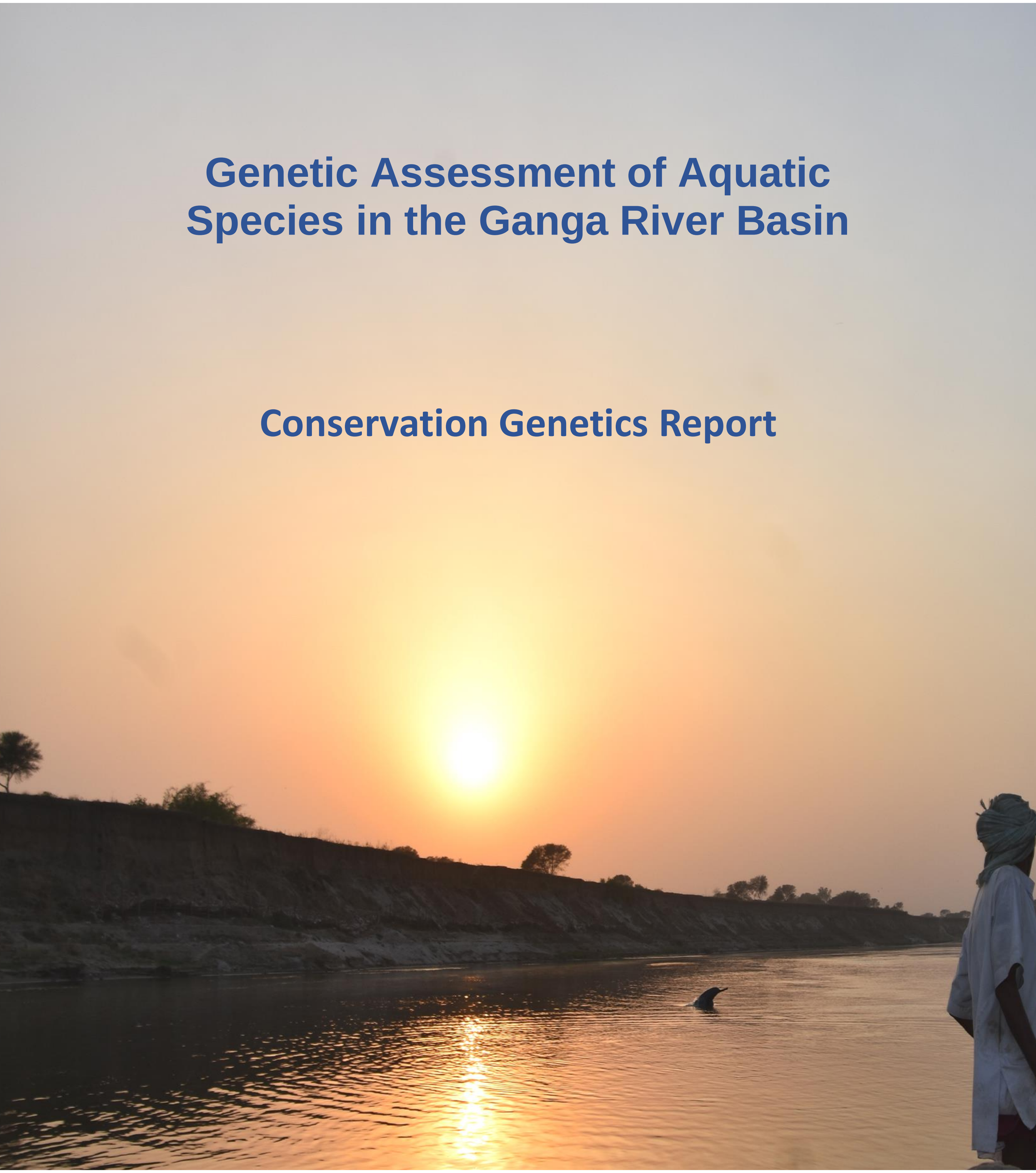
Genetic Assessment of Aquatic Species in the Ganga River Basin

Conservation Genetics Report



PLANNING AND MANAGEMENT FOR AQUATIC SPECIES CONSERVATION AND MAINTENANCE OF ECOSYSTEM SERVICES IN THE GANGA RIVER BASIN FOR A CLEAN GANGA

Genetic Assessment of Aquatic Species in the Ganga River Basin

Project Investigator's

Ruchi Badola, Dean & Nodal Officer NMCG-WII
S. A. Hussain, Project Manager

Component Investigator's

Sandeep Kumar Gupta, Scientist-E

Research Team

Ajit Kumar, Project Scientist-II
Ashish Kumar Panda, Senior Project Associate
Aftab Alam Usmani, Principal Project Associate
Naveen Negi, Project Associate
Anand Sharma, Project Intern
Maneesha Bhatt, Project Intern

Maps and GIS support

Zeeshan Ali

Citation

WII-GACMC (2024). Planning and Management for Aquatic Species Conservation and Maintenance of Ecosystem Services in the Ganga River Basin for a Clean Ganga.
Genetic Assessment of Aquatic Species in the Ganga River Basin

Wildlife Institute of India
Chandrabani, Dehradun - 248001,
Uttarakhand, India

This document is an output of the project
Planning and Management for Aquatic Species Conservation and Maintenance of Ecosystem Services in the Ganga River Basin for a Clean Ganga sponsored by the National Mission for Clean Ganga, Ministry of Jal Shakti, Government of India, New Delhi.

CONTENTS

1. Introduction

1.1 Ganga River System

The Ganga River system, the largest river in India, spans approximately 1087300 km^2 across India and Nepal. In recent years, the growing population in the Ganga basin has led to agricultural expansion, urbanization, and early-stage industrialization, resulting in extensive water usage for irrigation, industry, and public supply. Unfortunately, segments of the Ganga River suffer from pollution caused by sewage, untreated industrial discharges, and nutrient-rich runoff from agricultural fields. Additionally, climate change projections, including melting headwater glaciers and altered rainfall patterns, pose significant threats to the river's water availability and contribute to changed droughts and flooding patterns.

Scientific evidence has highlighted that impoundments and reservoir-induced habitat fragmentation hinder natural migration in aquatic fauna, particularly fish, leading to spatially and genetically fragmented populations. This fragmentation adversely affects the genetic characteristics of these populations, a critical concern in the biodiversity-rich Ganga River basin, where hydroelectric, irrigation, and recreational reservoirs impede connectivity in the aquatic system. With growing apprehensions about water scarcity, declining biodiversity, and effective river management in the face of environmental changes, the Indian Government has initiated plans for the Ganga River system. These plans aim at a comprehensive, long-term strategy to enhance environmental flow, restore biodiversity, and address water scarcity issues.

The health of aquatic habitats can be gauged by observing the abundance of aquatic life in and around river stretches, as diverse species indicate the well-being of the aquatic ecosystem. Globally, indicator species or communities have been employed to monitor aquatic ecosystem health, as their presence, absence, or abundance reflects specific environmental conditions. However, identifying suitable indicator species remains a challenge, requiring scientifically sound links between an organism's presence and an environmental condition to justify its use as a proxy.

To restore the ecological integrity of the aquatic ecosystem, it is crucial to recover and enhance its biodiversity. Science-based species management is essential to avoid the risk of inbreeding depression, maintain genetic diversity and fitness of individuals, and limit species introgression. Consequently, this will maintain and enhance the survival rate of species. Maintaining the genetic status of in aquaculture and the effective conservation of wild relatives are equally important.

1.2 Conservation Genetics

Conservation genetics plays a crucial role in addressing the challenges posed by habitat loss, climate change, and other threats to biodiversity. By integrating genetic information into conservation planning, scientists can make more informed decisions to ensure the long-term survival of endangered species and the health of ecosystems. The main goals of conservation genetics are to understand the genetic factors influencing the viability of populations, identify and manage threats to genetic diversity, and develop strategies for sustainable conservation.

Genetic diversity of species has been correlated with individual fitness and the adaptive capacity of populations in the face of environmental changes. The reduction in genetic variability ultimately heightens the risk of extinction. In particular, small populations face increased vulnerability to inbreeding and loss of genetic diversity due to genetic drift. Habitat fragmentation, another significant threat to various species, exacerbates these issues by impeding gene flow between populations. The levels of genetic diversity within populations and the divergence between them are anticipated to be influenced by the degree of fragmentation. As the geographic distance between fragments or the duration of isolation increases, the expectation is for a decline in genetic diversity within smaller populations, leading to population divergence.

Currently, a diverse set of molecular genetics techniques is being applied to evidence sourced from a broad range of taxa, encompassing aquatic life conservation. This expansive application holds immense potential for the preservation of many endangered aquatic species. The assessment of genetic diversity encompasses various methodologies, such as examination of karyotypes and single-locus markers (including allozymes, mitochondrial DNA restriction fragment length polymorphisms), employing random amplified polymorphic DNA, examining mtDNA sequences, scrutinizing nuclear DNA sequences, and utilizing microsatellites (simple sequence repeats). Currently, mitochondrial and nuclear DNA sequences data are the most informative methods for characterizing genetic variability at or above the level of populations. Leveraging molecular markers offers a powerful toolset for unraveling population genetic structures, deciphering phylogeographic relationships, identifying management units (MUs — defined by shared distinctive allele frequencies among populations), and delineating evolutionarily significant units (ESUs — characterized by robust phylogenetic structures based on multilocus mtDNA or nDNA variations).

1.2.1 Current Techniques used in Conservation genetics

Mitochondrial DNA Sequencing

Mitochondrial DNA (mtDNA) stands out as a prominent population genetic marker, owing to a confluence of intrinsic and technical characteristics. Its relatively rapid rate of base substitution, haploid nature, and maternal inheritance collectively provide invaluable insights into the genetic diversity, population structure, and evolutionary history of species. The process of DNA sequencing involves discerning each nucleotide (base) within a specified target region of DNA, commonly known as the genetic marker.

The sequencing depth of mtDNA yields a unique sequence specific to a particular species, facilitating precise and reliable identification. This utilization of mitochondrial DNA sequencing not only enhances our understanding of population dynamics but also serves as a powerful tool for species delineation and conservation efforts.

Microsatellite genotyping

Microsatellites are widely acceptable and versatile genetic markers, finding numerous applications in population genetics, conservation biology, and evolutionary studies. These genetic markers comprise arrays of DNA sequences characterized by tandem repeats of mono-, di-, tri-, and tetra-nucleotide units, distributed widely across the genomes of numerous eukaryotic species. Their codominant nature, high polymorphism, ease of typing, and mode of inheritance make them exceptionally well-suited for exploring population structures and conducting pedigree analyses. Additionally, microsatellites are adept at discerning differences among closely related species. The automated application of PCR for microsatellites allows for the efficient identification of simple sequence repeat polymorphisms. Furthermore, only small amounts of blood samples or alcohol-preserved tissue are requisite for their analysis. Notably, the majority of microsatellites are noncoding, rendering their variations independent of natural selection. These unique characteristics position microsatellites as ideal genetic markers for applications in conservation genetics and fisheries management.

1.3 Conservation challenges for freshwater biodiversity

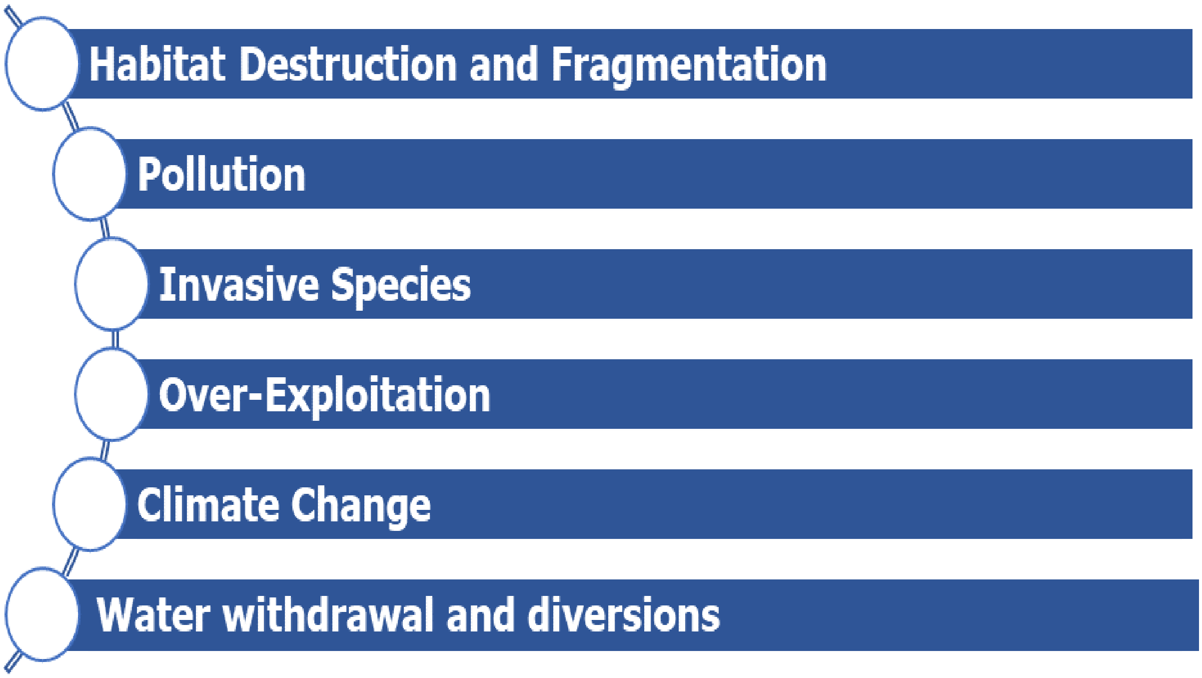
Freshwater ecosystems harbour over 40 percent of the globe's fish species, underscoring their critical ecological significance. However, despite their immense value, numerous lakes, rivers, and wetlands worldwide face severe degradation due to human activities. Alarmingly, these aquatic ecosystems are undergoing decline at a rate outpacing that of terrestrial ecosystems. This concerning trend highlights the urgent need for concerted conservation efforts to mitigate the detrimental impact of human activities on aquatic habitats.

1.4 Use of genetic information in aquatic conservation

Assessing genetic diversity and gene flow
Assessing population genetic structure
Identifying population demographic trends
Determine effective population size
Monitor inbreeding and outbreeding
Identify cryptic, alien or invasive species
Identify pure genetic stocks for population recovery programmes
Design breeding programmes for conservation and re-introduction plan

1.5 Threats

The historic distribution ranges of many aquatic species has almost shifted and is being increasingly confined to restricted zones. Primarily, human activities such as habitat destruction/fragmentation, hunting, and human-mediated translocation, have influenced the distribution, population structure, and genetic diversity of aquatic species. Therefore, a study on genetic variability is crucial to identify the fitness of individuals which regulates adaptation to change in environmental conditions and stress. Degradation of genetic diversity reduces the adaptive potential, increases chances of inbreeding which poses the risk of extinction of the species. There have been very few scientific efforts with almost no baseline information on the genetic status of threatened/endangered aquatic species of Ganga River.



1.6 Objectives

The primary goal of this project is to assess the genetic diversity, gene flow patterns, and population demography of select aquatic species in the Ganga River basin. To achieve this, we utilized mitochondrial and microsatellite markers.

The data collected from this research will provide crucial insights necessary for managing the genetic resources of these species and developing effective long-term conservation strategies.

2. UNDERSTANDING THE SPATIAL CONNECTIVITY IN ASIAN SILURID CATFISH (*Wallago attu*) IN THE GANGA RIVER SYSTEM

2.1 Background of the study

The Ganga River system (GRS) is the world's most populous river basin and also biodiversity hotspots. This extensive basin is intricately interwoven with a vast network of tributaries, which are threatened by anthropogenic pollution and degradation, significantly impacting aquatic biodiversity and their habitats (Kumar, 2017). It also faces the challenge of accommodating the highest human density. Consequently, these ecological disruptions hinder the conservation efforts aimed at preserving the river's delicate ecosystem (Singh and Singh, 2007). To comprehensively assess the implications of the contemporary environmental threats, losses for biodiversity management, harnessing the link between genetics and associated habitat in populations of a 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. This initiative's focus on ecological restoration and health resonates with the urgent need to ensure the long-term viability and stability of the GRS ecosystem.

The Asian Silurid catfish, *Wallago attu* (Bloch and Schneider, 1801; Figure 1) is widely distributed in the Indian subcontinent and several Southeast Asian countries including Thailand, Laos, Cambodia, Vietnam, Malaysia, and Indonesia (Roberts 1993). *W. attu* is a fast growing catfish that belongs to the Siluridae family and can reach a maximum length of 2 meter weighing more than 45 kg (Talwar and Jhingran, 1991; Roberts 2014).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(Higham and Kijnga). 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 (Mishra et al. 2009, Patra et al. 2005, Baran et al. 2018), and hence categorized as Vulnerable in the IUCN Redlist.

An understanding of the genetic background is essential for the conservation and management of wild species. For this purpose, molecular markers have been extensively used in population genetic studies (Sun et al., 2012; Zhao et al., 2013; Wei et al., 2013). Mitochondrial DNA is inherited maternally and has proven to be a useful molecular marker in studying fish populations due to its small size, relatively rapid base substitution rate, and lack of recombination (Guo et al., 2004). Furthermore, the varying evolution rate in different regions makes mtDNA suitable for research at various levels of evolution. The

Cytochrome *b* (Cyt *b*) and cytochrome oxidase subunit *I* (COI) genes are protein-coding regions within the mtDNA genome and typically exhibit a moderate level of intraspecific variation, hence used widely for investigating genetic diversity, phylogeography, and population genetics (Hebert et al., 2003; Mandal et al., 2012; Suneetha et al., 2000). Whereas, the control region (CR), also known as the D-loop, is a noncoding region and is recognized as the most variable part of the mtDNA genome. It commonly exhibits high intraspecific variation, making it suitable for studying genetic variability, inter and intraspecific levels variations and stock assessments (Donaldson and Wilson, 1999; Liang et al., 2011; Zhao et al., 2013). To date, limited research has been conducted on the genetic characteristics of *W. attu*, focusing mainly on the species' barcoding and phylogeny from the Indus river (Sajjad et al., 2023), and biology (Wu et al., 1999; He and Chen, 2007; Chen et al., 2009). 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 assess the genetic status and population structure of the species from tributaries of the GRS to cover a wide range of natural habitats threatened with anthropogenic alteration.

As such, the key objective of this study is to assess the genetic diversity, differentiation, population genetic structure, and demographic patterns of *W. attu*. To achieve this goal, we utilized three specific mtDNA regions, Cyt *b*, COI, and the control region to analyze the genetic characteristics of *W. attu* populations present in selected tributaries of the GRS. The information gathered through this research will provide fundamental knowledge necessary for the management of the species' genetic resources and the development of long-term conservation strategies.

2.2 Methodology

A total 176 samples of *W. attu* were collected from 19 sites covering seven tributaries of the Ganga river system namely Ganga (n=102; Seven site L1 to L7); Ramganga (n=5; L8); Sharda (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) (Figure 2). A comprehensive list of the sampling sites, sample numbers, sample codes, and corresponding GPS coordinates are provided in Table 1. A small piece of the pectoral fin was carefully clipped from each sample and stored in 95% ethanol until genomic DNA extraction was performed using the phenol chloroform method (Sambrook and Russell, 2001). The concentration and quality of the extracted DNA were evaluated using 0.8% agarose gel electrophoresis.

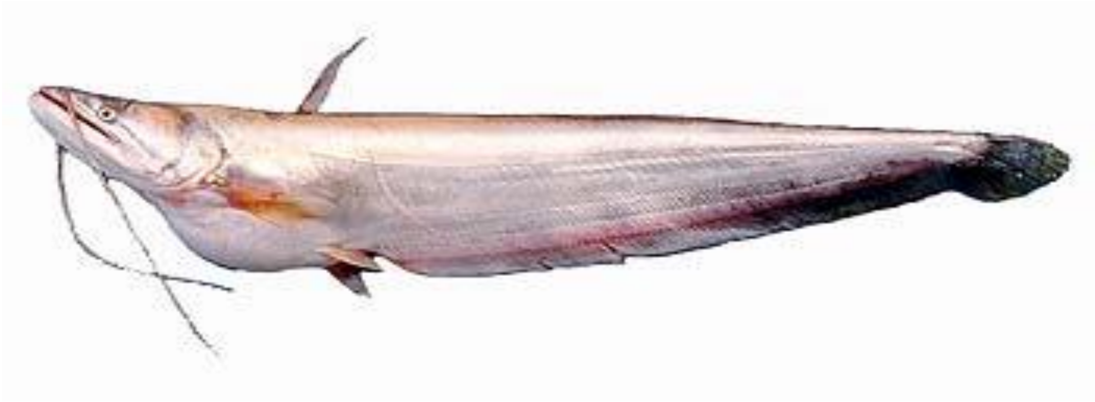


Figure 1: Image of Asian Silurid catfish, *Wallago attu* (Bloch and Schneider, 1801)

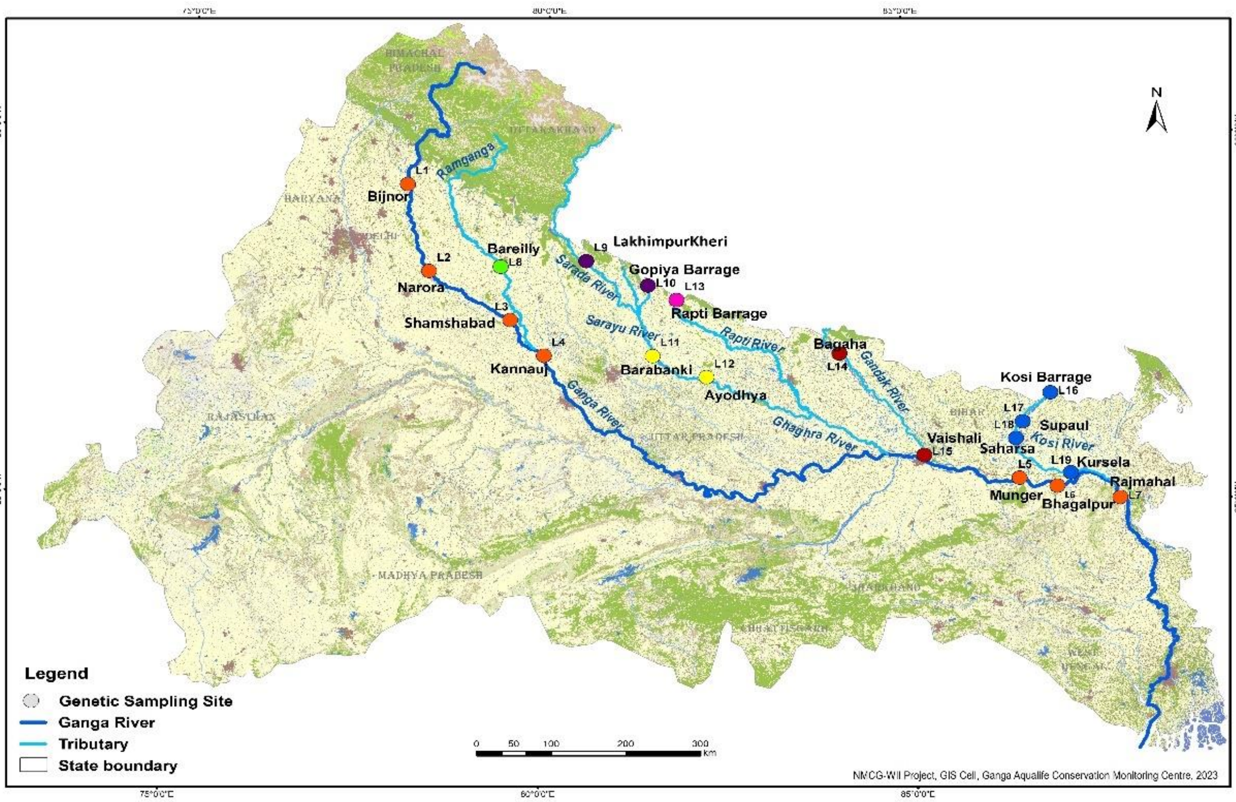


Figure 2: Sampling sites of *Wallago attu* samples in the Ganga river system

Table 1. Sample collection data of *Wallago attu* used in this study.

