institute of India (WII) in the successful completion of this research. We thank to Dr. J. Antony Johnson, Scientist-F at WII, for his insightful comments and dedicated efforts to enhance the quality of our manuscript. We appreciate the help of Ms. Aishwarya Ramachandran in map preparation, Mr. Madhanraj, Wildlife Forensic and Conservation Genetic Cell and Ms. Maneesha Bhatt for their support. Furthermore, we are grateful for the collaborative efforts of the NMCG-WII team whose assistance during the study greatly contributed to the accomplishment of our study objectives.

Author contributions S.A.H. and R.B. conceived the project, raised funds, and reviewed/edited the paper. S.K.G. designed the methodology for the molecular study of the project, supervised, validated, and reviewed/edited the paper. A.K. designed the methodology, analyzed the data, and wrote the paper. A.K., N.N., and N.Y. collected samples, performed the laboratory work and generated the data. All authors agreed and approved the final paper.

Funding This study was conducted as part of the project "Planning &

Management for Aquatic Species Conservation and Maintenance of Ecosystem Services in the Ganga River Basin," funded under reference number B-03/2015–16/1077/NMCG – NEW PROPOSAL by the National Mission for Clean Ganga (NMCG), Ministry of Jal Shakti, Government of India.

Data availability The newly generated sequences have been deposited in GenBank under accession numbers OR945533-OR945708 for the cytochrome c oxidase subunit *I*; OR943701-OR943876 for cytochrome b gene and OR943877-OR944052 for the control region and other details are presented in the Supplementary Table ST1.

Declarations

Ethical approval Institutional Animal Ethics Committee (Letter No. WII/IAEC/2017-18) approved the sampling protocols and has no objection to carrying out the research.

Competing interests The authors declare that they have no competing financial/personal interests.

References

1. Kumar D (2017) River ganges – historical, cultural and socioeconomic attributes. Aquat Ecosyst Health Manag 20(1–2):8–20
2. Singh M, Singh AK (2007) Bibliography of environmental studies in natural characteristics and anthropogenic influences on the Ganga River. Environ Monit Assess 129:421–432
3. Roberts TR (1993) Artisanal fisheries and fish ecology below the great waterfalls of the Mekong River in southern Laos. Nat Hist Bull Siam Soc 41:31–62
4. Talwar PK, Jhingran AG (1991) Inland fishes of India and adjacent countries. CRC, Florida, United States
5. Roberts TR (2014) *Wallago* Bleeker, 1851 and Wallagonia Myers, 1938 (Ostariophysi, Siluridae), distinct Genera of Tropical Asian catfishes, with description of *Wallago maemohensis* from the Miocene of Thailand. Bull Peabody Mus Nat Hist 55(1):35–47
6. Goswami PK, Devaraj M (1992) Breeding, age and growth of the freshwater shark *Wallago attu* (Bloch and Schneider) from the Dhir Beel of the Brahmaputra basin, Assam, India
7. Prasad H, Desai AY (2020) Fecundity, spawning, and ova diameter of the *Wallago attu* from Bhadar reservoir of Gujarat, India. J Entomol Zool Stud 8(4):540–542
8. Mishra SS, Acherjee SK, Chakraborty SK (2009) Development of tools for assessing conservation categories of siluroid fishes of fresh water and brackish water wetlands of South West Bengal, India. Environ Biol Fish 84(4):395–407
9. Patra MK, Acharjee SK, Chakraborty SK (2005) Conservation categories of siluroid fishes in North-East Sundarbans, India. Biodivers Conserv 14:1863–1876
10. Zafar MM, Abbas K, Basharat H, Ahmad T, Hussain U (2016) Microsatellite DNA marker based genetic polymorphism in *Wallago attu*. Int J Bioinform Biol Syst 1(2):18–24
11. Sun P, Shi Z, Yin F, Peng S (2012) Population genetic structure and demographic history of Pampus argenteus in the Indo-West Pacific inferred from mitochondrial cytochrome b sequences. Biochem Syst Ecol 43:54–63
12. Zhao L, Zhou X, Liu Q, Zhang H (2013) Genetic variation and phylogeography of *Sinibrama macrops* (Teleostei: Cyprinidae) in Qiantang river, China. Biochem Syst Ecol 49:10–20
13. Guo XH, Liu SJ, Liu Q, Liu Y (2004) New progress on mitochondrial DNA in fish. Acta Genet Sin 31:983–1000
14. Hebert PD, Ratnasingham S, De Waard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc Lond Ser B: Biol Sci 270:S96–S99
15. 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*. Mol Biol Rep 39(2):1765–1775
16. Liang JJ, Liu Y, Zhang XF, Zhang XY, Yue BS, Song ZB (2011) An observation of the loss of genetic variability in prenant's schizothoracin, *Schizothorax prenanti*, inhabiting a plateau lake. Biochem Syst Ecol 39:361–370
17. Sajjad A, Jabeen F, Ali M, Zafar S (2023) DNA barcoding and phylogenetics of *Wallago attu* using mitochondrial COI gene from the River Indus. Journal of King Saud University-Science, 35(6), p.102725
18. Kumar G, Kashyap A, Serajuddin M (2023) Length-weight relationships and condition factor of Asian sheat catfish, *Wallago attu* (Bloch & Schneider, 1801) inhabiting different rivers of India. J Fisheries 11:111203–111203
19. 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
20. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. Philosophical Trans Royal Soc Part B: Biol Sci 360(1462):1847–1857
21. Cheng YZ, Xu TJ, Jin XX, Tang D, Wei T, Sun YY, Meng FQ, Shi G, Wang RX (2012) Universal primers for amplification of the complete mitochondrial control region in marine fish species. Mol Biol 46:727–730
22. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876
23. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452
24. Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype network construction. Methods Ecol Evol 6(9):1110–1116
25. Tajima H (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595
26. Fu Y (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925 (1997)
27. Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mole Ecol Resour 10:564–567
28. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969–1973
29. Posada D, Crandall K (1998) Modeltest: testing the model of DNA substitution. Bioinform 14:817–818
30. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in bayesian phylogenetics using Tracer 1.7. Syst Biol 67(5):901–904
31. Sarkar UK, Pathak AK, Sinha RK, Sivakumar K, Pandian AK, Pandey A, Dubey VK, Lakra WS (2012) Freshwater fish biodiversity in the River Ganga (India): changing pattern, threats and conservation perspectives. Rev Fish Biol Fish 22:251–272
32. Ovenden JR, Berry O, Welch DJ, Buckworth RC, Dichmont CM (2015) Ocean's eleven: a critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries. Fish Fish 16(1):125–159