Rivers/Sites	Sites	Sample Code	Sampling Localities	Sample Size	Latitude (N)	Longitude (E)
Ganga	L1	BJ01-BJ23	Bijnor, Uttar Pradesh	23	29.391944	78.048056
Ganga	L2	NR01-NR24	Narora, Uttar Pradesh	24	28.216478	78.380189
Ganga	L3	FB01-FB10	Shamshabad, Uttar Pradesh	10	27.5656	79.518069
Ganga	L4	KJ01-KJ06	Kannauj, Uttar Pradesh	6	27.087931	79.990178
Ganga	L5	MG01-MG19	Munger, Bihar	19	25.369722	86.475278
Ganga	L6	BG01-BG07	Bhagalpur Bihar	10	25.245833	86.983611
Ganga	L7	JH01-JH10	Rajmahal, Jharkhand	10	25.058425	87.826489
Ramganga	L8	RG01-RG05	Bareillly, Uttar Pradesh	5	28.292617	79.370408
Sharda	L9	LK01-LK12	LakhimpurKheri, Uttar Pradesh	12	28.38331	80.5591
Sarada	L10	GB01-GB06	Gopiya Barrage, Bahraich	6	28.05411	81.41737
Sarayu/Ghaghra	L11	BK01-BK06	Barabanki, Uttar Pradesh	6	27.0929	81.4864
Sarayu/Ghaghra	L12	AD01-AD07	Ayodhya, Uttar Pradesh	7	26.80593	82.2264
Rapti	L13	RB01-RB07	Rapti Barrage, Uttar Pradesh	7	27.8607	81.81464
Gandak	L14	BH01-BH07	Bagaha, Bihar	7	27.117094	84.056972
Gandak	L15	HP01-HP05	Vaishali, Bihar	5	25.709764	85.188389
Kosi	L16	KB01-KB05	Kosi Barrage, Bihar	5	26.5225	86.934722
Kosi	L17	SP01-SP04	Supaul, Bihar	4	26.138056	86.543611
Kosi	L18	SH01-SH05	Saharsa, Bihar	5	25.911944	86.4455
Kosi	L19	KS01-KS05	Kursela, Bihar	5	25.421661	87.173464
			Total	**176**		

2.2.1 mtDNA amplification and sequencing

Polymerase Chain Reaction (PCR) was performed for all 176 samples to amplify three mtDNA regions, Cyt *b* using primers L14724 and H15915 (He and Chen, 2007),partial COI using primers Fish F1 and Fish R1 (Ward et al., 2005) and partial CR using primers Fish D-loopF and Fish D-loopR (Cheng et al., 2012) (Table 2). PCR was carried out in a final volume of 10µl reaction using a PCR buffer (10 mM Tri–HCl, pH 8.3, and 50 mM KCl), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.25 mM of each primer, 5 U of Taq polymerase and 1 µl of the template DNA (approximately 50ng). PCR thermal conditions were as follows: an initial denaturation at 95 C for 5 min, followed by 35 cycles of 95 C for 35 s, annealing temperature at 56˚C for 45 seconds, and extension at 72˚C for 75 seconds. The final extension was at 72˚C for 10 minutes. To monitor the effectiveness and consistency of the PCR reactions, positive controls were included. The amplified PCR products were run on a 2% agarose gel stained with ethidium bromide and visualized under UV light. To eliminate any residual primer, the amplified PCR products treated with Exonuclease I (EXO-I) and shrimp alkaline phosphatase (SAP) for 15 minutes each at 37˚C and 80˚C, respectively. Subsequently, the amplified PCR products were directly sequenced using the BigDye Terminator Kit (v3.1) and analyzed on an ABI 3500XL Applied Biosystems Genetic Analyzer. Both forward and reverse sequences were obtained for all products. The sequences were aligned and edited using Sequencer 4.7 software from Gene Code Corporation. Alignment of all raw sequences was performed using CLUSTAL W within the BioEdit v 7.2.5 software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

Table 2: Primes used for amplification and sequencing of *Wallago attu* samples

Gene	Primers	Reference
Cytochrome *b*	L14724: 5'-GACTTGAAAAACCACCGTTG-3' H15915: 5'-CTCCGATCTCCGGATTACAAGAC-3'	Xiao et al., 2001
Cytochrome c oxidase *I*	FishF1: 5'-TCAACCAACCACAAAGACATTGGCAC-3' FishR1: 5'-TAGACTTCTGGGTGGCCAAAGAATCA-3'	Ward et al., 2005
Control Region	Fish Dloop F, 5'- AGCACCGGTCTTGTAAACCG-3' Fish Dloop R, 5'- CTCCGGTTTGAACTCAGATC -3'	Cheng et al., 2012

2.3 Data analysis

The sequences of Cyt*b*, COI and CR were derived from the forward and reverse directions and edited using SEQUENCHER version 4.9 (Gene Codes Corporation, Ann Arbor, MI, USA). The alignment and analysis of each region was performed separately using the CLUSTAL X multiple sequence alignment program (Thompson et al., 1997), and the sequences were examined by visual inspection. Further, concatenate dataset of three mtDNA regions was generated to analyze the haplotype diversity (h), nucleotide diversity (p), and polymorphic sites (s) using DnaSP 5.0 (Librado and Rozas, 2009). The spatial distribution of the haplotypes was visualized through a median-joining (MJ) network, which was created using the PopART software (Leigh and Bryant, 2015). To determine whether the *W. attu* populations carried a signal of spatial range expansion or a stationary population history, Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1989) neutrality test was performed in DnaSP 5.0 (Librado and Rozas, 2009). To generate the trends in spatial demography history, the mismatch analysis was carried out using the population growth-decline model in DnaSP (Librado and Rozas, 2009). To evaluate the fit of the observed distribution the sum of squared deviations (SSD), the raggedness index (r) under the growth-decline model for each population, genetic differentiation (FST values) and Analysis of Molecular Variance (AMOVA) were calculated in ARLEQUIN v3.5 program (Excoffier and Lischer, 2010). The P-values were obtained from 1000 simulations on the basis of a selective neutrality test. Phylogenetic analyses were conducted in BEAST ver 1.7 (Drummond et al., 2012). 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 BEASTver 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 fit substitution model was determined using MrModeltest Version 3.7 (Posada and Crandall 1998) based on the Akaike Information Criterion (AIC). The chosen model was the Hasegawa-Kishino-Yano model (HKY) with gamma distribution for (G). To determine the effective population size over time coalescent Bayesian skyline plots (Drummond et al. 2012) were generated using BEAST v1.10.4 (Drummond et al. 2012) and visualized in Tracer version 1.7 (Rambaut et al. 2018); 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).

2.4 Results

2.4.1 Sequence variations, haplotype distribution and genetic structure

We generated the sequences of three mtDNA regions consisting of Cyt b: 1054bp, COI: 632bp and CR: 734bp from 176 samples. To perform the sequence variations and gene diversity estimates, the concatenated dataset of three mtDNA region (Cytb+COI+CR: 2420bp) was generated. In the whole dataset, eighty polymorphic sites were identified (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 and the generated sequences have been submitted to GenBank. 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 andHap68 shared with Sharda and Sarayu/Ghaghra. Interestingly, no haplotype sharing was observed in Ramganga, Kosi and Rapti and also each population had its own private haplotypes (Table 3). The median-joining network of the 103 haplotypes represented the distribution pattern of haplotypes among the *W. attu* populations (Figure 3). 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 provide valuable insights into the genetic relationships among populations in the region, suggesting complex interactions and potentially unique evolutionary processes of *W. attu* in the Ganga river basin. In phylogenetic analysis, the result was consistent with MJ network and three clades were formed (Figure 4). 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.

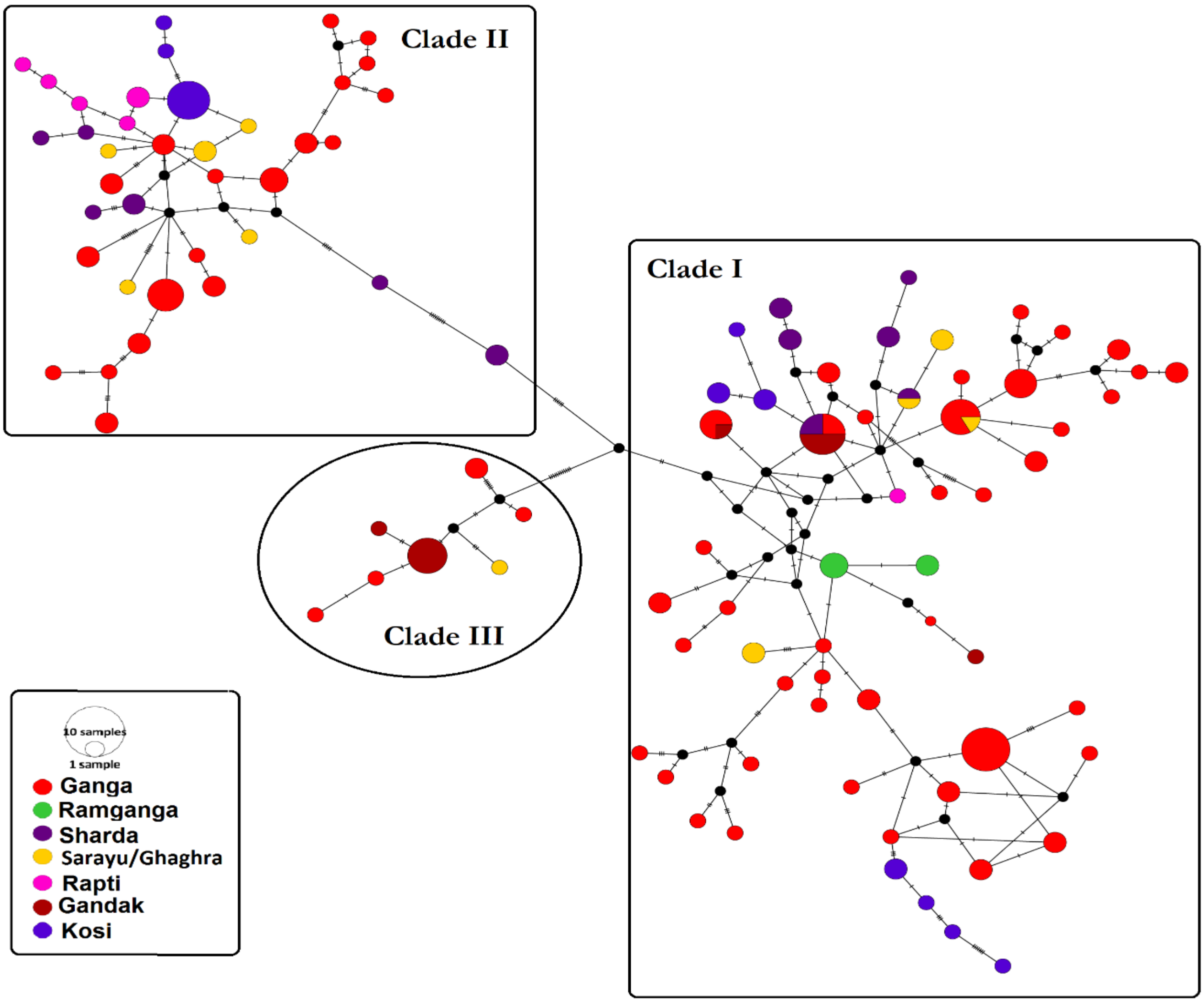


Figure 3. The median-joining network based on the concatenate dataset of three mtDNA regions COI, Cyt b and Control region of *Wallago attu*. The number of lines (bar) between nodes represents the mutation sites, and the size of the circle represents the number of individuals in each haplotype.

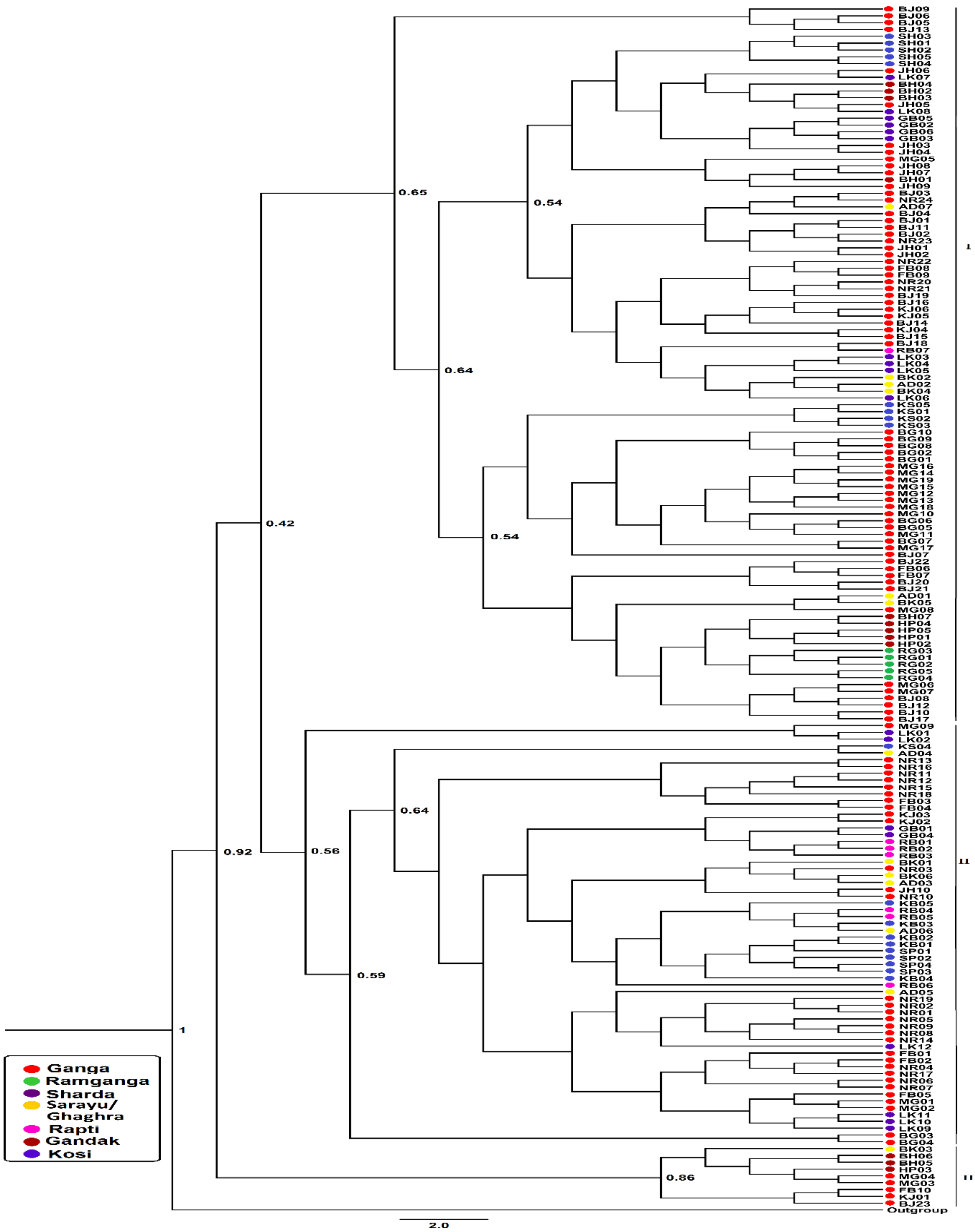


Figure 4. Bayesian phylogenetic tree for *Wallago attu* based on concatenated mitochondrial regions (COI, Cyt b, Control region). Samples collected in this study have symbols with colours corresponding to sampling rivers. Numbers next to nodes indicate their respective posterior probability (PP).

Table 3: Haplotype frequencies in *Wallago attu* populations in Ganga river system							
Haplotype:	**Ganga**	**Ramganga**	**Sharda**	**Saryu/Ghaghra**	**Rapti**	**Gandak**	**Kosi**
No of samples	**102**	**5**	**18**	**13**	**7**	**12**	**19**
Hap 1	5			1			
Hap 2	1						
Hap 3	2						
Hap 4	1						
Hap 5	1						
Hap 6	1						
Hap 7	1						
Hap 8	1						
Hap 9	1						
Hap 10	1						
Hap 11	1						
Hap 12	1						
Hap 13	4						
Hap 14	1						
Hap 15	1						
Hap 16	1						
Hap 17	1						
Hap 18	1						
Hap 19	1						
Hap 20	1						
Hap 21	3						
Hap 22	1						
Hap 23	4						
Hap 24	1						
Hap 25	2						
Hap 26	2						
Hap 27	2						
Hap 28	1						
Hap 29	1						
Hap 30	1						
Hap 31	1						
Hap 32	1						
Hap 33	1						
Hap 34	1						
Hap 35	2						
Hap 36	1						
Hap 37	2						

Hap 38	1						
Hap 39	1						
Hap 40	1						
Hap 41	2						
Hap 42	2						
Hap 43	2						
Hap 44	2						
Hap 45	1						
Hap 46	1						
Hap 47	1						
Hap 48	2						
Hap 49	1						
Hap 50	1						
Hap 51	1						
Hap 52	1						
Hap 53	9						
Hap 54	2						
Hap 55	2						
Hap 56	2						
Hap 57	1						
Hap 58	2						
Hap 59	2						
Hap 60	2						
Hap 61	2		2			3	
Hap 62	3					1	
Hap 63		3					
Hap 64		2					
Hap 65			2				
Hap 66			2				
Hap 67			1				
Hap 68			1	1			
Hap 69			1				
Hap 70			2				
Hap 71			1				
Hap 72			1				
Hap 73			2				
Hap 74			2				
Hap 75			1				
Hap 76				1			
Hap 77				1			
Hap 78				2			
Hap 79				2			

Hap 80				2			
Hap 81				1			
Hap 82				1			
Hap 83				1			
Hap 84					1		
Hap 85					1		
Hap 86					1		
Hap 87					2		
Hap 88					1		
Hap 89					1		
Hap 90						2	
Hap 91						4	
Hap 92						1	
Hap 93						1	
Hap 94							1
Hap 95							1
Hap 96							7
Hap 97							2
Hap 98							1
Hap 99							2
Hap 100							2
Hap 101							1
Hap 102							1
Hap 103							1

1.4.2 *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. The overall Hd value was high 0.947 ± 0.006 across all populations, however, π value was found to be low 0.0062 (Table 4). Pairwise Fst comparison of populations showed significant genetic differences between the populations. Comparatively low Fst between Ganga and Sarayu/Ghaghra (0.047) to high between Ramganga and Rapti (0.735) was observed (Figure 5). 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 5).

Table 4: 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.

Table 5. AMOVA analyses of mtDNA sequences for seven populations of *Wallago attu* from Ganga river basin

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			

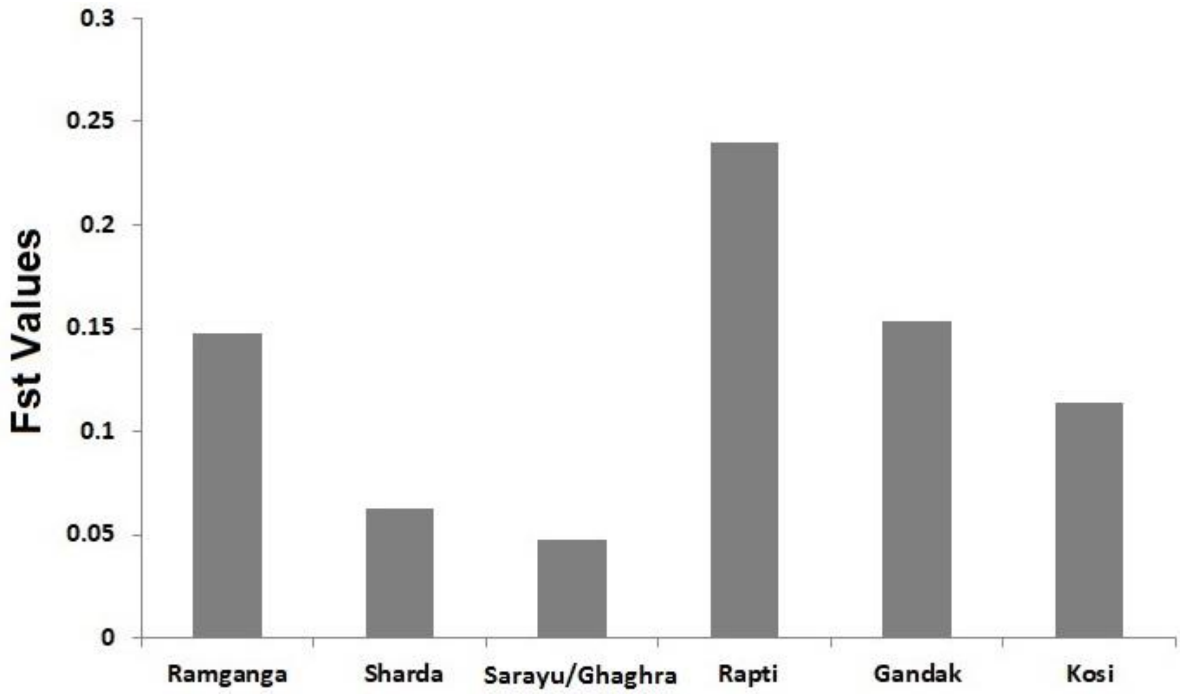


Figure 5: Genetic differentiation among the *Wallago attu* populations. The pairwise genetic distance values based on the concatenate dataset of three mtDNA regions COI, Cyt *b* and Control region.

2.4.3 Population demography

The neutrality test of Tajima's D and Fu's Fs tests and mismatch distributions analysis were carried out to deduce the demographic history of the *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 (Figure 6A and Table 2). The demographic scenario was also supported by the generalized least square procedure and the raggedness index of the distribution in studied rivers (Table 4). 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 a pronounced population expansion at approximately 0.04 Ma (Figure 6B).

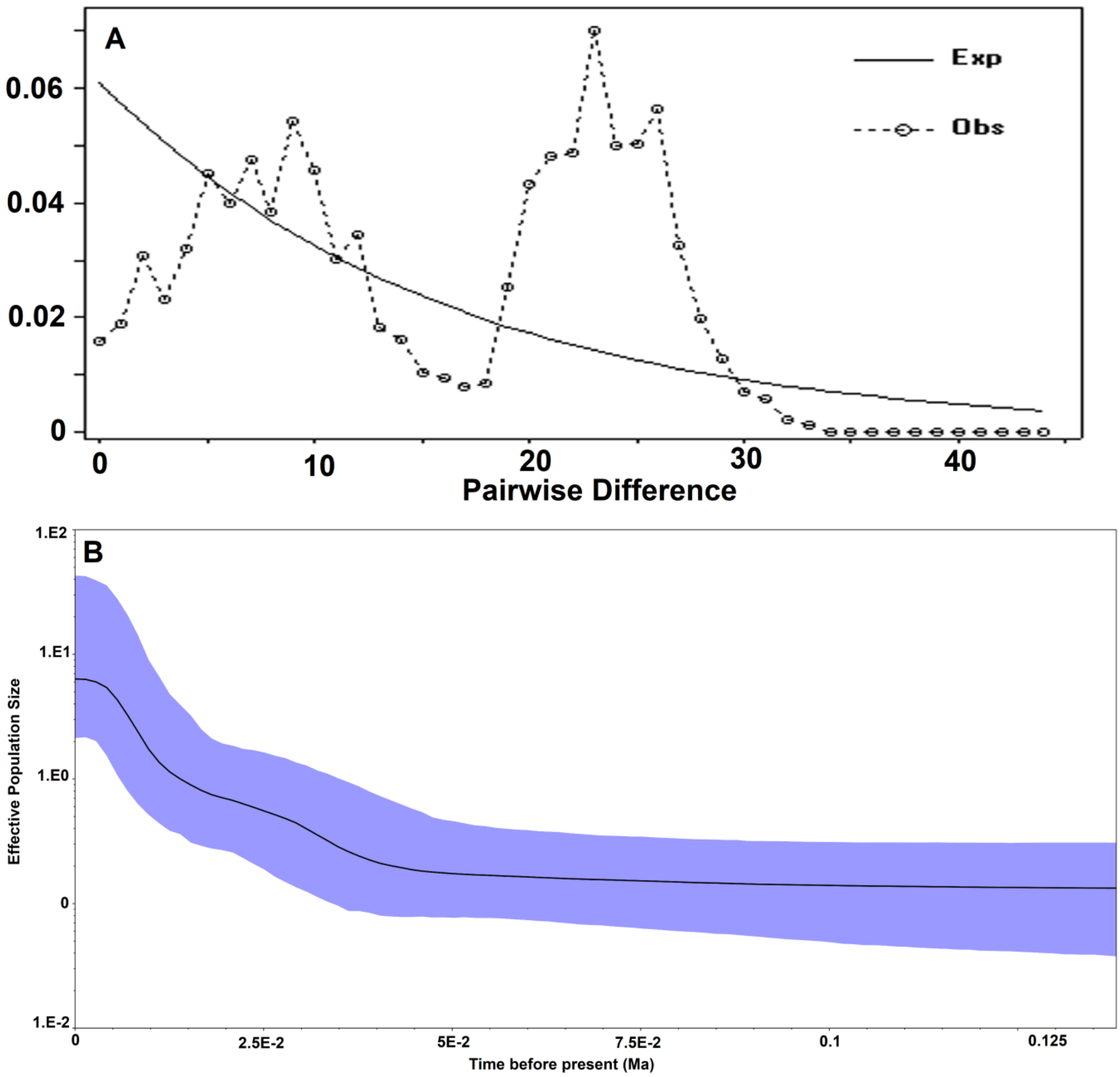


Figure 6: (A) Mismatch distribution graph for overall populations of *Wallago attu* in the Ganga river basin. The X and Y axis show the number of pairwise differences and the frequency of the pairwise comparisons, respectively. The observed frequencies are represented by dotted line. The frequency expected under the hypothesis of constant population model is depicted by solid line. (B) Bayesian Skyline plot showing 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 of females; τ = generation time), and the shaded portion around the solid median estimates show the 95% highest posterior density estimate of the historic effective population size.

2.5 Discussion

One of largest river ecosystem in the Asia, the GRS is characterized by high species richness and has been identified as one of the important centers of biodiversity (Sarkar et al., 2012). The GRS has been identified as a biodiversity hotspot, requiring immediate protection to preserve its rich variety of species. The region faces numerous anthropogenic threats, making conservation efforts urgent and crucial to safeguard its unique biodiversity. Hence, understanding the stock composition and population dynamics of exploited species is of utmost importance in ensuring sustainable fisheries management (Ovenden et al., 2015). The present study provides the first detailed insights into the genetic characteristics of *W. attu* and significantly expand our knowledge of their demography and levels of differentiations in GRS.

2.5.1 Genetic diversity and population structure

Investigating the mtDNA variations indicated high haplotype diversity with low nucleotide diversity within seven populations of *W. attu*. A high level of haplotype diversity combined with low nucleotide diversity often indicates that the population likely experienced a period of low effective population size followed by a significant expansion. *W. attu*, a widely distributed and frequently observed fish species in GRS, the large population size may account for the high levels of 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, coupled with significant variations in *W. attu* populations, could be attributed to the diverse environmental conditions and distinct selective pressures across different river systems (Modeel et al. in 2023). Furthermore, the rapid increase in population size may have facilitated the accumulation of mutations, and the period of expansion may have allowed ample time for the accumulation of their haplotype diversity, but insufficient for an increase in nucleotide diversity (Guo et al. in 2014).

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*. This indicates 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* (Guo et al., 2014), *Ancherythroculter nigrocauda* in the upper Yangtze River (Zhai et al., 2019), *Tor putitora* from the Upper Ganga region in India (Yadav et al., 2020), and the *Pethia* genus in Sri Lanka (Sudasinghe et al., 2021).

2.5.2 Genetic distance

Genetic distances between seven populations indicated that genetic differentiation ranged from low to high. As the distance between tributaries increase, so does the genetic differentiation among the populations. In general, river currents play an important role in dispersing the fishes resulting in low genetic divergences with reference to 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, whilst there was moderate differentiation among populations ($P < 0.001$). The level of genetic differentiation, and AMOVA showed consistent results, which was 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.

2.5.3 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. Typically, significantly negative values of D and Fs can indicate either population expansion or purifying selection (Liu et al. 2009; Ren et al. 2017). In this analysis, both 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 the analysis of the Bayesian Skyline Plot (BSP), 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* (Guo et al., 2014). 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 (Dynesius and Jansson, 2000). It is plausible that these recurrent glacial-interglacial changes within the Ganges River system 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.

2.6 Conclusion

Gaining insights into the genetic diversity and population structure holds the potential to enhance both the management of fisheries and the 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 revelation of high population genetic diversity, coupled with moderate genetic differentiation between populations, indicates the initiation of divergence within the *W. attu* species 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 the assessment of 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.

2.7 References

- Chen, F., Chen, Y.F., He, D.K., 2009. Age and growth of *Schizopygopsis younghusbandi younghusbandi* in the Yarlung Zangbo River in Tibet, China. Environ. Biol. Fish. 86, 155e162
- Cheng, Y.Z., Xu, T.J., Jin, X.X., Tang, D., Wei, T., Sun, Y.Y., Meng, F.Q., Shi, G. and Wang, R.X., 2012. Universal primers for amplification of the complete mitochondrial control region in marine fish species. Molecular Biology, 46, pp.727-730.
- Donaldson, K.A., Wilson, R.R., 1999. Amphi-Panamaic geminates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rates in the mitochondrial DNA control region of fishes. Mol. Phylogenet. Evol. 13, 208e213.
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969–1973.
- 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
- Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linuxand Windows. Mole Ecol Resour.10, 564–567 (2010).
- Fu, Y. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147,915–925 (1997).
- Giri, S.S., Sahoo, S.K., Sahu, B.B., Sahu, A.K., Mohanty, S.N., Mukhopadhyay, P.K. and Ayyappan, S., 2002. Larval survival and growth in *Wallago attu* (Bloch and Schneider): effects of light, photoperiod and feeding regimes. Aquaculture, 213(1-4), pp.151-161.
- Guo, S.S., Zhang, G.R., Guo, X.Z., Wei, K.J., Yang, R.B. and Wei, Q.W., 2014. Genetic diversity and population structure of *Schizopygopsis younghusbandi* Regan in the Yarlung Tsangpo River inferred from mitochondrial DNA sequence analysis. Biochemical Systematics and Ecology, 57, pp.141-151.
- Guo, X.H., Liu, SJ., Liu, Q, Liu, Y., 2004. New progress on mitochondrial DNA in fish. Acta Genet. Sin. 31, 983e1000
- Gupta, S.D., Reddy, P.V.G.K., Natrajan, E., Sar, U.K., Sahoo, S.K., Rath, S.C., Dasgupta, S., 1992. On the second breeding and some aspect of culture of *Wallago attu* (Schneider). J. Aquacult. 2, 1 – 6.
- 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. https:// doi.org/ 10.1098/ rsbl. 2003. 0025.
- Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype network construction. Methods Ecol Evol 6(9):1110– 1116.
- Li WLS, Drummond AJ. 2012. Model Averaging and Bayes Factor Calculation of Relaxed Molecular Clocks in Bayesian Phylogenetics. Mol Biol Evol. 29:751–761.
- Liang, J.J., Liu, Y., Zhang, X.F., Zhang, X.Y., Yue, B.S., Song, Z.B., 2011. An observation of the loss of genetic variability in prenant's schizothoracin, *Schizothorax prenanti*, inhabiting a plateau lake. Biochem. Syst. Ecol. 39, 361e370.
- Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451e1452.
- Liu, J.Y., Brown, C., Yang, T.B., 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.

- 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. https:// doi.org/ 10. 1007/ s11033- 011- 0917-9.
- Mishra, S. S., Acherjee, S. K. and Chakraborty, S. K. 2009. Development of tools for assessing conservation categories of siluroid fishes of fresh water and brackish water wetlands of South West Bengal, India. Environmental Biology of Fishes 84(4): 395–407.
- Modeel, S., Joshi, B.D., Yadav, S., Bharti, M. and Negi, R.K., 2023. Mitochondrial DNA reveals shallow population genetic structure in economically important Cyprinid fish *Labeo rohita* (Hamilton, 1822) from South and Southeast Asia. Molecular Biology Reports, 50(6), pp.4759-4767.
- Ng, H.H., de Alwis Goonatilake, S., Fernado, M. & Kotagama, O. 2019. Wallago attu (errata version published in 2020). The IUCN Red List of Threatened Species 2019: e.T166468A174784999.
- 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 and Fisheries 16(1):125–159.
- Patra, M. K., Acharjee, S. K. and Chakraborty, S. K. 2005.Conservation categories of siluroid fishes in North-East Sundarbans, India. Biodiversity and Conservation 14: 1863–1876.
- Posada, D., Crandall, K., 1998. Modeltest: testing the model of DNA substitution. Bioinform 14, 817–818.
- 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. https://doi.org/10.1093/sysbio/ syy032
- 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.
- Roberts, T. R. (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. Bulletin of the Peabody Museum of Natural History, 55(1), 35-47.
- Roberts, T.R. (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. Bulletin of the Peabody Museum of Natural History, 55 (1): 35-47.
- Sajjad, A., Jabeen, F., Ali, M. and 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.
- Sambrook, J., Russell, D.W., 2001. Molecular Cloning: a Laboratory Manual, third ed. Cold Spring Harbor Laboratory Press, New York.
- Sarkar, U.K., Pathak, A.K., Sinha, R.K., Sivakumar, K., Pandian, A.K., Pandey, A., Dubey, V.K. and Lakra, W.S., 2012. Freshwater fish biodiversity in the River Ganga (India): changing pattern, threats and conservation perspectives. Reviews in Fish Biology and Fisheries, 22, pp.251-272.
- Singh M and Singh AK (2007).Bibliography of environmental studies in natural characteristics and anthropogenic influences on the Ganga River. Environmental Monitoring and Assessment, 129, pp.421-432.
- 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. https:// doi. org/ 10. 1186/ s12862- 021- 01923-5.

- Sun, P., Shi, Z.H., Yin, F., Peng, S.M., 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, 54e63.
- Suneetha KB, Dahle G, N.vdal G (2000) Analysis of mitochondrial DNA sequences from two *maurolicus* taxa: evidence forseparate species? J Fish Biol 57(6):1605–1609. https:// doi.org/ 10.1111/j. 1095- 8649. 2000. tb022 36.x
- Tajima, H. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595 (1989).
- Talwar PK, Jhingran AG (1991) Inland fishes of India and adjacent countries. CRC Press, Florida, United States
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876.
- Ward R.D., Zemlak T.S., Innes B.H., Last P.R., Hebert P.D.N. 2005. DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society, Part B: Biological Sciences 360 (1462): 1847–1857. DOI: 10.1098/rstb.2005.1716.
- Wei, K.J., Wood, A.R., Gardner, J.P.A., 2013. Population genetic variation in the New Zealand greenshell mussel: locus-dependent conflicting signals of weak structure and high gene flow balanced against pronounced structure and high self-recruitment. Mar. Biol. 160, 931e949
- Wu, Y.F., Kang, B., Men, Q., Wu, C.Z., 1999. Chromosome diversity of Tibetan fishes. Zool. Res. 20, 258e264. Wu, Y.F., Wu, C.Z., 1992. The Fishes of the Qinghai-xizang Plateau. Sichuan Publishing House of Science and Technology, Chengdu
- Xiao, W., Zhang, Y. and Liu, H., 2001. Molecular systematics of Xenocyprinae (Teleostei: Cyprinidae): taxonomy, biogeography, and coevolution of a special group restricted in East Asia. Molecular phylogenetics and evolution, 18(2), pp.163-173.
- 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. https:// doi. org/ 10. 1371/ journal. pone. 02343 77.
- 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. https:// doi. org/ 10. 1111/ jai. 13897.
- Zhao, L.J., Zhou, X.Y., Liu, Q.G., Zhang, H., 2013. Genetic variation and phylogeography of Sinibrama macrops (Teleostei: Cyprinidae) in Qiantang River, China. Biochem. Syst. Ecol. 49, 10e20.

3. MOLECULAR EVIDENCE INDICATES THE EXISTENCE OF MULTIPLE LINEAGES OF *SPERATA SPECIES* IN INDIAN RIVERS

3.1 Background

The genus *Sperata* is one of the largest catfish that belongs to family Bagridae and widely distributed in India, Pakistan, Bangladesh, Afghanistan, and Nepal (Talwar & Jhingharan, 1991). The genus *Sperata* was previously known as *Aorichthys*/*Mystus* and was recognized as one of the commercially important catfish species (Saigal & Motwani, 1961; Froese & Pauly, 2011). Due to high nutritional values and the low number of intramuscular bones, *Sperata* species have a huge demand in South Asian countries (Mohanty *et al.,* 2012). Based on the morphological characteristics, four species of *Sperata* are currently recognized from Asia (Day, 1878; Talwar & Jhingharan, 1991; Ferraris & Runge, 1999; Ng & Kottelat, 2013). These four species are named as, *S. seenghala* (Sykes, 1839), *S. aor* (Hamilton, 1822), *S. acicularis* (Ferraris & Runge, 1999), and *S. aorella* (Blyth, 1858). Recently, there are one more species, *S. sarwari,* which is recognized from the Indus river system of India and Pakistan (Nawaz *et al*., 1994). The catfish can tolerate a high range of temperature (38-40 °C), low dissolved oxygen, and salinity along with variable water conditions, because of their robust air-breathing system (Yadav *et al*., 2017; 2018). Among the four species of the genus *Sperata*, *S. seenghala* (Fig. 7) (the giant river-catfish) and *S. aor* (Fig. 8) (the long-whiskered catfish) have a wide distribution in the Indian subcontinent (Gupta, 2015). In the Ganga river, due to the high anthropogenic activities such as extensive fishing, water abstraction, pollution, siltation, and invasion of exotic species, are threatening the *Sperata* populations. *S. seenghala* and *S. aor* are recognized as "least concern" in the RedList of International Union for Conservation of Nature (IUCN). Previous studies on giant river catfish gradually focused on the genetic assessment of *S. seenghala* using Randomly amplified polymorphic DNA (RAPD) marker (Saini *et al*., 2010; Garg *et al*., 2014), mtDNA markers (Kumari *et al*., 2017), and microsatellite marker (Acharya *et al*., 2019). The microsatellite-based stock discrimination of *S. aor* revealed the existence of three genetic stocks from the connected vicinity of the Ganga river (Nazir & Khan, 2017). Though both the species of *Spereta* are abundantly found in Indian rivers; no study has been done so far to differentiate them.

Day (1878) documented the morphological characteristics for the identification of these catfish species. In *S. aor* maxillary barbels are longer in size and prolonged or beyond to the base of the caudal fin. It poses a supraoccipital spine, and its interneural shield shares almost the same length. It is also characterized by 10-11 pectoral-fin rays, 19-20 gill rakers, and its orbit prolongs through the middle of the length of the head. However, in *S. seenghala*, the snout is chisel-shaped; maxillary barbels are not extended beyond the middle body; and supraoccipital spine shorter than interneural shield. It posses 8-9 pectoral-fin rays, 13-15 gill rakers, and its orbit completely present in the head's anterior half (Gupta, 2015; Miyan *et al*., 2016). Despite being significant variations, the identification between these two species is quite difficult. Hence, reliable sampling is essential for resolving the population genetic structure and diversity. Therefore, an appropriate management plan for a different stock can be prepared by incorporating reliable knowledge about existing lineages from species distribution ranges. Due to the maternal mode of transmission, and more rapid evolution than the nuclear genome, mitochondrial DNA (mtDNA) provides reliable information on the relationships among closely related species and populations (Brown *et al.,* 1979). In particular, the mtDNA control region (CR) is a highly variable portion and estimated to be five times more substitution rate than that of the rest sequences in mtDNA (Aquadro & Greenburg, 1983). For this reason, this marker is widely acceptable for population genetics studies (Wan *et al*., 2004; Gupta *et al*., 2018).

We used *S. seenghala* and *S. oar* from the Ganga river to improve our knowledge of contemporary genetic relationships, diversity, and demography of *Sperata*. Then we compared it with *S. seenghala* populations of Mahanadi and Brahmaputra. We used wide-range geographic sampling over different locations of the Ganga River and DNA sequences of other major rivers for comprehensive genetic insight into *Spereta* lineages.

Figure 7: *Sperata seenghala* collected from Bijinor Uttar Pradesh

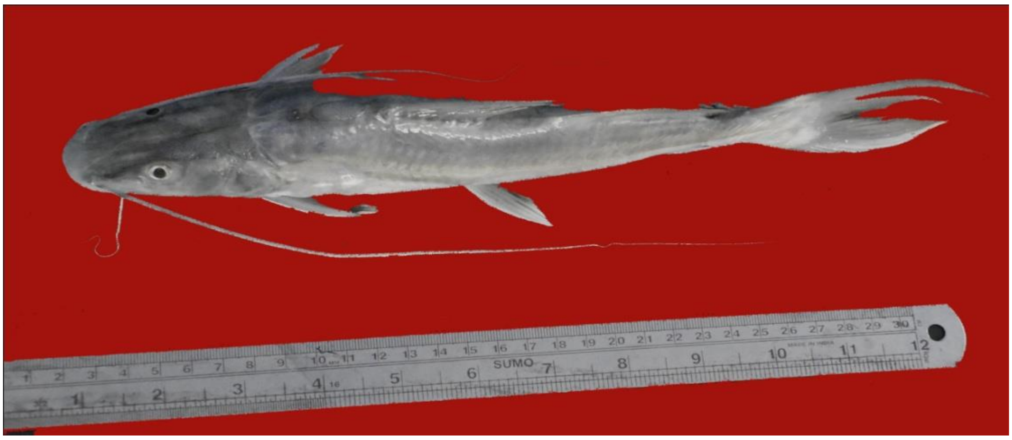


Figure 8: ***Sperata aor*** **collected from Bijinor Uttar Pradesh**

3.2 Methodology

A small portion of fin tissues was collected from *S. seenghala* (n=131) from eight locations and *S. aor* (n=78) from four locations of the Ganges river and stored in 70% ethanol at 4°C. (Table 6, Fig. 9). Total genomic DNA was extracted from the tissue samples using the phenol-chloroform extraction protocols with a final elution volume of 100 µl (Sambrook *et al*., 1989). The extracted DNA was checked on 0.8% agarose gel and quantified in QIAxcel and diluted in a final concentration of 30ng/µl for PCR amplification.