33. Modeel S, Joshi BD, Yadav S, Bharti M, Negi RK (2023) Mitochondrial DNA reveals shallow population genetic structure in economically important Cyprinid fish *Labeo rohita* (Hamilton, 1822) from South and Southeast Asia. Mol Biol Rep 50(6):4759–4767
34. Guo SS, Zhang GR, Guo XZ, Wei KJ, Yang RB, Wei QW (2014) Genetic diversity and population structure of *Schizopygopsis younghusbandi* Regan in the Yarlung Tsangpo River inferred from mitochondrial DNA sequence analysis. Biochem Syst Ecol 57:141–151
35. Zhai D, Zhang Z, Zhang F, Liu H, Cao W, Gao X (2019) Genetic diversity and population structure of a cyprinid fish (*Ancherythroculter nigrocauda*) in a highly fragmented river. J Appl Ichthyol 35(3):701–708
36. Yadav P, Kumar A, Hussain SA, Gupta SK (2020) Evaluation of the effect of longitudinal connectivity in population genetic structure of endangered golden mahseer, *Tor putitora* (Cyprinidae), in himalayan rivers: implications for its conservation. PLoS ONE 15(6):e0234377
37. Sudasinghe H, Ranasinghe T, Herath J, Wijesooriya K, Pethiyagoda R, Ruber L, Meegaskumbura M (2021) Molecular phylogeny and phylogeography of the freshwater-fish genus *Pethia* (Teleostei: Cyprinidae) in Sri Lanka. BMC Ecol Evol 21:1–24
38. Liu JY, Brown C, Yang TB (2009) Population genetic structure and historical demography of grey mullet, *Mugil cephalus*, along the coast of China, inferred by analysis of the mitochondrial control region. Biochem Syst Ecol 37:556–566
39. Dynesius M, Jansson R (2000) Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. Proc Natl Acad Sci U S A 97:9115–9120