Table 6: Sampling locations of ***S. seenghala*** **and** ***S. aor*** **used in this study. N, number of samples.**

Rivers	Species	Sites in map	Sampling sites	N	Latitude	Longitude
Ganga	*S. seenghala*	1	Bijinor, Uttar Pradesh	21	29.3919	78.0480
	S. seenghala	2	Tigri Ghat, Uttar Pradesh	15	28.8164	78.1461
	S. seenghala	3	Narora, Uttar Pradesh	20	28.2164	78.3800
	S. seenghala	4	Samshabad, Uttar Pradesh	18	27.5655	79.5180
	S. seenghala	5	Kanpur, Uttar Pradesh	15	26.4765	80.3683
	S. seenghala	6	Allahabad, Uttar Pradesh*	25	25.45	81.85
	S. seenghala	7	Varanasi, Uttar Pradesh	12	25.3225	83.1344
	S. seenghala	8	Munger, Bihar	16	25.2229	86.2748
	S. seenghala	9	Sahebganj, Jharkhand	14	25.2511	87.6600
Ganga	*S. aor*	1	Bijinor, Uttar Pradesh	41	29.3919	78.0480
	S. aor	2	Tigri Ghat, Uttar Pradesh	11	28.8164	78.1461
	S. aor	3	Narora, Uttar Pradesh	12	28.2164	78.3800
	S. aor	7	Varanasi, Uttar Pradesh	14	25.3225	83.1344
Bramhaputra	*S. seenghala*	10	Raha, Assam*	25	26.35	92.69
Mahanadi	*S. seenghala*	11	Sambalpur, Odisha*	25	21.47	83.97

Note: *Sequences taken from Kumari et al., (2017)

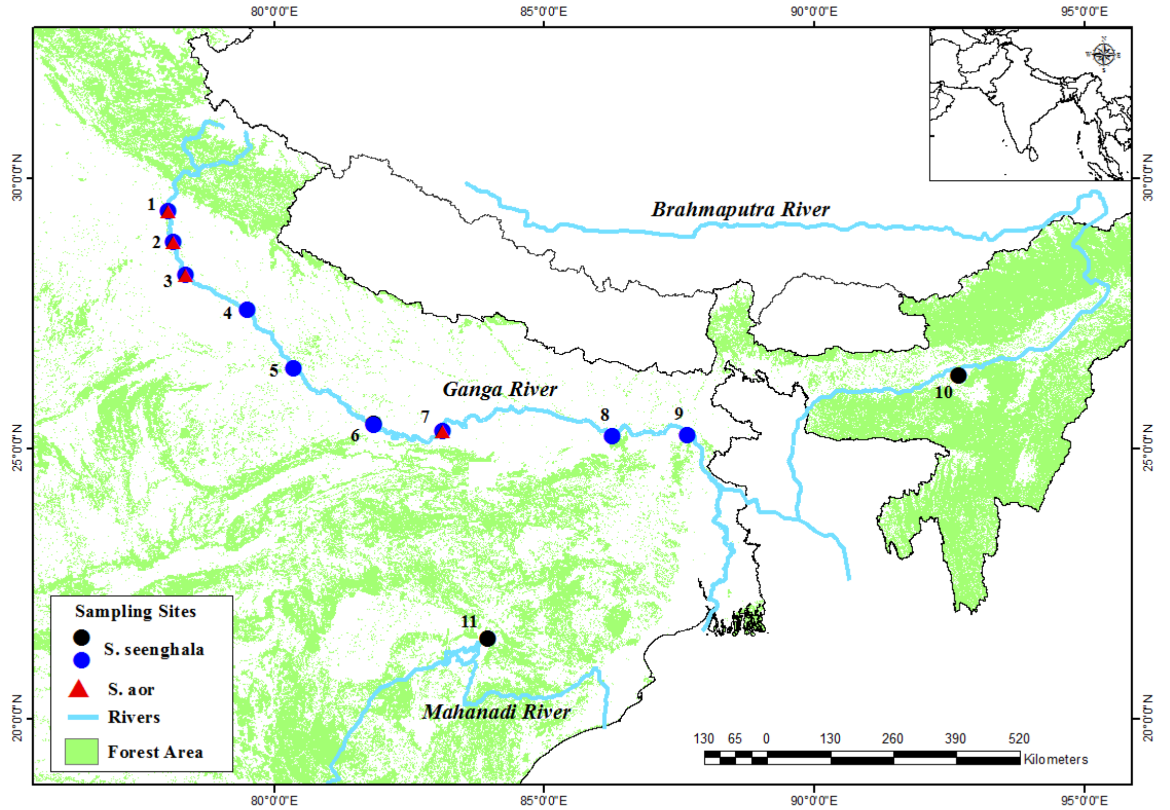


Figure 9:The Ganga, Brahmaputra and Mahanadi river system in India and the geographic origin of *Spereta seenghala* and *S. aor* samples used in this study.

3.2.1 *PCR amplification and sequencing*

PCR amplification was performed using mtDNA control region (CR) primers SSDLoop F:5'-CACCCCTAACTCCCAAAGC-3', and SSDloop R:5'-GGTTTAGGGGTTTAACAGG-3'(Kumari *et al.*, 2017). PCR reaction was performed in total reaction volumes of 20 µl using a 1 X PCR buffer (10 mM Tri –HCl, pH 8.3, and 50 mM KCl), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 2 pmol of each primer, 5 U of Taq DNA polymerase and 1 µl (~30 ng) of the template DNA. The PCR conditions were as follows: 95°C for 5 minutes, followed by 35 cycles of denaturation at 95°C for 45 seconds, annealing at 58°C for 45 seconds, and extension at 72°C for 75 seconds. The final extension was at 72°C for 10 minutes. The consistency of the PCR amplification was monitored using positive and negative controls. The PCR amplicons were checked on 2% agarose gel by gel electrophoresis and visualized under the UV transilluminator. Exonuclease I (EXO-*I*) and shrimp alkaline phosphatase (SAP) treatments were given to the amplified PCR products (USB, Cleveland, OH) for 15 min at 37ºC and 20 min at 80ºC, to degrade any residual primer and dNTPs. The amplified PCR products were sequenced from both forward and reverse directions using the BigDye® Terminator kit (v3.1) and analyzed on an Applied Biosystems Genetic Analyzer ABI 3500XL.

3.3 Data analysis

All the raw sequences obtained from the forward and reverse directions were aligned and edited with SEQUENCHER®version 4.9 (Gene Codes Corporation, Ann Arbor, MI, USA) to get the consensus sequence. All the consensus sequences were aligned using the program CLUSTAL X multiple sequence alignment (Thompson *et al*., 1997) and examined by visual inspection. DnaSP 5.0 (Librado & Rozas, 2009) was used to analyze the haplotype diversity (h), nucleotide diversity (p), and polymorphic sites (s). The numbers of nucleotide substitutions per site were estimated for multiple substitutions using the Tamura-3 parameter method in MEGA X (Kumar *et al*., 2018). For the genetic distance, we used the Tamura-3 parameter using a discrete Gamma distribution (TN92+G) with the lowest BIC score value using MEGA X (Kumar *et al*., 2018). Phylogenetic analyses were conducted in BEAST ver 1.7 (Drummond *et al*., 2012). The spatial distribution of haplotypes was visualized by a median-joining network, was created with the PopART software (Leigh & Bryant, 2015). MEGA X (Kumar *et al*., 2018) was used to estimate within-group genetic distance and between-group mean distance between the populations of *S. seenghala* and *S. aor*. In addition, molecular variance analysis (AMOVA) was performed to test the genetic differentiation between geographical units. The significance values generated by AMOVA were tested by random permutations of sequences among populations. DnaSP 5.0 (Librado & Rozas, 2009) was used to generate the mismatch distribution plot for trends in spatial demography history of *S. seenghala* and *S.aor* populations. Besides, to determine the demography history of each population of both the species, we have performed neutrality test, Tajima's D (Tajima, 1989), Fu's Fs test (Fu, 1997), the sum of squared deviations (SSD) and Raggedness index (r) under the demographic expansion model for each population using the program Arlequin 3.5 (Excoffier & Lischer, 2010). The P-values were obtained by 1000 simulations based on a selective neutrality test.

3.4 Results

3.4.1 mtDNA control region sequence polymorphism and haplotype diversity

We obtained 870 bp sequences for CR from 131 samples of *S. seenghala* from eight different sites and 78 samples of *S. aor* from four sites of the Ganga river (Table 6). Besides, 25 sequences of *S. seenghala,* each from the Ganga river, Mahanadi river, and Brahmaputra river were included from NCBI GenBank (Table 6, Table 7). Therefore, a dataset of 156 sequences of *S. seenghala* from the Ganga river were used in the present analysis. The high number of segregating sites was found in the Brahmaputra (173) followed by Mahanadi (136), *S. seenghala* of Ganga (86), and *S. aor* of Ganga (42) population. When sequences of all three rivers were aligned with our data, all were grouped into 77 haplotypes (Table 7). Out of these, 37 haplotypes (Hap1 to Hap 37) were observed in *S. seenghala* from the Ganga, 28 haplotypes (Hap38 to Hap65) were found in *S. aor*, four haplotypes (Hap 66 to Hap 69) were observed in Mahanadi, nine haplotypes (Hap3, Hap70 to Hap77) were observed in Brahmaputra river. Interestingly, four haplotype (Hap34 to Hap37) from the Ganga river sustain distinct genetic signature and highly diverse than the sequences of *S. seenghala* from the Ganga. The newly identified haplotypes of *S. seenghala* and *S. aor* from the Ganga River were submitted in GenBank (MT707741- MT707801).

The haplotype diversity (hd) of *S. seenghala* and *S.aor* from the Ganges river were comparable and high 0.943 and 0.934, respectively. In contrast, the haplotype diversity of the Mahanadi and Brahmaputra rivers was 0.717 and 0.820, respectively. The nucleotide diversity (**π**) was ranging from 0.009 in *S.aor* to 0.049 in the Mahanadi population. The ratio of transition and transversion rate (Ts/Tv) was high in Mahanadi (94/49), whereas it was low in Brahmaputra 11/15 (Table 8).

Table 7: Haplotype distribution of *Sperata seenghala* and *Sperata aor* used in this study.						
Haplotypes	Sample Size	Species	River	Sample origin	Accession Number	Reference
Hap.1	1	*S. seenghala*	Ganga	Bijnor (1)	MT707741	This study
Hap.2	11	*S. seenghala*	Ganga	Bijnor (1), Tigarighat (1), Shamshabad (2), Kanpur (1), Allahabad*(6)	MT707742, KT022186, KT022189, KT022192, KT022198	This study & *Kumari et al., 2016
Hap.3	14	*S. seenghala*	Ganga and Brahmaputra	Bijnor (4), Narora (2), Shamshabad (5), Kanpur (1), Brahmaputra*(2)	MT707743, KT022220, KT306656	This study & *Kumari et al., 2016
Hap.4	2	*S. seenghala*	Ganga	Bijnor (2)	MT707744	This study
Hap.5	2	*S. seenghala*	Ganga	Bijnor (1), Narora (1),	MT707745	This study
Hap.6	16	*S. seenghala*	Ganga	Bijnor (2), Shamshabad (2), Kanpur (2), Allahabad*(2), Varanasi (3), Sahebganj (5)	MT707746, KT022193, KT022197	This study & *Kumari et al., 2016
Hap.7	1	*S. seenghala*	Ganga	Bijnor (1)	MT707747	This study
Hap.8	6	*S. seenghala*	Ganga	Bijnor (1), Tigarighat (4), Narora (1)	MT707748	This study
Hap.9	24	*S. seenghala*	Ganga	Bijnor (1), Tigarighat (5), Narora (6), Shamshabad (3), Kanpur (1), Allahabad*(5), Sahebganj (3)	MT707749, KT022185, KT022188, KT022194, KT022195, KT022196	This study & *Kumari et al., 2016
Hap.10	8	*S. seenghala*	Ganga	Bijnor (4), Tigarighat (3), Kanpur (1),	MT707750	This study
Hap.11	1	*S. seenghala*	Ganga	Bijnor (1)	MT707751	This study
Hap.12	4	*S. seenghala*	Ganga	Bijnor (2), Tigarighat (2)	MT707752	This study

Hap.13	7	*S. seenghala*	Ganga	Narora (6), Shamshabad (1)	MT707753	This study
Hap.14	3	*S. seenghala*	Ganga	Narora (3)	MT707754	This study
Hap.15	3	*S. seenghala*	Ganga	Narora (1), Kanpur (2)	MT707755	This study
Hap.16	2	*S. seenghala*	Ganga	Shamshabad (2)	MT707756	This study
Hap.17	2	*S. seenghala*	Ganga	Shamshabad (1), Kanpur (1)	MT707757	This study
Hap.18	2	*S. seenghala*	Ganga	Shamshabad (2)	MT707758	This study
Hap.19	2	*S. seenghala*	Ganga	Kanpur (2)	MT707759	This study
Hap.20	3	*S. seenghala*	Ganga	Kanpur (3)	MT707760	This study
Hap.21	1	*S. seenghala*	Ganga	Kanpur (1)	MT707761	This study
Hap.22	6	*S. seenghala*	Ganga	Allahabad*(3), Varanasi (3)	MT707762, KT306667 , KT306668, KT022180	This study & *Kumari et al., 2016
Hap.23	3	*S. seenghala*	Ganga	Allahabad (3)	KT306669, KT306670, KT022182	Kumari et al., 2016
Hap.24	3	*S. seenghala*	Ganga	Allahabad (3)	KT306671, KT022187, KT022190	Kumari et al., 2016
Hap.25	1	*S. seenghala*	Ganga	Allahabad (1)	KT022181	Kumari et al., 2016
Hap.26	1	*S. seenghala*	Ganga	Allahabad (1)	KT022184	Kumari et al., 2016
Hap.27	2	*S. seenghala*	Ganga	Allahabad (1); Varanasi (1)	MT707763, KT022191	This study & Kumari et al., 2016
Hap.28	3	*S. seenghala*	Ganga	Varanasi (2); Sahebganj (1)	MT707764	This study
Hap.29	1	*S. seenghala*	Ganga	Varanasi (1)	MT707765	This study

Hap.30	2	*S. seenghala*	Ganga	Varanasi (2)	MT707766	This study
Hap.31	2	*S. seenghala*	Ganga	Sahebganj (1), Munger (1)	MT707767	This study
Hap. 32	1	*S. seenghala.*	Ganga	Munger (1)	MT707768	This study
Hap.33	4	*S. seenghala*	Ganga	Munger (4)	MT707769	This study
Hap.34	6	*Sperata spp.*	Ganga	Munger (4), Sahebganj (2)	MT707770	This study
Hap.35	4	*Sperata spp.*	Ganga	Munger (2), Sahebganj (2)	MT707771	This study
Hap.36	2	*Sperata spp.*	Ganga	Munger (2)	MT707772	This study
Hap.37	2	*Sperata spp.*	Ganga	Munger (2)	MT707773	This study
Hap. 38	2	*S. aor*	Ganga	Bijnor(1), Narora(1)	MT707774	This study
Hap.39	2	*S. aor*	Ganga	Bijnor(1), Narora(1)	MT707775	This study
Hap.40	1	*S. aor*	Ganga	Bijnor(1)	MT707776	This study
Hap.41	1	*S. aor*	Ganga	Bijnor(1)	MT707777	This study
Hap.42	2	*S. aor*	Ganga	Bijnor(2)	MT707778	This study
Hap.43	1	*S. aor*	Ganga	Bijnor(1)	MT707779	This study
Hap.44	1	*S. aor*	Ganga	Bijnor(1)	MT707780	This study
Hap.45	2	*S. aor*	Ganga	Bijnor(2)	MT707781	This study
Hap.46	1	*S. aor*	Ganga	Bijnor(1)	MT707782	This study
Hap.47	2	*S. aor*	Ganga	Bijnor(2)	MT707783	This study

Hap.48	1	*S. aor*	Ganga	Bijnor(1)	MT707784	This study
Hap.49	7	*S. aor*	Ganga	Bijnor(7)	MT707785	This study
Hap. 50	8	*S. aor*	Ganga	Bijnor(5), Narora(3)	MT707786	This study
Hap. 51	15	*S. aor*	Ganga	Bijnor(10), Tigarighat(3), Narora(2)	MT707787	This study
Hap. 52	4	*S. aor*	Ganga	Bijnor(2), Tigarighat(2)	MT707788	This study
Hap. 53	4	*S. aor*	Ganga	Bijnor(2), Narora(1); Tigarighat(1)	MT707789	This study
Hap. 54	1	*S. aor*	Ganga	Bijnor(1)	MT707790	This study
Hap. 55	2	*S. aor*	Ganga	Tigarighat(2)	MT707791	This study
Hap. 56	2	*S. aor*	Ganga	Tigarighat(2)	MT707792	This study
Hap. 57	2	*S. aor*	Ganga	Narora(1); Tigarighat(1)	MT707793	This study
Hap. 58	3	*S. aor*	Ganga	Narora(1); Varanasi (2)	MT707794	This study
Hap. 59	1	*S. aor*	Ganga	Narora(1)	MT707795	This study
Hap. 60	1	*S. aor*	Ganga	Narora(1)	MT707796	This study
Hap. 61	2	*S. aor*	Ganga	Varanasi (2)	MT707797	This study
Hap. 62	5	*S. aor*	Ganga	Varanasi(5)	MT707798	This study
Hap. 63	2	*S. aor*	Ganga	Varanasi(2)	MT707799	This study
Hap. 64	1	*S. aor*	Ganga	Varanasi(1)	MT707800	This study
Hap. 65	2	*S. aor*	Ganga	Varanasi(2)	MT707801	This study

Hap. 66	7	*S. seenghala*	Mahanadi	Mahanadi (7)	KT306645, KT022202-KT022207	Kumari et al., 2016
Hap. 67	11	*S. seenghala*	Mahanadi	Mahanadi (11)	KT306640- KT306643, KT022211-KT022217	Kumari et al., 2016
Hap.68	4	*S. seenghala*	Mahanadi	Mahanadi (4)	KT306644, KT022208-KT022210	Kumari et al., 2016
Hap.69	3	*S. seenghala*	Mahanadi	Mahanadi(3)	KT022199-KT022201	Kumari et al., 2016
Hap.70	2	*S. seenghala*	Brahmaputra	Brahmaputra(2)	KT022221, KT306657	Kumari et al. 2016
Hap.71	8	*S. seenghala*	Brahmaputra	Brahmaputra(8)	KT306664, KT306663, KT306661, KT306659, KT306652, KT306651, KT306649, KT306647	Kumari et al. 2016
Hap. 72	2	*S. seenghala*	Brahmaputra	Brahmaputra(2)	KT022218, KT306654	Kumari et al. 2016
Hap. 73	3	*S. seenghala*	Brahmaputra	Brahmaputra(3)	KT306665, KT306666, KT306653	Kumari et al. 2016
Hap. 74	2	*S. seenghala*	Brahmaputra	Brahmaputra(2)	KT306658, KT306646	Kumari et al. 2016
Hap. 75	2	*S. seenghala*	Brahmaputra	Brahmaputra(2)	KT306660, KT306648	Kumari et al. 2016
Hap. 76	2	*S. seenghala*	Brahmaputra	Brahmaputra(2)	KT306662, KT306650	Kumari et al. 2016
Hap. 77	2	*S. seenghala*	Brahmaputra	Brahmaputra(2)	KT306655, KT022219	Kumari et al. 2016

Table 8: Haplotype diversity of *S. seenghala* and *S. aor* within rivers.

Parameters	*Sperata seenghala* Ganga	Mahanadi	Brahmaputra	*Sperata aor* Ganga
n	158	25	25	78
N_h	37	4	9	28
S	86	136	173	42
Ts/Tv	69/19	94/49	11/15	37/5
h (±SD)	0.943 (0.008)	0.717(0.056)	0.820 (0.065)	0.934(0.016)
π (±SD)	0.012 (0.001)	0.049(0.012)	0.034(0.017)	0.009(0.0008)
Neutrality Analysis				
Tajima's D (P)	-1.314(0.061)	0.498(0.669)	-1.149(1.02)	-0.295(0.399)
Fu's Fs (P)	-4.662(0.152)	31.053(1.00)	1.107(0.717)	-4.900(0.109)
Mismatch analysis				
SSD (P)	0.2500(0.770)	0.171(0.00)	0.492(0.60)	0.014(0.480)
r-index (P)	0.0062(0.900)	0.261(0.00)	0.058(1.00)	0.007(0.910)

Note: Sample size (n); number of haplotypes (N_h); segregating sites (S); haplotype diversity (h); nucleotide diversity (**π**); Standard diversity (SD); sum of square deviation (SSD); raggedness index (r)

3.4.2 *Molecular phylogeny and network*

The Bayesian consensus tree was generated to access the genetic relationships between all haplotypes belongs to Ganga, Mahanadi, and Brahmaputra (Fig. 10). The phylogenetic tree showed that *S. seenghala* from the Ganga formed two distinct clades: clade-I A and clade-I B, whereas Mahanadi forms a basal clade II. The *S. aor* and the Brahmaputra sequences formed distinct clade-III and IV. Two sequences of Brahmaputra (Hap 3) were merged with the haplotype of Ganga *seenghala,* and three sequences of Mahanadi (Hap 69) were clustered with haplotype of *S. aor* (Fig 10). The median-joining network of all recognized haplotypes strongly indicated the presence of multiple lineages in *S. seenghala* from the studied river (Fig. 11). The clustering pattern obtained from network analysis is highly supportive of phylogenetic analysis.

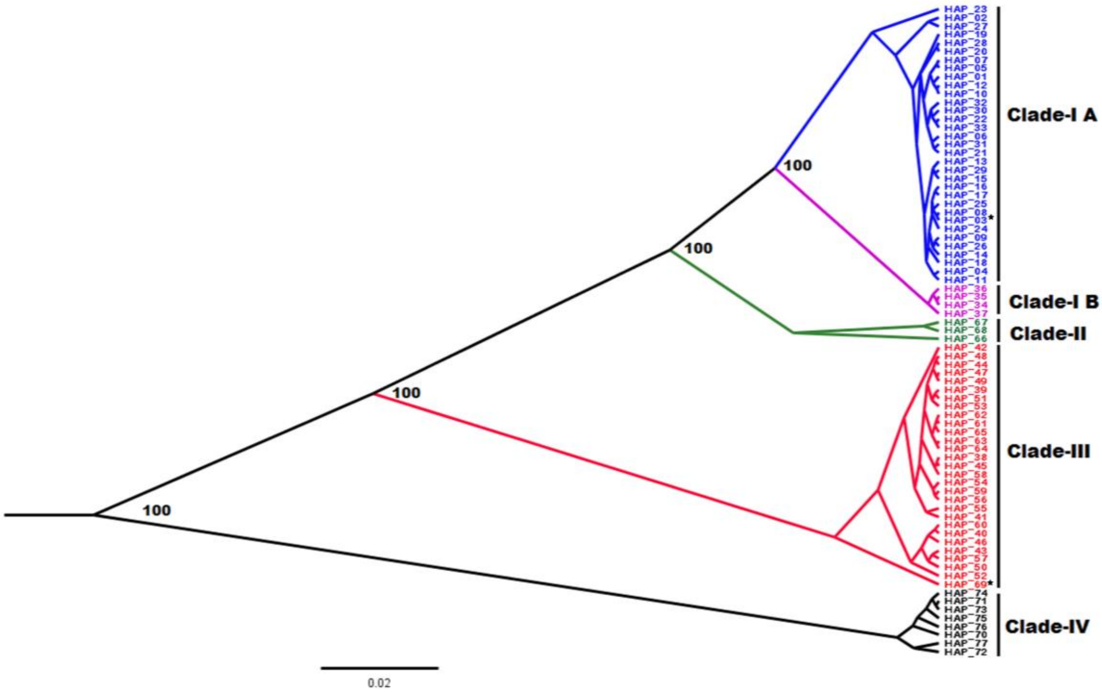


Figure 10: The Bayesian tree for the phylogenetic resolution between the *S. seenghala* form Ganges, Mahanadi, Brahmaputra, and *S. aor* Ganges. Numeric value at each node represents the posterior probability. Asterisk (*) represents, Hap3 belongs to Brahmaputra and Hap69 belongs to Mahanadi population identified by Kumari et al., (2016).

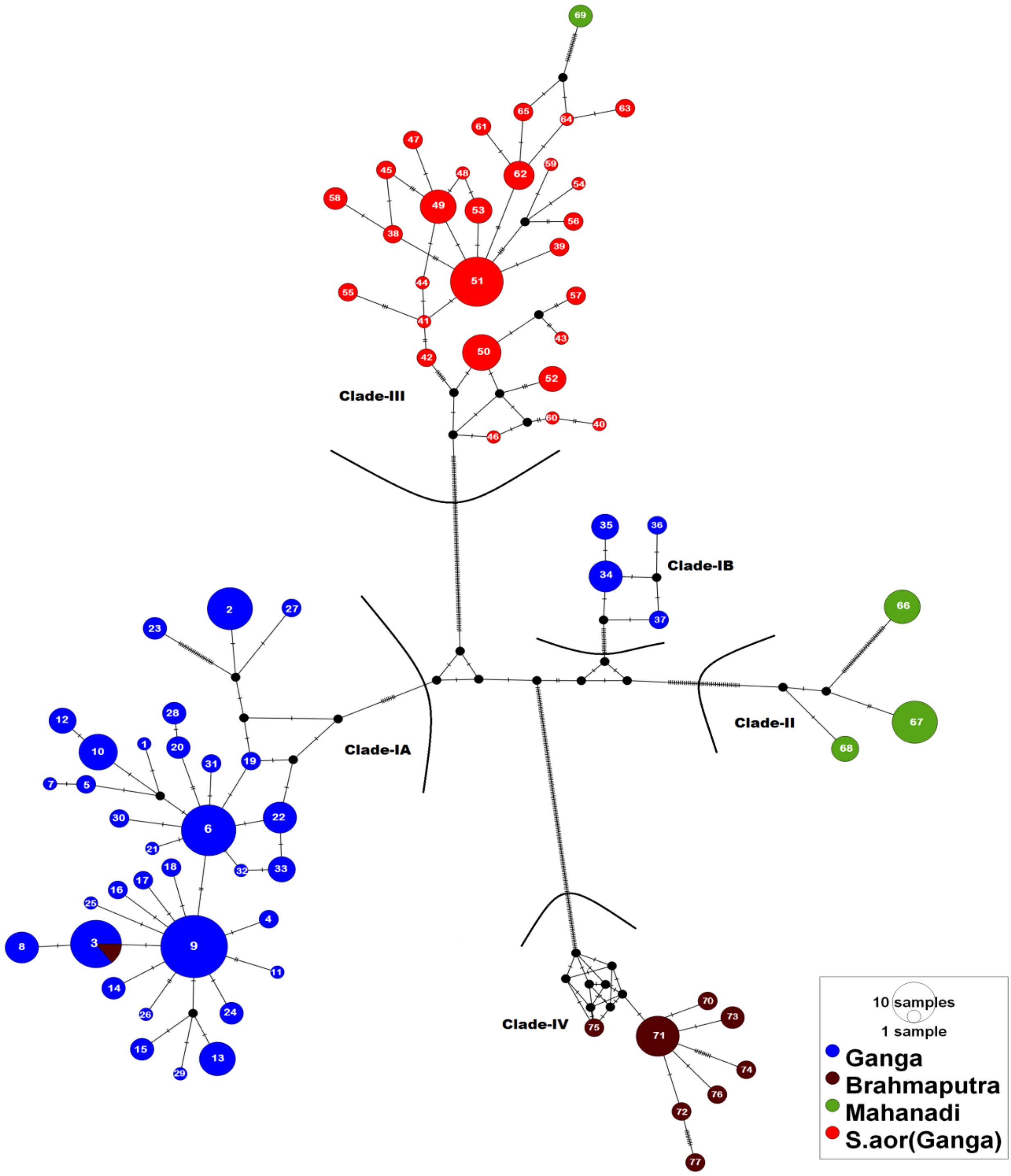


Figure 11: Median-joining network of *Sperata* species indicates the presence of three distinct clades. Clade-*I* consists of haplotypes of *S. seenghala* from the Ganges; Clade-*II* consists of Brahmaputra, Clade-*III* belongs to Mahanadi, and Clade-*IV* consists of haplotypes of *S. aor* from the Ganges.

3.4.3 Genetic differentiation

A few sequences of Brahmaputra and Mahanadi were merged with Ganga *seenghala* and *S. aor*, respectively, therefore the factual mean pairwise genetic distance was estimated based on their clustering pattern. The analyses demonstrated significant genetic differentiation between all the lineages of *Sperata*. The genetic distinction between two known species, i.e, *S. seenghala* (Clade IA) and *S. aor* (Clade III) is 0.175. Considering these estimates as a reference for species-level variation, the Brahmaputra population is genetically highly diverse (0.25 to 0.28) than the other studied lineages. The genetic differentiation between two lineages of *S. seenghala* (Clade IA and IB) from Ganga is 0.048, whereas it was comparably high with Mahanadi 0.076 to 0.080 (Table. 9). The analysis of molecular variance indicated a high degree of structuring among the populations. A large proportion of genetic variation was attributable to the difference among the group, 62.56% and the differentiation among populations within the groups, 29.46% (Table 10). The F_{SC} and Fst values were found to be significant P<0.05.

Table 9: Estimates of evolutionary divergence over sequence pairs between groups of *Spereta* from four different rivers. Standard error estimate(s) are shown above the diagonal.

Clade	IA	IB	II	III	IV
IA		0.008	0.010	0.017	0.023
IB	0.048		0.010	0.017	0.022
II	0.080	0.076		0.016	0.022
III	0.175	0.174	0.165		0.024
IV	0.253	0.253	0.257	0.287	

Table 10: Hierarchical analysis of molecular variance (AMOVA) for *S. Seenghala* from Ganges, Mahanadi, and Brahmaputra

Source of variation	d.f.	Sum of squares	Variance components	% of variation	Fixation indices	P values
Among groups	3	11195.790	45.5	62.56	FCT: 0.625	0.11±0.01
Among populations within groups	1	552.709	21.43	29.46	FSC: 0.786	0.000
Within populations	279	1619.695	5.80	7.98	FST: 0.920	0.000

3.4.4 Demographic history

The demographical dynamics of the *Sperata* was inferred from the neutrality test and mismatch distributions (Table 8 and Fig.12). No statistical significance values (P > 0.10) for Tajima's D and Fu's Fs were observed in *Sperata* populations. The bimodal shaped graphs were observed in *seenghala* of Ganga and Brahmaputra, whereas it was multimodal and ragged-shaped in *S.aor* and Mahanadi, suggested population subdivision and indications of demographic equilibrium. To assess the fitment of our data, we calculated the SSD and raggedness statistic (r-index) under the demographic expansion model for each population. However, these values were not statistically significant except Mahanadi, which indicates that neither the neutrality test nor the mismatch distribution test supported the hypothesis that the *Sperata* populations of the have passed through population expansions (Table 8).

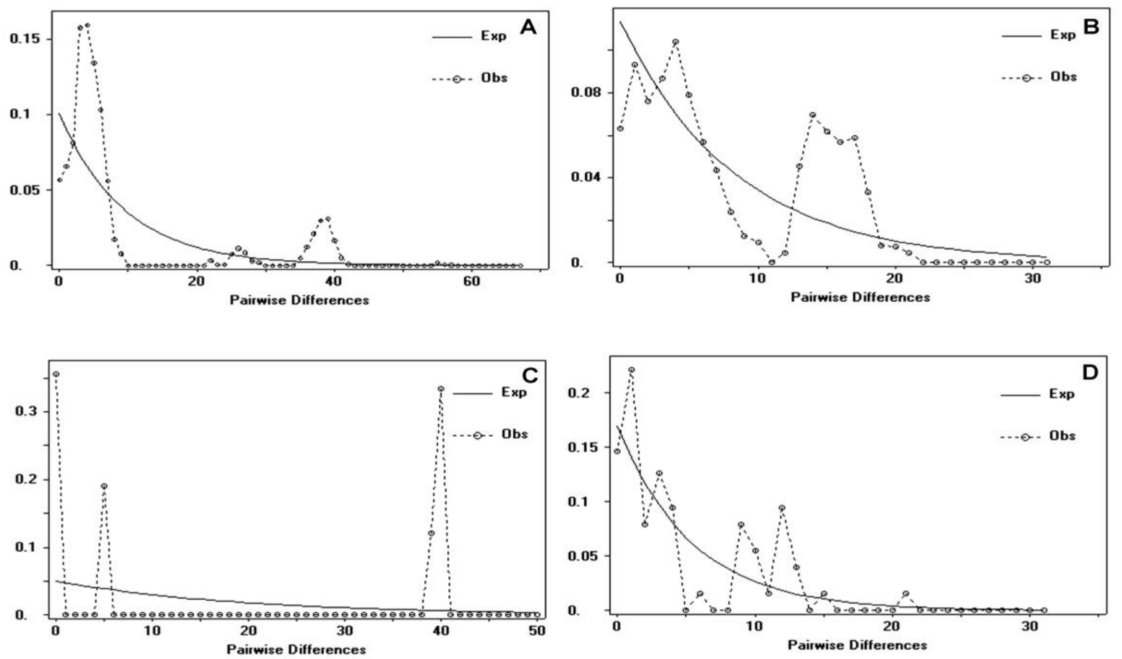


Figure 12: Mismatch distribution graph for *Sperata* populations A: Ganges *S. seenghala*; B: Ganges *S. aor*; C: Brahmaputra; and D: Mahanadi population. The x-axis shows the number of pairwise differences; the y-axis shows the frequency of pairwise comparisons. The observed and expected frequency under the population expansion model is represented by a dotted and continuous line.

3.5 Discussion

The genetic structure and diversity analysis of any species that inhabits under high threats give us vital information for their development and effective management (Kumar *et al*., 2017; Gupta *et al*., 2018). For the conservation and to maintain the pure genetic stock, knowledge about the existing lineages is essential. Through this study, we first time generated the molecular data of *S. aor* and compared it with other existing populations of *Sperata spp*. Our data revealed a high haplotype diversity (>0.71) in *Sperata* species found in three Indian rivers. The pairwise genetic distance, phylogenetic, and AMOVA analysis indicated population structuring in *Sperata* spp. The phylogenetic relationship and network revealed the presence of two mtDNA lineages (clade IA and IB) of the *S. seenghala* from the Ganga river. Clade IA widely distributed throughout the Ganga, whereas Clade IB confined to lower stretch of Ganga, i.e, Munger district, Bihar. Interestingly, when we compared the data with Kumari *et al*. (2017), the two sequences (accession no. KT022220 and KT306656) from the Brahmaputra were merging with the Ganga haplotype (Hap.3). In contrast, three sequences (accession no. KT022199, KT022220, and KT022201) from Mahanadi (Hap. 69) were clustered with *S. aor* of the Ganga river. It indicated that the Brahmaputra River also sustains the gene pool of the Ganga river. The possible explanation for the existence of the Brahmaputra lineage in the Ganga is that both the rivers merge at the Sunderbans area of Bangladesh and provide the potential route for migration. Besides, the presence of similar gene-pool of Ganga *S. seenghala* in Brahmaputra river could be the effect of human-mediated propagation, and there is also a possibility of cross-contamination or cataloging error while doing genetic analysis. Further, three samples collected in a previous study from the Mahanadi river as *S. seenghala* belongs to *S. aor*. The clustering of Mahanadi *seenghala* with *S. aor* might be the result of the misidentification of species during sample collection. Therefore, the genetic differentiation was estimated between the different identified lineages observed in the study. Considering two known species *S. seenghala* and *S. aor* as a reference, we observed high genetic differentiation value (17.5%) than the newly identified lineage of *seenghala* from Ganga (4.8%) and Mahanadi (8.0%). Interestingly, the genetic differentiation between Ganga-Brahmaputra was much higher (25.3%) than the values observed from *S. seenghala* and *S. aor*. This result indicated the presence of highly diverse lineages in the Brahmaputra than the *S. seenghala*, that needs to be confirmed based on their detailed morphological characteristics. The existence of structuring in *Sperata* is supported with a previous RAPD based study where population sub-structuring was observed in giant river catfish of Sutlej and Beas from the Indus river system (Saini, Dua & Mohindra, 2010). The study on stock discrimination using microsatellite markers from the middle to lower stretch of river Ganga, i.e, Narora–Kanpur, Varanasi, and Bhagalpur, revealed three different stocks of *S. aor*

(Nazir and Khan, 2017). However, this result was contradictory with our mtDNA based study where we did not found any significant barrier to gene flow and structuring in *S.aor* across the sampling sites of the Ganga river.

Moreover, the existence of two distinct lineages of *S. seenghala*, which is diverse than *S.aor* and Mahanadi lineage, indicated the presence of multiple clades of *Sperata*. In the stretch of Varanasi to Sahebganj of the Ganga river, a large number of tributaries such as Gomti, Ghaghara, Gandak, Son, Kiul and Kosi merge with the Ganga that might have resulted in the migration and confinement of the different gene-pool of these rivers to a particular habitats (Nazir and Khan, 2017). Recently, Acharya *et al* (2019) reported population sub-structuring in the Brahmaputra, Ganga, Godavari, Mahanadi, and Narmada. The results indicated the low genetic divergence between Ganga–Brahmaputra than the Mahanadi populations, with comparatively low genetic diversity from the Ganga river due to the presence of excess homozygotes. In contrast to this, our mtDNA results showed Ganga *seenghala* is genetically close to Mahanadi than the Brahmaputra, which is supported by a previous study based on cytochrome *b* and D-loop region (Kumari *et al*. 2017). Though, incongruence results obtained from mtDNA and microsatellite analysis, suggesting that sex-biased dispersal behaviors that could contributed to disparities in genetic structuring of species. It also appears that the Ganga and Mahanadi are geographically more isolated than Brahmaputra and all three rivers obviously on an independent evolutionary trajectory; however, our study indicated that Brahmaputra lineages qualify the species level variation and adequately address the highly diverse lineages among S*pereta*. The non-significant Tajima's D and Fu's Fs tests suggested that there was no historical reduction in effective population size in these rivers. The mismatch distribution analysis showed bi-modal and multimodal mismatch distribution indicates the demographic stability or results of population admixture in S*pereta* that could be due to their reliability with allopatric divergence (Rogers & Harpending, 1992; Zhao et al. 2008).

The present genetic features in Ganga as well as other existing *Spereta* populations are the consequence of long-term geographical isolation and adaptation to the local environment.

3.6 Conclusions

The present study provides the reference database for the identification and genetic differentiation of *S. aor* and *S. seenghala* from the Ganga river. The study highlighted the presence of two distinct lineages of *S. seenghala* from the Ganga river. Also, it indicated that the Brahmaputra population is much diverse than *S. seenghala* and *S. aor*. Therefore, the present study highlighted that the Brahmaputra giant river catfish is a highly diverse lineage that genetically qualifies the status of distinct species. However, it requires a detailed morphological study from more locations for the identification of multiple lineages existing in the Indian river that can formulate the appropriate conservation and management plans. This study indicated the presence of multiple lineages of *Spereta* from Ganga, Mahanadi, and Brahmaputra that considered as a distinct Evolutionary Significant Unit (ESUs). These findings will help understand the fish biology and implement the proper conservation, management plan, for development catfishes. However, detailed comparative morphological evidence with support of molecular data is required to be able to support their consideration as separate species.

3.7 References

- Acharya, A. P., Pavan-Kumar, A., Joshi, C. G., Namrata, P., Gireesh-Babu, P., Chaudhari, A., & Krishna, G. (2018). Development and characterization of 15 novel polymorphic microsatellites for Giant river-catfish *Sperata seenghala* (Sykes, 1839) using next generation sequencing approach. *J Appl Ichthyol*, 34, 971-973. https://doi.org/10.1111/jai.13683
- Acharya, A.P., Pavan-Kumar, A., Gireesh-Babu, P., Joshi, C.G., Chaudhari, A. & Krishna, G. (2019). Population genetics of Indian giant river-catfish, *Sperata seenghala* (Sykes, 1839) using microsatellite markers. *Aquatic Living Resources*, 32, 4. https://doi.org/10.1051/alr/2019002
- Day, F. (1878). The fishes of India being a natural history of the fishes known to inhabit the seas and fresh waters of India, Burma and Ceylon. William Dowson and Sons, London 778.
- Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution*, 29(8), 1969-1973. https://doi.org/10.1093/molbev/mss075
- Excoffier, L. & Lischer, H.E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, 10(3), 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x
- Excoffier, L., Estoup, A. & Cornuet, J.M. (2005). Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics*, 169(3), 1727-1738. https://doi.org/10.1534/genetics.104.036236
- Ferraris, J.C.J. & Runge, K.E. (1999). Revision of the South Asian Bagrid catfish genus *Sperata*, with the description of a new species from Myanmar. *Proceedings of the California Academy of Sciences*, 51(10), 397–424.
- Froese, R. & Pauly, D. (2011). Species of *Sperata* in FishBase. Retrieved from http://www.fishbase.org.
- Fu, Y.X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147(2), 915-925.
- Garg, R.K., Sairkar, P., Chouhan, S., Batav, N., Silawat, N., Sharma, R. *et al*. (2014). Characterization of genetic variance within and among five populations of *Sperata seenghala* (Skyes, 1839) revealed by random amplified polymorphic DNA markers. *Journal of Genetic Engineering and Biotechnology*, 12(1), 7-14. http://doi.org/10.1016/j.jgeb.2013.11.001
- Gupta, S. (2015). Review on *Sperata seenghala* (Sykes, 1839), a freshwater catfish of Indian subcontinent. *J Aquac Res Development*, 6(290), 2. http://doi.org/10.4172/2155-9546.1000290
- Gupta, S.K., Kumar, A., Angom, S., Singh, B., Ghazi, M.G.U., Tuboi, C. & Hussain, S.A. (2018). Genetic analysis of endangered hog deer (*Axis porcinus*) reveals two distinct lineages from the Indian subcontinent. *Scientific reports*, 8(1), 1-12. https://doi.org/10.1038/s41598-018-34482-9
- Khan, M. A., & Nazir, A. (2019). Stock delineation of the long-whiskered catfish, Sperata aor (Hamilton 1822), from River Ganga by using morphometrics. *Marine and Freshwater Research*, 70(1), 107-113. https://doi.org/10.1071/MF17306
- Kumar, A., Ghazi, M.G.U., Hussain, S.A., Bhatt, D. & Gupta, S.K. (2017). Mitochondrial and nuclear DNA based genetic assessment indicated distinct variation and low genetic exchange among the three subspecies of swamp deer (*Rucervus duvaucelii*). *Evolutionary Biology*, 44(1), 31-42. https://doi.org/10.1007/s11692-016-9387-2
- Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*, 35(6), 1547-1549. https://doi.org/10.1093/molbev/msy096

- Kumari, P., Pavan-Kumar, A., Kumar, G., Alam, A., Parhi, J., Gireesh-Babu, P. *et al*. (2017). Genetic diversity and demographic history of the giant river catfish *Sperata seenghala* inferred from mitochondrial DNA markers. *Mitochondrial DNA Part A*, 28(6), 920-926. http://doi.org/10.1080/24701394.2016.1209195
- Leigh, J.W. & Bryant, D. (2015). Popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9), 1110-1116. https://doi.org/10.1111/2041-210X.12410
- Librado, P. & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25(11), 1451-1452. https://doi.org/10.1093/bioinformatics/btp187
- Miyan, K., Khan, M.A., Patel, D.K., Khan, S. & Prasad, S. (2016). Otolith fingerprints reveal stock discrimination of *Sperata seenghala* inhabiting the Gangetic river system. *Ichthyological research*, 63(2), 294-301. http://doi.org/10.1007/s10228-015-0500-9
- Mohanty, B. P., Paria, P., Das, D., Ganguly, S., Mitra, P., Verma, A. *et al*. (2012). Nutrient profile of giant river-catfish *Sperata seenghala* (Sykes). *National Academy Science Letters*, 35(3), 155-161. http://doi.org/10.1007/s40009-012-0014-1
- Nawaz, H., Sharif, M.H. & Mirza, M.R. (1994). Food of the singhari, *Aorichthys uor sarwari* (Pisces: Bagridae). BIOLOGIA-LAHORE-, 40, 103-108.
- Nazir, A. & Khan, M.A. (2017). Stock discrimination of *Sperata aor* from river Ganga using microsatellite markers: implications for conservation and management. *Aquatic Living Resources*, 30, 33. https://doi.org/10.1051/alr/2017033
- Ng, H.H. & Kottelat, M. (2013). Revision of the Asian catfish genus Hemibagrus Bleeker, 1862 (Teleostei: Siluriformes: Bagridae). *Raffles Bulletin of Zoology*, 61(1).
- Rogers, A.R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular biology and evolution*, 9(3), 552-569. https://doi.org/10.1093/oxfordjournals.molbev.a040727
- Rogers, A.R. (1995). Genetic evidence for a Pleistocene population explosion. *Evolution*, 49(4), 608-615. https://doi.org/10.1111/j.1558-5646.1995.tb02297.x
- Saigal, B.N. & Motwani, M.P. (1961). Studies on the fishery and biology of the commercial catfishes of the ganga river system I. Eirly Life-History, Bionomics and Breeding of *Mystus* (Osteobagrus) *seenghala* (Syke). *Indian Journal of Fisheries*, 8(1), 60-74.
- Saini, A., Dua, A. & Mohindra, V. (2010). Genetic variability analysis of Giant river catfish (*Sperata seenghala*) populations from Indus river system by RAPD-PCR. *Genetika*, 46(8), 1102-1107. http://doi.org/10.1134/S1022795410080107
- Sambrook, J.E., Fritsch, E.F., Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual 2nd edn *Cold Spring Harbor Laboratory Press NY*.
- Sarkar, U.K., Gupta, B.K. & Lakra, W.S. (2010). Biodiversity, ecohydrology, threat status and conservation priority of the freshwater fishes of river Gomti, a tributary of river Ganga (India). *The Environmentalist*, 30(1), 3-17. http://doi.org/10.1007/s10669-009-9237-1
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), 585-595.
- Talwar, P.K. & Jhingharan, A.G. (1991). Inland fishes of India and adjacent countries, vol II. Oxford and IBH Publishing Co, New Delhi 19, 1158.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids research*, 25(24), 4876-4882. https://doi.org/10.1093/nar/25.24.4876

- Wan, Q.H., Wu, H., Fujihara, T. and Fang, S.G. (2004). Which genetic marker for which conservation genetics issue?. *Electrophoresis*, 25(14), pp.2165-2176.
- Yadav, P., Tripathi, R.K. & Mohindra, V. (2018). CmZNF384-like gene in hypoxia-tolerant Indian catfish, *Clarias magur* (Hamilton 1822) in hypometabolic states associated with acute hypoxia and summer aestivation. *Agri Gene*, 10, 1-11. https://doi.org/10.1016/j.aggene.2018.11.001
- Yadav, P., Tripathi, R.K., Singh, R.K. & Mohindra, V. (2017). Expression profile and in silico characterization of novel RTF2h gene under oxidative stress in Indian catfish, *Clarias magur* (Hamilton 1822). *Molecular biology reports*, 44(1), 63-77. http://doi.org/10.1007/s11033-016-4081-0
- Zhao, L., Zhang, J., Liu, Z., Funk, S.M., Wei, F., Xu, M., Li, M. (2008). Complex population genetic and demographic history of the Salangid, Neosalanx taihuensis, based on Cytochrome b sequences. *BMC Evolutionary Biology*, 8(1), 201.

4. POPULATION GENETIC STRUCTURE, DEMOGRAPHY CHANGES AND EVOLUTIONARY DIVERGENCE AMONG ASIAN FLAP SHELL TURTLE (*LISSEMYS SPECIES*)

4.1 Background of the study

The softshell turtles are an ancient, morphologically highly diverse and have globally affected by a variety of direct and indirect threats (Ernst and Barbour 1989; Choudhury et al. 2000; Li et al. 2017) . The presence of shiny leathery and pliable skin, flattened body shape with fleshy lips, and highly modified skulls are a unique feature that distinguished from hard-shell turtles (Meylan 1987; Ernst and Barbour 1989). The extant softshell turtles belong to Trionychidae that is widely distributed in North America, Europe, Africa and Asia (Engstrom and Okamura 2004). Despite the large distribution range, softshell turtles are extremely vulnerable to rapid decline, therefore knowledge of genetic variation and gene flow is of critical importance for the long term persistence of populations (Willi and Hoffmann 2009).

Asian flapshell turtle (AFT) belongs to family Trionychidae; genus *Lissemys* is confined to South Asia and western Southeast Asia (Webb 1982; Hossain et al. 2010; Uetz et al. 2016). Currently, three subspecies of *L. punctata* are recognized from India: *L. p. andersoni* (spotted northern Indian flapshell turtle) (Fig. 19) are confined to Pakistan, north India, Nepal, Bangladesh, western Myanmar*; L. p. vittata* (unspotted, central Indian flapshell turtle) (Fig. 20) found in central to peninsular India and *L. p. punctata* (unspotted, southern Indian flapshell turtle) (Fig. 21) restricted to southern peninsular India (Talukdar 1979; Das and Gupta 2011). *L. scutata* distributed from the Ayeyarwady (Irrawady), Sittaung, and Thanlwin (Salween) River systems of Myanmar and also reported from westernmost Thailand and Yunnan, China (Ernst and Barbour, 1989; Ernst et al. 2000; Fritz and Havaš 2007; Rhodin et al. 2010). The generic status of *L. scutata* is still under debate (Annandale 1912; Webb and RG 1980; Webb 1982; Rhodin et al. 2010), whereas Sri Lankan flapshell turtle (*L. ceylonensis*) is considered as separate species (Deraniyagala 1953; Praschag et al. 2011). The species found in a wide variety of aquatic habitats, ranging from rivers and streams to basins, lakes, swamplands, ponds, lakes, and salt marshes (Boulenger 1889; Siebenrock 1909; Smith 1931; Deraniyagala 1939; Mertens and Wermuth 1955; Wermuth and Mertens 1961; Wermuth and Mertens 1977; Webb and RG 1980; Webb 1982; Ernst and Barbour 1989; Ernst et al. 2000; Fritz and Havaš 2007; Rhodin et al. 2010). *Lissemys* are heavily poached for their meat, eggs and traditional medicine trade that appears major threats (Bhupathy et al. 2014). The latest analysis by TRAFFIC's during September 2009-September 2019, has found at least 1,11,312 tortoises and freshwater turtles reported in illegal wildlife trade in India. Out of 48.5%, documented cases, 15% belonged to AFT. Despite wide exploitation and destruction by anthropogenic pressures since a long time, *L. punctata* is still fairly communal and relatively stable populations and not yet seriously threatened (Hossain et al. 2010). Hence, the species is listed under IUCN Red List as “Least Concern” and protected in the Schedule *I* of the Indian Wildlife (Protection)

Act, 1972, included in Appendix *II* of the Convention on International Trade in Endangered Species of Flora and Fauna (CITES). Previous morphological and molecular results suggested that *Lissemys* are closely related to African flapshell turtle (Meylan 1987; Engstrom and Okamura 2004). Praschag et al. (2011) investigated the phylogenetic relationships among AFT and indicated five divergent clades with conflicting taxonomic arrangements of spotted and unspotted *Lissemys* species. Virtually all previous studies dealing with the systematics and intraspecific variations (Praschag et al. 2011; Lalitha and Chandavar 2018; Bhaskar and Mohindra 2019). The detailed genetic diversity within the subspecies of *Lissemys* genus was never questioned. The riverine ecosystem of India, particularly in one of the largest river Ganges supports rich fauna and flora, including a major population of northern subspecies of *Lissemys.* The Ganga river has suffered from intense human intervention; the continuous monitoring of such a population of freshwater turtles should guide the critical issues related to the conservation.

Low genetic variability and gene flow lead to the genetic bottlenecks and failure of species to endure changes in the environment (Keller and Waller 2002). Using various molecular markers, a wild stock can be successfully identified and characterized for better stock management, species-recovery, and reintroduction program (Gupta et al. 2018). Genetic data evaluating the variability within and among populations would significantly contribute to developing more powerful conservation strategies. In this context, we undertook the study to assess genetic diversity, gene flow and the population genetic structure of *Lissemys* from Ganga river using mitochondrial DNA (mtDNA) and nuclear gene (nuDNA). We also examine the genetic relationship, demographic pattern and lineage diversification of AFT populations. This study seeks to develop meaningful recommendations for conservation policies and the preservation of genetic stock of AFT.

Figure 13: ***Lissemys punctata andersoni***

Figure 14: ***Lissemys punctata vittata***

Figure 15: ***Lissemys punctata punctata***

4.2 Methodology

4.2.1 *Sampling*

A total of 69 samples were collected during the periods of 2017 to 2019, covering seven different sites of Ganga river. These sites were Rishikesh (n=5) from Uttarakhand; Bijinor (n=4), Kannauj (n=1), Kanpur (9), Fathepur (n=3), Prayagraj (n=8) and Sarnath turtle sanctuary, Varanasi (39) from Uttar Pradesh (Fig. 22). Most of the samples were collected by swabbing of mouth and clipping off a small piece of the toe and further released in natural habitat. Opportunistically, samples were also collected from carcasses from above-mentioned sites. The samples were preserved in 95% ethanol and stored at -20°C until processing.

4.2.2 *Laboratory procedures*

Total genomic DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). The complete cytochrome b (Cyt b) gene was amplified using the primers: CytbG:5'-AACCATCGTTGTWATCAACTAC-3' (Spinks and Shaffer 2004) and mt-f-na3:5'-AGGGTGGAGTCTTCAGTTTTTGGTTTACAAGACCAATG-3' (Lenk et al. 1999). For nuclear gene amplification, we used partial coding region for the oocyte maturation factor Mos (Cmos gene): Cmos1:5'-GCCTGGTGCTCCATCGACTGGGATCA-3' and Cmos3: 5'-GTAGATGTCTGCTTTGGGGGTGA-3' (Le et al. 2006) and Intron 1 of the RNA fingerprint protein 35 (R35) gene: R35Ex1:5'-ACGATTCTCGCTGATTCTTGC-3' and R35Ex2:5'-GCAGAAAACTGAATGTCTCAAAGG-3' (Fujita et al. 2004). PCR reactions were performed in total reaction volumes of 20 µl using a PCR buffer (10 mM Tri–HCl, pH 8.3, and 50 mM KCl), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 2 pmol of each primer, 5.0 U of Taq DNA polymerase and 1 µl (~30 ng) of the template DNA. All reactions were run along with negative controls. The PCR conditions were 95°C for 5 min followed by 35 cycles at 95°C for 45 s, annealing 54-56°C for 45 s and extension 72°C for 1 min, with a final extension of 72°C for 10 min. The effectiveness and consistency of the PCR reactions were monitored using positive controls. The amplified PCR amplicons were visualized in UV light on 2% agarose gel stained with ethidium bromide. Exonuclease I (EXO-*I*) and shrimp alkaline phosphatase (SAP) treatments were given to the amplified PCR products (USB, Cleveland, OH) for 15 minutes each at 37°C and 80°C, respectively, to eliminate any residual primer. The amplified PCR products were directly sequenced using the BigDye® Terminator Kit (v3.1) and analyzed on an ABI 3500XL Applied Biosystems Genetic Analyzer. All the products were sequenced in both directions.

4.3 Data analysis

4.3.1 Genetic diversity, phylogenetic and demographic analysis

Sequences of Cyt *b*, *Cmos* and *R35* were derived from the forward and reverse directions. In addition, 55 sequences of mtDNA Cyt *b* from different localities of known origin were downloaded from GenBank database National Center for Biotechnology Information, USA: NCBI submitted by Praschag et al. (2011) and Bhaskar and Mohindra (2019) (Fig. 22; Table 16). The mtDNA sequences of *L. p. andersoni* were divided into two groups, group I consisted of 69 samples of Ganga river and Group II consisted of sequences belongs to Uttar Pradesh near Nepal border, Assam, Haryana (Yamuna River), Bangladesh and Myanmar.

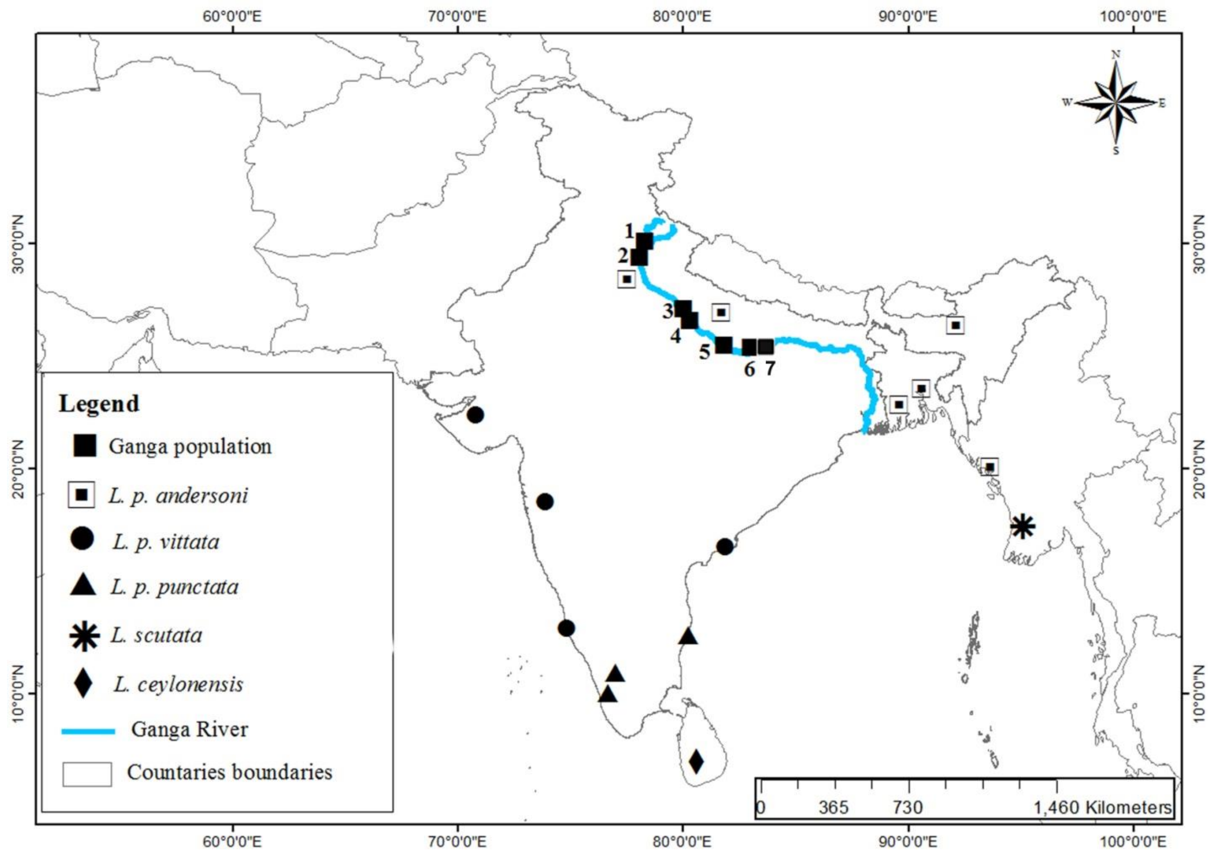


Figure 16:Map showing the sample collection sites of *Lissemys* and sequences used from type localities.

Table 11: Distribution of haplotypes in *Lissemys* populations

Hap	N	S.ID/Locations	Species	Clade	Accession No	Reference
Hap 1	34	Site4 (4); Site7(26); Site6 (3)	*L.P.andersoni*	I	MT465665	This study
		Uttar Pradesh(2)	*L.P.andersoni*		FR850617-FR850618	Praschag et al. 2011
Hap 2	7	Site7 (5); Site6 (2)	*L.P.andersoni*	I	MT465666	This study
Hap 3	1	Site4 (1)	*L.P.andersoni*	I	MT465667	This study
Hap 4	1	Site6 (1)	*L.P.andersoni*	I	MT465668	This study
Hap 5	1	Site2 (1)	*L.P.andersoni*	I	MT465669	This study
Hap 6	23	Site4 (3); Site7 (7); Site6 (2); Site5(2)	*L.P.andersoni*	I	MT465670	This study
		Bangladesh: Dhaka (5)	*L.P.andersoni*		FR850609-FR850613	Praschag et al. 2011
		Assam: Brahmaputra	*L.P.andersoni*		FR850614	Praschag et al. 2011
		Haryana: Yamuna River	*L.P.andersoni*		FR850615	Praschag et al. 2011
		Myanmar (2)	*L.P.andersoni*		FR850619-FR850620	Praschag et al. 2011
Hap 7	2	Site2(2)	*L.P.andersoni*	I	MT465671	This study
Hap 8	4	Site1(2),Site2 (2)	*L.P.andersoni*	I	MT465672	This study
Hap 9	2	Site1 (2)	*L.P.andersoni*	I	MT465673	This study
Hap 10	1	Site1 (1)	*L.P.andersoni*	I	MT465674	This study
Hap 11	2	Site4(1),Site7 (1)	*L.P.andersoni*	I	MT465675	This study
Hap 12	2	Site3(1); Site5(1)	*L.P.andersoni*	I	MT465676	This study
Hap 13	1	Godavari River:Andhra Pradesh	*L.p.vittata*	II	FR850621	Praschag et al. 2011#
Hap 14	7	Godavari River,AP (3); Rajkot, Gujrat (4)	*L.p.vittata*	II	FR850622-FR850624 FR850627-FR850630	Praschag et al. 2011#
Hap 15	1	Goa	*L.p.vittata*	II	FR850625	Praschag et al. 2011#
Hap 16	1	Gujarat: Rajkot	*L.p.vittata*	II	FR850626	Praschag et al. 2011#
Hap 17	1	Karnataka: Mangalore	*L.p.vittata*	II	FR850631	Praschag et al. 2011#
Hap 18	2	Maharashtra: Pune	*L.p.vittata*	II	FR850635,FR850636	Praschag et al. 2011#
Hap 19	3	Karnataka: Mysore	*L. punctata*	II	KY946735, KY946742; MK466375	Lalitha & Chandavar, 2018; Bhaskar & Mohindra, 2019
Hap 20	2	Karnataka: Mysore	*L. punctata*	II	KY946736;KY946740	Lalitha & Chandavar, 2018
Hap 21	1	Karnataka: Mysore	*L. punctata*	II	KY946737	Lalitha & Chandavar, 2018
Hap 22	1	Karnataka: Mysore	*L. punctata*	II	KY946738	Lalitha & Chandavar, 2018
Hap 23	1	Karnataka: Mysore	*L. punctata*	II	KY946739	Lalitha & Chandavar, 2018
Hap 24	1	Karnataka: Mysore	*L. punctata*	II	KY946741	Lalitha & Chandavar, 2018
Hap 25	1	Karnataka	*L. punctata*	II	MK466374	Bhaskar & Mohindra, 2019
Hap 26	1	Karnataka	*L. punctata*	II	MK466376	Bhaskar & Mohindra, 2019
Hap 27	4	Kerala (3); Tamil Nadu(1)	*L.p.punctata*	III	FR850632-FR850634 MK028164	Praschag et al. 2011* Bhaskar & Mohindra, 2019
Hap 28	1	Tamil Nadu	*L.p.punctata*	III	FR850642	Praschag et al. 2011*
Hap 29	3	Tamil Nadu	*L.p.punctata*	III	FR850643, FR850644, MK028163	Praschag et al. 2011* Bhaskar & Mohindra, 2019
Hap 30	1	Tamil Nadu	*L.p.punctata*	III	MK028165	Bhaskar & Mohindra, 2019
Hap 31	2	Sri Lanka	*L.ceylonensis*	IV	FR850645;FR850648	Praschag et al. 2011
Hap 32	1	Sri Lanka	*L.ceylonensis*	IV	FR850646	Praschag et al. 2011
Hap 33	1	Sri Lanka	*L.ceylonensis*	IV	FR850647	Praschag et al. 2011

Hap 34	1	Sri Lanka	*L.ceylonensis*	IV	FR850649	Praschag et al. 2011
Hap 35	1	Myanmar	*L.scutata*	V	FR850650	Praschag et al. 2011
Hap 36	2	Myanmar	*L.scutata*	V	FR850651;FR850653	Praschag et al. 2011
Hap 37	2	Myanmar	*L.scutata*	V	FR850652;JQ361816	Praschag et al. 2011 Nie and Liu,(unpublished)
Hap 38	1	Myanmar	*L.scutata*	V	AY259567	Engstrom et al., 2004

Note: Wrongly submitted: #Sequences mentioned as *L.p.punctata*; *mentioned as *L.p.vittata* in Genebank.

The *Cmos* and *R35* sequence of *Lissemys* (except one from *L. punctata* and two from *L. scutata*) was not available in NCBI database. All the sequences were aligned using CLUSTAL W (Thompson et al. 1994), as implemented in the BioEdit v 7.2.5 software (Hall 1999) and manually checked. Newly generated sequences were deposited in GenBank under accession numbers (MT465659- MT465681).

The number of polymorphic sites (S), number of haplotypes (h), haplotype diversity (hd) and nucleotide diversity (π) within the Ganga population and other known global populations of *Lissemys* were computed using the software DNASPv5.0 (Librado and Rozas 2009). Maximum-likelihood (ML) tree was performed with RAxML implemented in raxmlGUI 2.0 beta using GTR with a gamma distribution (+G) and a proportion of invariable sites (+I) model of best-fit nucleotide substitution under the Bayesian Information Criterion (BIC). The bootstrap support was obtained by running 1000 pseudoreplicates (Edler et al. 2019). One sequence of *C. senegalensis* (FR850654) was used as an outgroup for better insights into the phylogenetic relationships.

The resulting phylogenetic trees were visualized in FigTreev1.4.0 (Rambaut 2012) (http://tree.bio.ed.ac.uk/software/figtree/). The spatial distribution of the haplotypes was visualized through a median-joining network, which was created using the PopART software (Leigh and Bryant 2015).

To determine the demography history of *Lissemys* population, we performed different statistical approaches such as Tajima's *D* (Tajima 1989), Fu's F*s* test (Fu 1997). Mismatch distribution analysis, a sum of squared deviations (SSD), and the raggedness index (r) were also used to demonstrate the pattern of population stability or population expansion in the past under the sudden demographic expansion models using the ARLEQUIN v3.5 program (Excoffier and Lischer 2010). The *P*-values were obtained from 1000 simulations on the basis of a selective neutrality test.

A Bayesian skyline plot (BSP) was constructed using the Monte Carlo Markov Chain (MCMC) method with 100 million generations using BEAST ver 1.7.5 (Drummond et al. 2012). The temporal trends in the effective population size of the *Lissemys* overtime/generations were estimated using a coalescent BSP. Nucleotide substitution rates were assumed to be 0.125%/lineage/myr according to the previous study (Avise et al. 1992).

4.3.2 Estimating divergence dating

To estimate the divergence time of the *Lissemys,* we inferred genealogies using a relaxed-clock method in BEAST v.1.7.5 (Drummond et al. 2012). We performed the analysis using 13 softshell turtle sequences belongs to 12 genera downloaded from the NCBI database: *Trionyx triunguis* (AB477345), *Chitra indica* (JQ406951), *Chitra chitra* (AY259562), *Amyda cartilaginea* (AY259550), *Nilssonia formosa* (AY259547), *Apalone spinifera* (NC021371), *Pelochelys cantorii* (JN016747), *Rafetus swinhoei* (NC017901), *Apalone ferox* (NC014054), *Palea steindachneri* (NC013841), *Dogania subplana* (NC002780), *Pelodiscus sinensis* (NC006132) and *Carettochelys insculpta* (NC014048).

Divergence time estimation was calibrated between Trionychidae and Carettochelyidae, so the node age was estimated to be 146 MA (million years ago), with a 95% confidence interval running from 126 to 173 MA (Li et al. 2017). Tree prior categories were set to the Yule-type speciation. A HKY model using gamma + invariant sites with five gamma categories was used. We conducted two independent analyses to confirm the convergence of the analysis and Markov chain Monte Carlo (MCMC) was run for 150 million generations and sampled every 1000 generations. All the runs were evaluated in Tracer v. 1.6. The final phylogenetic tree was visualized in FigTree v.1.4.2. (http://tree.bio.ed.ac.uk/software/figtree/) (Rambaut 2012).

4.3.3 *Genetic differentiation*

The conventional *F*ST i.e. Wright's fixation index of population subdivision, was calculated to test for differences in haplotype frequencies (Weir and Cockerham 1984) and ϕ statistics was calculated which incorporate the information of nucleotide differences between haplotypes (Excoffier et al.1992). The significant P values of the statistics were computed using a nonparametric permutation approach with 1000 permutations.

4.4 Results

4.4.1 *Sequence information and genetic diversity*

We generated the mtDNA cyt *b* dataset comprised 1133bp, two nuDNA datasets *Cmos* of 602bp, and *R35* of 985bp from 69 samples of *L. p. andersoni* from seven sites of Ganga river. In addition, we included 55 cyt *b* sequences from GenBank, which were consisted of *L. p. andersoni* (11), *L. p. punctata* (9), *L. p. vittata* (24), *L. ceylonensis* (5), and *L. scutata* (6) (Table 16). Among the 124 cyt *b* sequences, we observed 229 variable sites, of these 29 polymorphic sites were singletons, and 200 were parsimony informative. These sequences were grouped into 38 haplotypes (H), and there were different numbers of individuals in each haplotype (Table 16). Of these, 12 haplotypes (H1 to H12) were observed in Ganga and two haplotypes (H1 to H6) were observed in other populations of *L. p. andersoni*. The H1 and H6 is the core haplotypes, where H1 is shared with the sequences of Ganga and other northern group and H6 is more common in the north to southeast populations. Fourteen haplotypes (H13 to H26) were detected in *L. p. vittata,* four (H27 to H30) were observed in *L.p. punctata*, four (H31 to H34) were found in Sri Lankan population, *L. ceylonensis*, and four (Hap 35 to 38) were detected from Myanmar lineages*, L. scutata* (Table 16 and Table 17). The haplotypes (*hd*) and nucleotide diversity (π) of *L. punctata* are varied and the values are ranged from *hd*=0.327 ± 0.053 and π=0.0002 ± 0.0001 in *L. p. andersoni* group II to *hd*=0.906 ± 0.046 and π=0.013 ± 0.001 in *L. p. vittata*. The *hd* and π of Ganga population was 0.694 ± 0.054 and 0.0009 ± 0.0001, respectively. The gene diversity of overall *L. p. andersoni* population was high and estimates as *hd*=0.732 ± 0.036 and π=0.00095 ±0.00009. The genetic diversity indices of all AFT presented in Table 17.

In two nuDNA, observed potentially informative sites were 5 (*CMOS*) and 3 (*R35*). The number of haplotypes were 6 (*CMOS*) and 5 (*R35*). Overall, *Hd* and π, were as follows: *CMOS* (*Hd*= 0.585 ± 0.063, π=0.0014 ± 0.0001); and it was higher in *R35* (*Hd*= 0.721 ± 0.045, π=0.0009 ± 0.00009).

The pairwise differences (PD) derived from the mtDNA significantly differentiated ($P<0.05$) all the subspecies and species of *L. punctata* in an identical manner (Table 18). We found low PD value between the Ganga and other *andersoni* group (0.117). The PD value further indicates that the genetic differentiation between *andersoni* and *scutata* was much higher (0.988) than the *L. p. vittata* (0.904), *L. p. punctata* (0.967), and *L. ceylonensis* (0.986).

Table 12: Sample size (n), number of haplotypes (H), haplotype diversity (h), nucleotide diversity (p), Fu's Fs, Tajima's D, mismatch distribution test (SSD) and raggedness index (r) in *Lissemys* populations

	L. p. andersoni		*L. p. punctata*	*L. p. vittata*	*L. ceylonensis*	*L. scutata*
	Group I	**Group II**				
n	69	11	9	24	5	6
h	12	2	4	14	4	4
S	10	1	9	60	7	7
hd (±SD)	0.694 (0.054)	0.327 (0.053)	0.750 (0.112)	0.906 (0.047)	0.900 (0.161)	0.867 (0.129)
π (±SD)	0.0009 (0.0001)	0.0002 (0.0001)	0.003 (0.0005)	0.013 (0.001)	0.0030 (0.0008)	0.0028 (0.00067)
Ts/Tv	7/3	1/0	6/3	37/25	6/1	6/1
Neutrality tests						
Tajima's *D* (P)	-1.310 (0.05)	-1.100 (0.25)	0.957 (0.83)	-0.190 (0.46)	0.082 (0.52)	0.128 (0.47)
Fu's Fs (P)	-6.84 (0.00)	0.356 (0.35)	2.00 (0.83)	0.909 (0.646)	-0.12801 (0.368)	0.31383 (0.548)
Mismatch analysis						
SSD (P)	0.003 (0.35)	0.002 (0.35)	0.181 (0.05)	0.030 (0.03)	0.05 (0.46)	0.315 (0.62)
Rd	0.063 (0.44)	0.226 (0.39)	0.523 (0.02)	0.042 (0.03)	0.130 (0.83)	0.115 (0.68)

Note: n, sample size; h, number of haplotypes; S, number of polymorphic sites; hd, haplotype diversity; π, nucleotide diversity; Ts, transition sites; Tv, transversion sites; sum of squared deviations (SSD), and the raggedness index (r); SD, standard deviation; P, probability value.

Table 13: mtDNA-based estimates of pairwise difference between matrilines of *Lissemys* populations

	L. p. andersoni G-I	*L.p. andersoni* G-II	*L. p. punctata*	*L. p. vittata*	*L. ceylonensis*	*L. scutata*
***L. p. andersoni* G-I**	-					
***L. p. andersoni* G-I**	0.117	-				
L. p. vittata	0.904	0.806	-			
L. p. punctata	0.967	0.960	0.761	-		
L. ceylonensis	0.986	0.987	0.866	0.964	-	
L. scutata	0.988	0.988	0.887	0.969	0.971	-

Note: Number of permutations: 1023; Significance Level=0.05.

4.4.2 Phylogenetic and network relationship

To understand the phylogenetic relationships among the identified subspecies and species of AFT, we constructed a ML tree based on 38 haplotypes. The tree represented the five distinct clusters with significant bootstrap support values (>57%) (Fig.23). Clade-*I* consisted of twelve haplotypes (H1 to H12) represented the sequences of Ganga clustered with other *andersoni* populations. Clade-*II* representing the haplotype of *L. p. vittata* form the sister group with Clade-*I* with bootstrap support ~71%*.* Clade-*III* comprised the haplotypes belong to *L. p. punctata* whereas Clade-*IV* and Clade-*V* represents the sequences of *L. ceylonensis* from Sri Lanka and *L. scutata* from Myanmar, respectively. The distribution and connectivity of AFT haplotypes were inferred from the median-joining network. The results of the network analysis were consistent with those of the phylogenetic analyses, illustrating five clustered clade. Two haplotypes H1 and H6 shared a majority of genepool from Ganga sites as well as the sequences belongs other localities of Uttar Pradesh, Assam, Haryana (Yamuna River), Bangladesh and Myanmar. No clear structuring was observed in seven sites of Ganga River. The haplotypes of *L. p. vittata*, *L. p. punctata*, *L. ceylonensis*, and *L. scutata* formed their cluster in network analysis (Fig. 24).

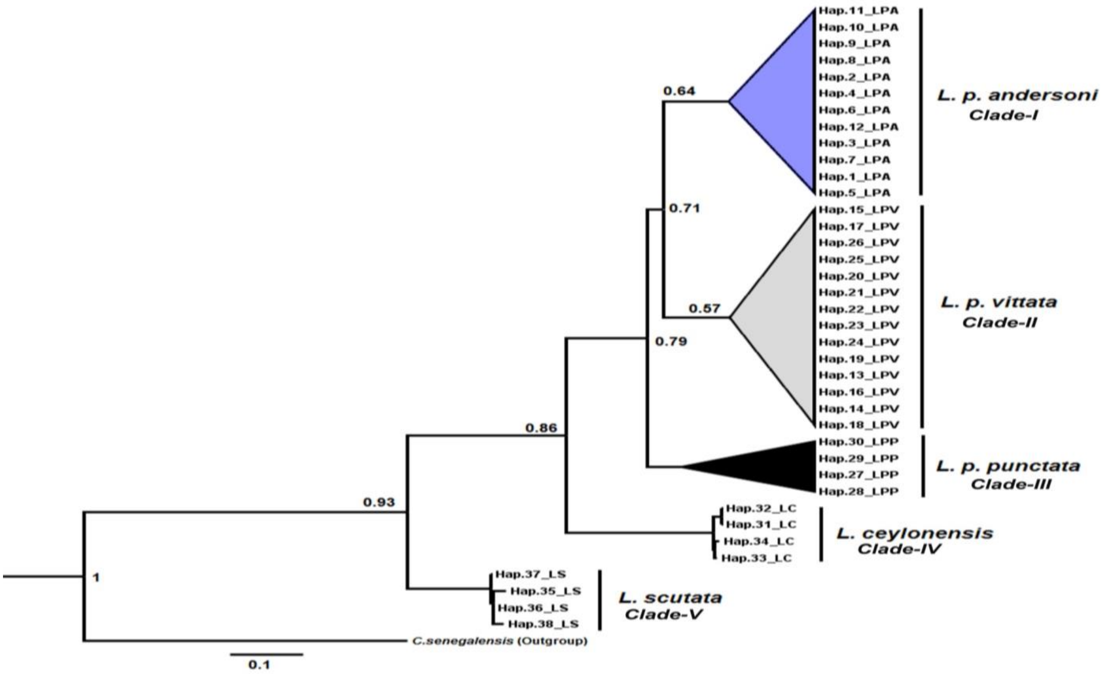


Figure 17: Phylogenetic tree inferred from maximum likelihood (raxmlGUI 2.0 beta) analyses using complete cytochrome *b* gene among *Lissemys*. Bootstrap support values for maximum likelihood (≥ 50% BS) are shown at the node of the tree. *Cyclanorbis senegalensis* (FR850654) was used as an outgroup.

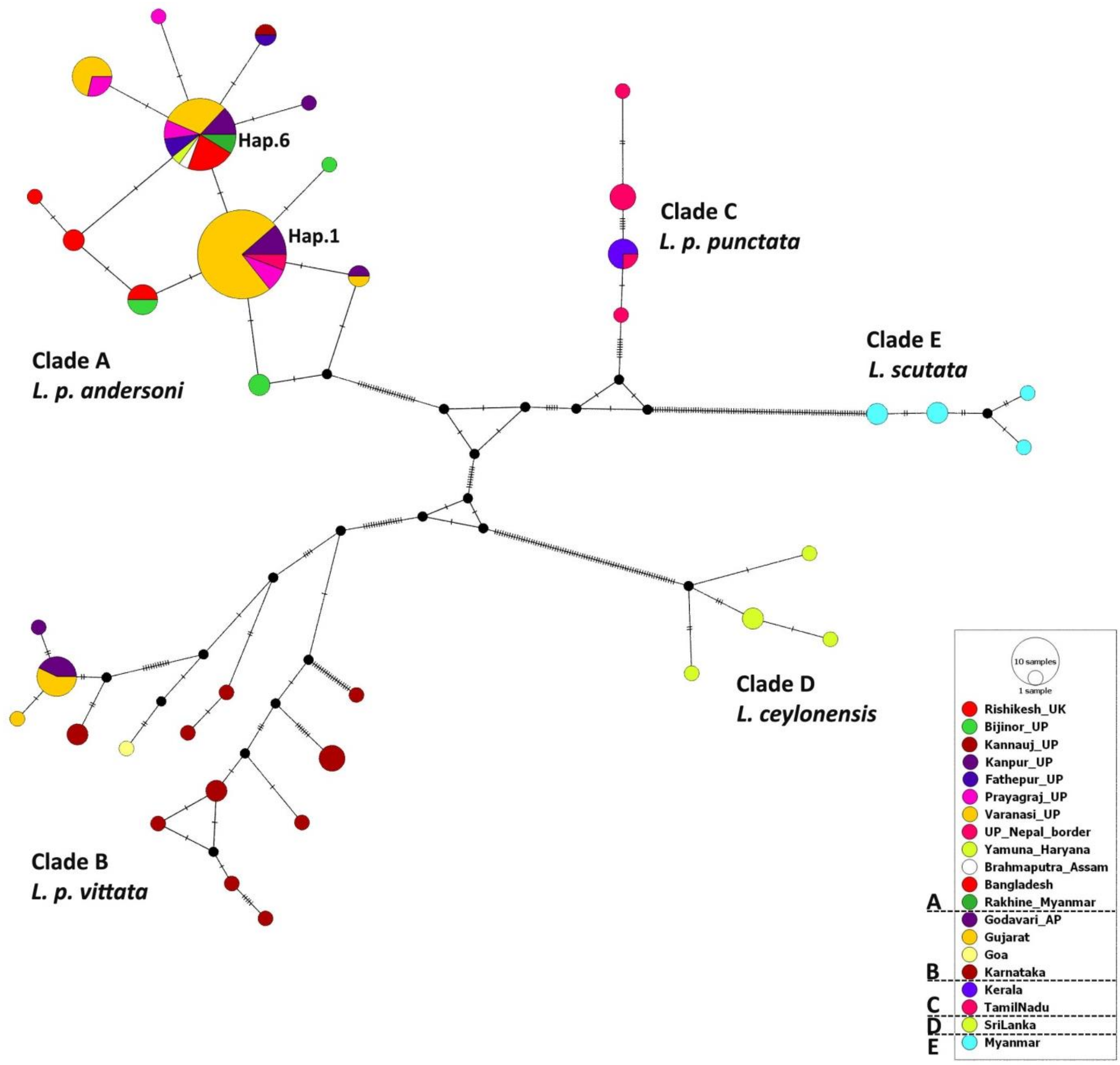


Figure 18: Median Joining network showing the distribution of haplotypes among *Lissemys*. Sharing of haplotypes were labeled among the *L. andersoni*.

4.4.3 Divergence dating and demographic history

The divergence time between Trionychidae and Carettochelyidae was estimated at 146 MA (95% high posterior density [HPD] = 126-173). On this basis, we estimated that the genus *Lissemys* split earlier from the other listed softshell genus during Cretaceous, around 116.32 MA (95% HPD = 94.66–139.59). Within *Lissemys*, the diversification events were mainly occurred in two periods, namely Oligocene and Miocene (Fig.25). The *L. scutata* split first around 29.96 MA (95% HPD = 22.72–37.70), then *L. ceylonensis* is about 23.21 MA (95% HPD = 17.74–29.71). The molecular divergence between lineages of *L. punctata* suggests a divergence date of approximately 9.38 MA (95% HPD = 6.67–12.33).

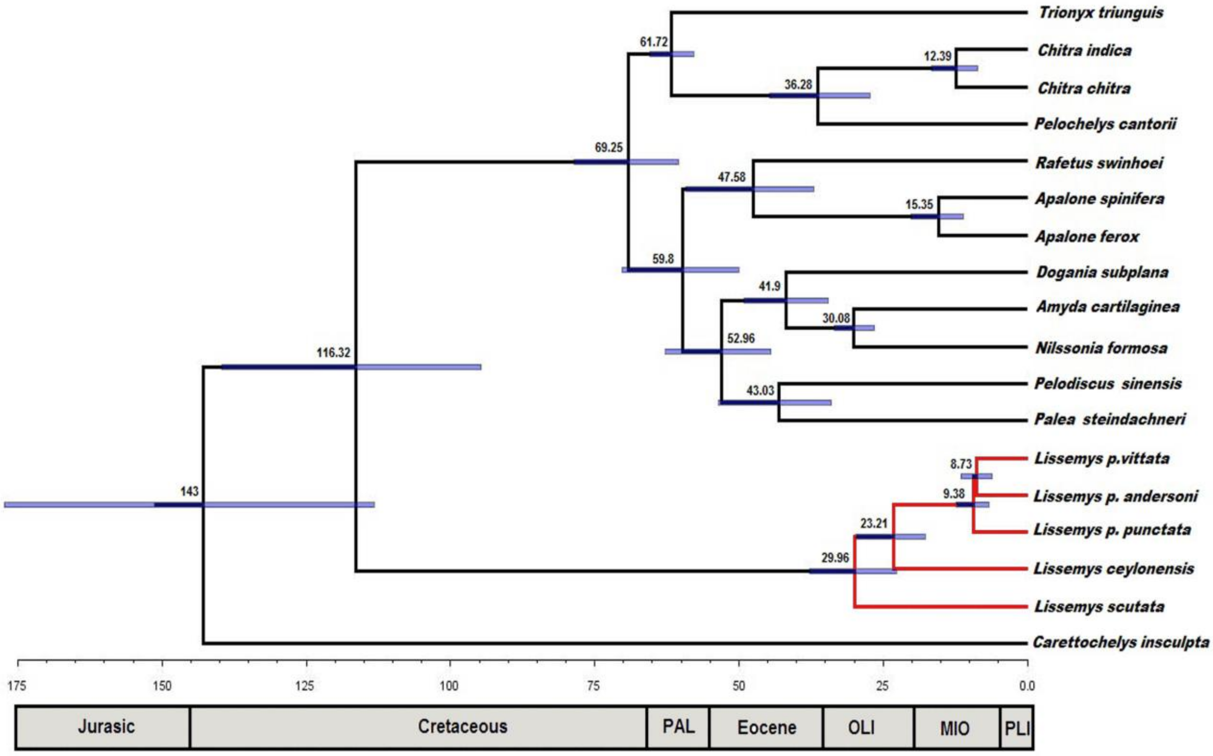


Figure 19: Divergence times estimation using the program BEAST 1.8. The blue lines indicate the 95% confidence interval values of each node. The red clade represents the lineage of Asian flapshell turtles. Note: The periods boundary is according to Geologic Time Scale v. 5.0 (Walker et al., 2018). PAL Paleocene, OLI Oligocene, MIO Miocene, PLI Pliocene.

We performed the Tajima's *D* and Fu's F*s* tests to infer the demographic history of all *L. punctata* (Table 17). The significant negative Tajima's *D* value was observed in Ganga populations whereas it was non-significant negative for *L. p. andersoni* group-II and *L. p. vittata*. The non-significant positive values for Tajima's *D* were observed in *L. p. punctata, L. ceylonensis* and *L. scutata*. No statistical significance values for Fu's F*s* were observed in all AFT populations except Ganga where it was negative. The significant negative values for Tajima's *D* and Fu's F*s* in Ganga indicating an excess of rare nucleotide site variants and will be taken as evidence of deviation caused by population growth or selection. The pairwise mismatch distribution analysis of Ganga populations also supported the demographic expansion which showed unimodal shaped curve. The multimodal and ragged-shaped graphs was observed in rest of the AFT suggests a structured and stable population size, which are an indications of a demographic equilibrium, whereas the mismatch distribution plot for the *L. p. andersoni* group-II was smooth and unimodal, indicating a sign of population expansion (Fig. 26). We also calculated the sum of square deviation (SSD) and raggedness index (Rd) under the demographic expansion model for each AFT populations. We found that all populations had a non-significant raggedness index except *L. p. punctata* and *L. p. vittata* (Table 17) (Harpending 1994). Bayesian skyline plot (BSP) analyses also supported the hypothesis of population expansion in *andersoni* populations and a long phase demographic stability were observed in the rest of the AFT (Fig. 27) from late Pleistocene onwards.

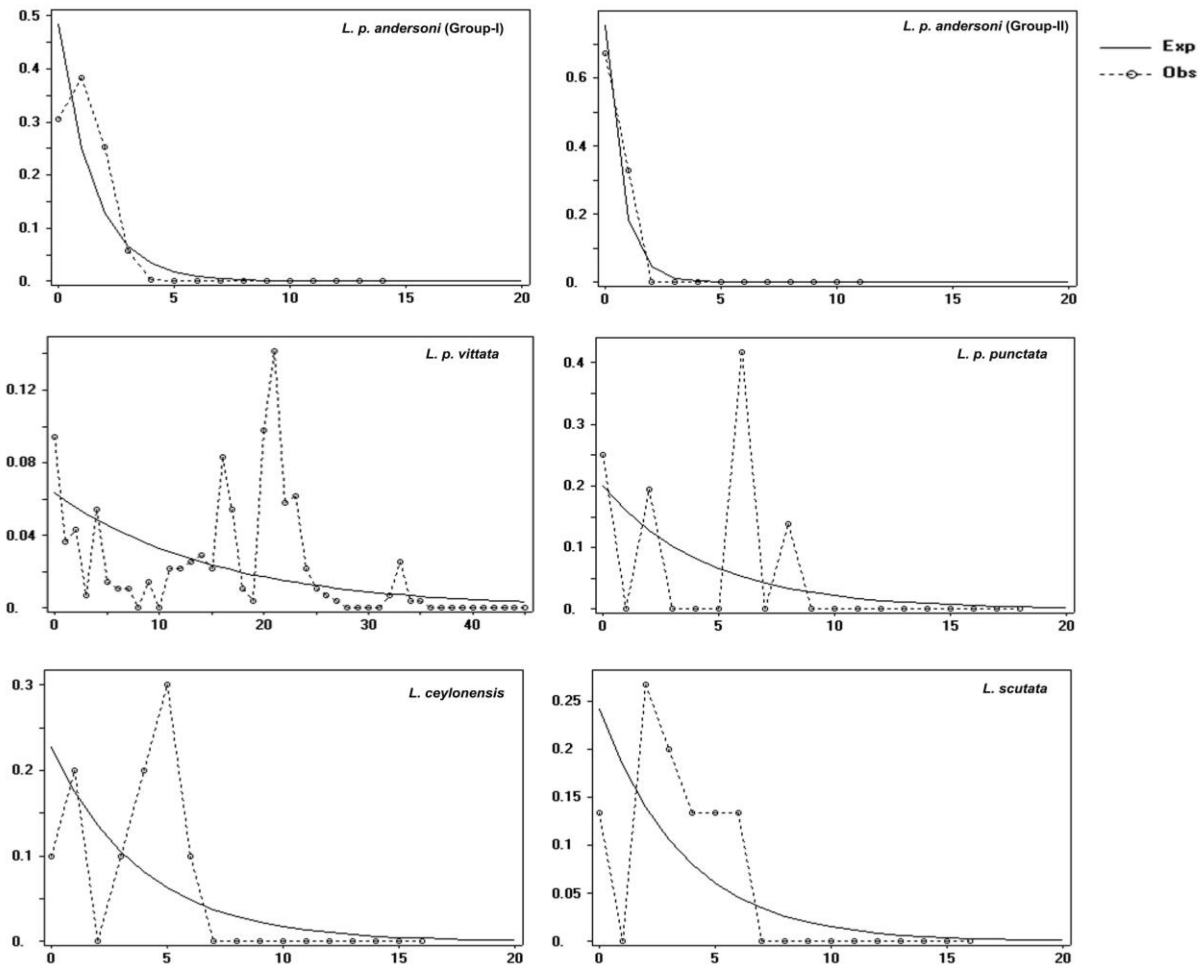


Figure 20: Observed and expected mismatch distributions for *Lissemys*. Dashed line showing observed distribution; solid line showing the theoretical expected distribution under a population expansion model.

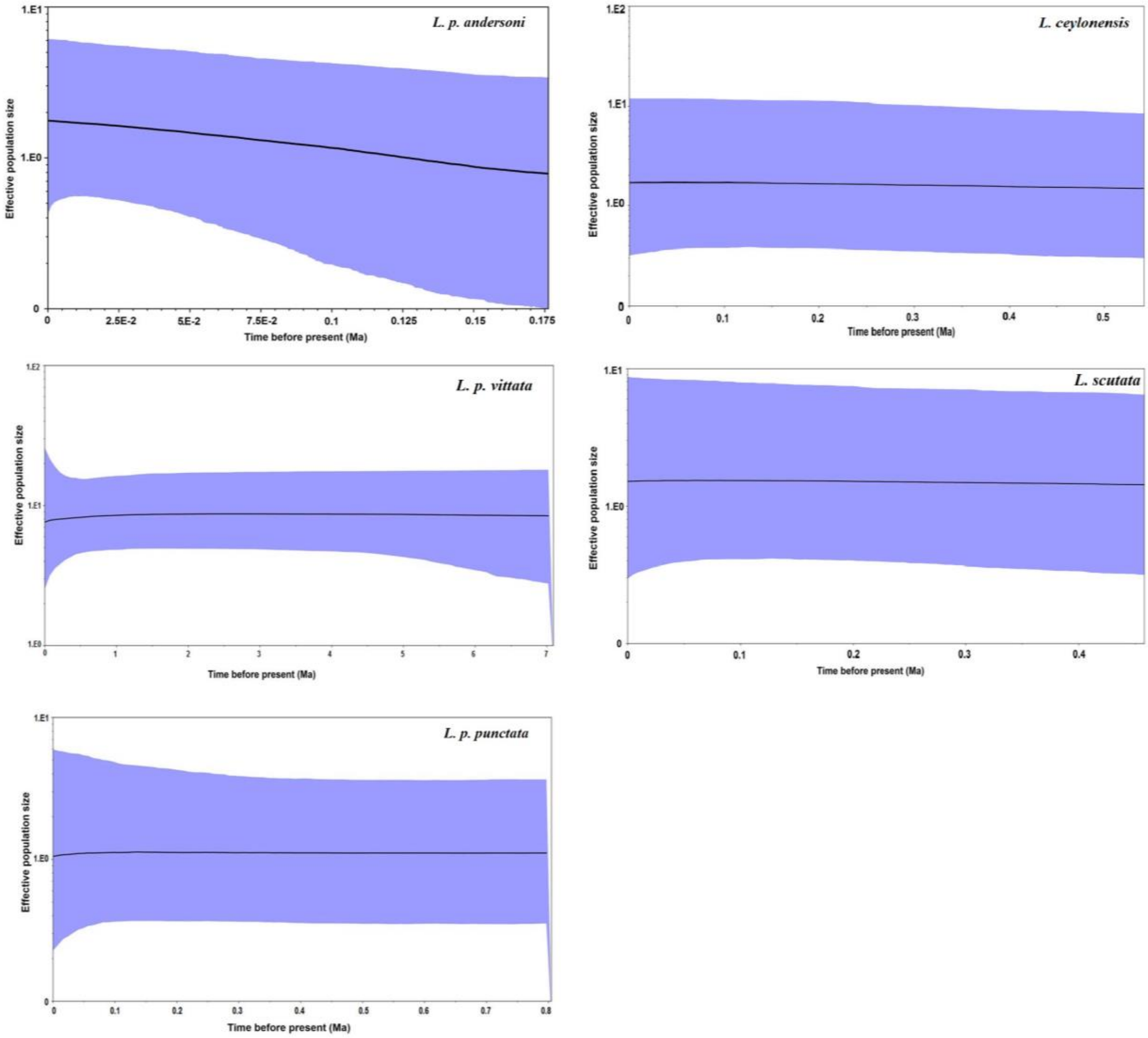


Figure 21: Bayesian skyline plots (BSPs) estimated by BEAST for *Lissemys* populations in the *X*-axis shows time before present (Ma). The *Y*-axis (logarithmic scale) expressed the effective population size (Ne) estimates multiplied by generation time (τ). The solid line indicates the median of population size, and the 95% highest posterior density (HPD) credibility interval is depicted in blue.

4.5 Discussion

The Ganga is the largest river of India, and its basin supports home to more than 600 million peoples (World Bank Report No. STEP705 2011). The Ganga river is passages through a distinct biogeographic zone of India and is notable for rich turtle diversity (Rodgers and Panwar 1988). Understanding the knowledge about genetic diversity, geneflow and demography is of critical importance for implementing the conservation management plan (Fraser and Bernatchez 2001). Species having wide distribution range are faced dynamic environment; thus identification of lineage and its diversification over evolutionary time are of great importance (Frankham 1995; Gupta et al. 2018). Previous molecular study suggested the existences of five distinct AFT lineages (Praschag et al. 2011). Our analysis agrees with this study, but one the major limitation is that the sequences of *L. p. punctata* from Kerala and Tamil Nadu were wrongly submitted as *L. p. vittata* whereas the *L. p. vittata* is mentioned as *L. p. punctata* in NCBI Genbank. Therefore, caution will be taken while assessing these sequences with proper confirmation of the source of origin. In view of this, prior knowledge of the range, source of origin of extant lineage and its genetic diversity is useful to understand the present scenario of population trend and molecular tracking confiscated seizures.

This study highlights together, for the first time analysis of mitochondrial and nuclear DNA variations among northern populations covering seven different localities of Ganga River to better understand the genetic diversity, population structure, geneflow and demography history in relation to the other extant AFT.

4.5.1 Genetic diversity and population structure

The pattern of genetic diversity varies within the AFT lineages found in different geographic ranges. The major *andersoni* populations of Ganga River exhibited a moderate level of genetic diversity in mtDNA cyt *b* region and *CMOS*, whereas it was high in *R35* region. In contrast, a relatively very low level of genetic diversity was observed in the *andersoni* group *II*, where widely distributed samples belongs to three different localities of India, Bangladesh and Myanmar were clustered within two haplotypes. Interestingly, sharing of sequences within Ganga and other *andersoni* populations indicates high geneflow, possibly due to high contemporary migration between Pakistan, northern India, Nepal, Bangladesh and Rakhine State of western Myanmar. We did not detect any significant barrier of geneflow within *andersoni* populations. The comparatively high levels of genetic diversity was observed in *L. p. vittata* from central to peninsular India and *L. ceylonensis* which is only restricted to the Sri Lanka. The moderate to high level of haplotype diversity couple with low nucleotide diversity was observed in Ganga can be a signature of sudden population growth due to accumulation of new mutations from a small effective population size (Avise et al. 1984; Rogers and Harpending 1992). The results were consistent with the ecological status of turtles in Asia, where half of the species are under endangered and 28% are considered under Critically Endangered categories (Van Dijk 2000; Engstrom et al. 2002). Furthermore, low nucleotide diversity might be suspected in other AFT (except *L. p. vittata*) where population growth suffering due to high anthropogenic pressure results in habitat destruction, poaching and overexploitation.

The phylogenetic and network analysis also revealed the five distinct mtDNA lineages of AFT from its distribution ranges. The distribution ranges of AFT, both the spotted and unspotted group, sometimes overlap. These species are quite different in morphological features such as pattern of head, neck, carapace shape and color (Smith 1931; Webb and RG 1980; Webb 1982; Ernst and Barbour 1989; Ernst et al. 2000; Praschag et al. 2011). Despite the clear morphological character between spotted subspecies *L. p. andersoni* and unspotted *L. p. vittata*, intergradations between these subspecies was recorded in few localities of Odisha, Indian state (Praschag et al. 2011). The identification of hybrid individuals is sometimes challenging because of morphological similarity, hence establishing a genetic barcode for molecular tracking and can help for guiding the adequate conservation strategies.

4.5.2 Genetic differentiation and divergence estimation

The significant pairwise genetic differentiation was detected between five extant AFT lineages, due to their specific habitat, dispersal ability and reproductive isolation. Praschag et al. (2011) reported the high sequences divergence based on uncorrected *p* distances between *L. p. andersoni* and *L. p. vittata* than the *L. p. punctata*, whereas in our study, we observed its vice-versa (high PD value between *L. p. andersoni* and *L. p. punctata*). This high genetic divergence between these two lineages might be the results of amalgamation of hybrid origin (*L. p. andersoni* × *L. p. vittata*) samples from Odisha, Indian state. In comparison to *L. p. andersoni*, the complete cyt *b* shows variations from ~4.9% for *L. p. vittata* to 12.9% for *L. scutata.* These sequences were used to estimate the divergence time along with that of other extant softshell turtles. The evolution of Asian softshell turtle is still in uncertainty and it has been validated thoroughly in previous studies (Le et al. 2014; Li et al. 2017). The *Carettochelys insculpta* is the only extant member of the family Carettochelyidae; it is native to northern Australia and southern New Guinea (Li et al. 2017). We estimated divergence time between Trionychidae and Carettochelyidae was to be around at 143 MA (HPD=113.26–177.4) for this node. The last common ancestor of extant AFT existed in South East Asia and split at around 116.32 MA (HPD=94.66–139.59) during the Cretaceous (Joyce et al. 2013; Le et al. 2014; Li et al. 2017). The diversification rate of softshell turtles was boosted in South East Asia to India due to active during the period of global warming from mid of the Oligocene to Miocene and due to paleoclimate change between the Miocene and Pliocene (Duellman 1999; Le et al. 2014). The diversification of softshell turtles, plants, other aquatic species and mammals showed similar dispersal and speciation pattern during the warmer climate (Duellman 1999; Moreau et al. 2006; Smith and Rose 2006). Due to change in climate, *L. ceylonensis* of Sri Lanka were diverged from *L. scutata* of Myanmar at around 30 MA (HPD=22.72–37.70) followed by *L. punctata* at around 23.21 MA (HPD=17.74–29.71). During this period, a Palaeogene biogeographic link was survived between South-East Asia and India. The northward drifting of Indian plate towards Asia occupied the north-east corner of the subcontinent coming into contact with Sumatra and Burma at around 57 MA ago (early Eocene), which was followed by a hard collision at around 35 MA (early Oligocene) with Asia (Ali & Aitchison 2008, Beck et al. 1995). Further, the *L. punctata* were dispersed and split to *L. p. punctata, L. p. vittata* and *L. p. andersoni* during the late Miocene between mean split age approx. 8 - 10 MA (CI=6.75-12.33). However, Praschag et al. (2011) reported considerably lower divergence estimated around 4.2-4.5 MA for these nodes. The discrepancy in results of the earlier study was might be due to the calibration of divergence time using much younger fossil group than the age of species.

The results from BSP, neutrality and mismatch distribution strongly indicate population expansion of *L. p. andersoni* in the Ganga River and other northern populations from the ~1.25 MA in late Pleistocene (Fig 27). The scenario in the *L. p. vittata*, *L. p. punctata*, *L. ceylonensis* and *L. scutata* indicate a stable demographic situation in the effective population size with no sign of population expansion. This result was consistent with previous findings of Hawksbill turtle (Vargas et al. 2016), leatherback turtles (Molfetti et al. 2013), *Schizothorax o'connori* from the Yarlung Tsangpo River (Guo et al. 2016); *Schizothorax nukiangensis* from Nujiang River (Chen et al. 2015), where climatic fluctuations during the Holocene and Pleistocene play a major role in dramatic demographic changes over species growth. The sea-level fluctuation in the Pleistocene period greatly influences the environment, alternation of land, river dimension and existence of natural sources (Simanjuntak 2006). These change conditions call for adaptation processes and play a significant role in influencing the activity and mobility of species over time. However, in our analysis, low sample size in few AFT populations are likely to reduce the power for estimating the demographic history (Ho and Shapiro 2011). Hence, assumptions involved in BSP analysis suggest that this should be interpreted cautiously (Grant 2015).

4.6 Conclusion

Assessments of genetic diversity, geneflow and population demography are competent aspects for conservation programs. Maintaining a high level of genetic diversity is assumed to be essential for the conservation of viable populations (Reed et al. 2002). Despite low nucleotide diversity, the Ganga population shows signs of expansion in effective population size, whereas the rest of AFT population showed a stable phase. It necessitating urgent targeted conservation efforts to augment the diversity and maintain gene flow of the species. The resolution power in delimiting taxa, especially in cases when hybridizations are now occurring and extant lineages are morphologically very similar, guiding the adequate conservation strategies for any lineages is a difficult task. Therefore, the signature of mtDNA and nuclear gene can play a pivotal role in wildlife forensic to track the origin of poached species. However, due to limited sampling under the current study, more research is required to compare and resolve the genetic story of the species with other tributaries of Ganga. To save these jewels of evolution, all stakeholders must come together and act to know more about the turtle populations and their demographic history to better understand the threats and how to combat them. Specialized administration strategies to conserve and maintain safe habitats for these ancient creatures stand as the need of the hour.

4.7 References

- Ali, J. R., & Aitchison, J. C. (2008). Gondwana to Asia: plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through latest Eocene (166–35 Ma). *Earth-Science Reviews*, 88, 145-166.
- Annandale, N. (1912). The Indian mud-turtles (Trionychidae). *Records of the Indian Museum*, 7, 151-180.
- Avise, J. C., Bowen, B. W., Lamb, T., Meylan, A. B., & Bermingham, E. (1992). Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. *Molecular biology and evolution*, 9, 457-473. https://doi.org/10.1093/oxfordjournals.molbev.a040735
- Avise, J. C., Neigel, J. E., & Arnold, J. (1984). Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *Journal of Molecular Evolution*, 20, 99-105. https://doi.org/10.1007/BF02257369
- Beck, R. A., Burbank, D. W., Sercombe, W. J., Olson, T. L., & Khan, A. M. (1995). Organic carbon exhumation and global warming during the early Himalayan collision. *Geology*, 23, 387-390. https://doi.org/10.1130/0091-7613(1995)023<0387:OCEAGW>2.3.CO;2
- Bhaskar, R., & Mohindra, V. (2019). Phylogenetic relationships among Indian freshwater turtles (family Trionychidae and Geoemydidae) with special reference to *Lissemys punctata*, inferred from mitochondrial cytochrome b gene sequences. *Meta Gene*, 22, 100610. https://doi.org/10.1016/j.mgene.2019.100610
- Bhupathy, S., Webb, R. G., & Praschag, P. (2014). *Lissemys punctata* (Bonnaterre 1789)– Indian Flapshell turtle. Conservation Biology of freshwater Turtles and Tortoises. *Chelonian Research Monographs*, 076-1.
- Boulenger, G. A. (1889). Catalogue of the chelonians, rhynchocephalians, and crocodiles in the British Museum (Natural History). order of the Trustees..
- Chen, W., Du, K., & He, S. (2015). Genetic structure and historical demography of *Schizothorax nukiangensis* (Cyprinidae) in continuous habitat. *Ecology and evolution*, 5, 984-995. https://doi.org/10.1002/ece3.1413
- Choudhury, B. C., Bhupathy, S., & Hanfee, F. (2000). Status information on the tortoises and freshwater turtles of India. In Asian Turtle Trade: Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia. *Chelonian Research Monographs*, 2, 86-94.
- Das, K. C., & Gupta, A. (2011). Site records of softshell turtles (Chelonia: Trionychidae) from Barak Valley, Assam, northeastern India. *Journal of Threatened Taxa*, 1722-1726. https://doi.org/10.11609/JoTT.o2487.1722-6
- Deraniyagala, P. E. P. (1939). The Tetrapod Reptiles of Ceylon: Volume 1 Testudinates and Crocodilians. *Colombo Museum*.
- Deraniyagala, P. E. P. (1953). A colored atlas of some vertebrates from Ceylon: *tetrapod Reptilia* (Vol. 2). Ceylon Govt. Press.
- Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution*, 29, 1969-1973. https://doi.org/10.1093/molbev/mss075
- Duellman, W. E. (Ed.). (1999). Patterns of distribution of amphibians: a global perspective. *JHU Press*.
- Edler, D., Klein, J., Antonelli, A., & Silvestro, D. (2019). raxmlGUI 2.0 beta: a graphical interface and toolkit for phylogenetic analyses using RAxML. bioRxiv, 800912. https://doi.org/10.1101/800912

- Engstrom, D. W., & Okamura, A. (2004). Working with survivors of torture: Approaches to helping. *Families in Society*, 85, 301-309. https://doi.org/10.1177/104438940408500304
- Engstrom, T. N., Shaffer, H. B., & McCord, W. P. (2002). Phylogenetic diversity of endangered and critically endangered southeast Asian softshell turtles (Trionychidae: Chitra). *Biological Conservation*, 104, 173-179. https://doi.org/10.1016/S0006-3207(01)00161-6
- Ernst, C. H., & Barbour, R. W. (1989). Turtles of the World Smithsonian Institution Press. Washington, *DC*, 313.
- Ernst, C. H., Altenburg, R. G. M., & Barbour, R. W. (2000). Turtles of the world. World biodiversity database, version 1.2. *Biodiversity Center of ETI, Amsterdam*, CD-ROM.
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, 10, 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x
- Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131, 479-491.
- Frankham, R. (1995). Conservation genetics. *Annual review of genetics*, 29, 305-327.
- Fraser, D. J., & Bernatchez, L. (2001). Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular ecology*, 10, 2741-2752. https://doi.org/10.1046/j.0962-1083.2001.01411.x
- Fritz, U., & Havaš, P. (2007). Checklist of chelonians of the world. *Museum of Zoology*.
- Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147, 915-925.
- Fujita, M. K., Engstrom, T. N., Starkey, D. E., & Shaffer, H. B. (2004). Turtle phylogeny: insights from a novel nuclear intron. *Molecular phylogenetics and evolution*, 31, 1031-1040. https://doi.org/10.1016/j.ympev.2003.09.016
- Grant, W. S. (2015). Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography. *Journal of Heredity*, 106, 333-346. https://doi.org/10.1093/jhered/esv020
- Guo, X. Z., Zhang, G. R., Wei, K. J., Yan, R. J., Ji, W., Yang, R. B., ... & Gardner, J. P. (2016). Phylogeography and population genetics of *Schizothorax o'connori*: strong subdivision in the Yarlung Tsangpo River inferred from mtDNA and microsatellite markers. *Scientific reports*, 6, 1-15. https://doi.org/10.1038/srep29821
- Gupta, S. K., Kumar, A., Angom, S., Singh, B., Ghazi, M. G. U., Tuboi, C., & Hussain, S. A. (2018). Genetic analysis of endangered hog deer (*Axis porcinus*) reveals two distinct lineages from the Indian subcontinent. *Scientific reports*, 8, 1-12. https://doi.org/10.1038/s41598-018-34482-9
- Hall, T. A. (1999, January). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, *Nucleic acids symposium series* 41, 95-98. [London]: Information Retrieval Ltd., c1979-c2000.
- Harpending, H. C. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human biology*, 591-600. https://www.jstor.org/stable/41465371
- Ho, S. Y., & Shapiro, B. (2011). Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular ecology resources*, 11, 423-434. https://doi.org/10.1111/j.1755-0998.2011.02988.x
- Hora, S. L. (1951). Some observations on the palaeogeography of the Garo-Rajmahal Gap as evidenced by the distribution of Malayan fauna and flora to peninsular India. *Proceedings of the Natural Institute of Sciences of India*, 16, 437-444.

- Hossain, M. L., Sarker, S. U., & Sarker, N. J. (2010). Breeding biology aspects of spotted flapshell Turtle, *Lissemys punctata* (Lacepede 1788), in Bangladesh. *Saudi journal of biological sciences,* 17, 5-12. https://doi.org/10.1016/j.sjbs.2009.12.002
- https://doi.org/ 10.1016/j.earscirev.2008.01.007
- Joyce, W. G., Parham, J. F., Lyson, T. R., Warnock, R. C., & Donoghue, P. C. (2013). A divergence dating analysis of turtles using fossil calibrations: an example of best practices. *Journal of Paleontology*, 87, 612-634. https://doi.org/10.1666/12-149
- Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in ecology & evolution*, 17, 230-241.
- Lalitha, R., & Chandavar, V. R. (2018). Intraspecific variations in *Cyt b* and D-loop sequences of Testudine species, *Lissemys punctata* from south Karnataka. *Journal of advanced research*, 9, 87-95. https://doi.org/10.1016/j.jare.2017.10.007
- Le, M., Duong, H. T., Dinh, L. D., Nguyen, T. Q., Pritchard, P. C., & McCormack, T. (2014). A phylogeny of softshell turtles (Testudines: Trionychidae) with reference to the taxonomic status of the critically endangered, giant softshell turtle, *Rafetus swinhoei*. *Organisms Diversity & Evolution*, 14, 279-293. https://doi.org/10.1007/s13127-014-0169-3
- Le, M., Raxworthy, C. J., McCord, W. P., & Mertz, L. (2006). A molecular phylogeny of tortoises (Testudines: Testudinidae) based on mitochondrial and nuclear genes. *Molecular Phylogenetics and Evolution*, 40, 517-531. https://doi.org/10.1016/j.ympev.2006.03.003
- Leigh, J. W., & Bryant, D. (2015). popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6, 1110-1116. https://doi.org/10.1111/2041-210X.12410
- Lenk, P., Fritz, U., Joger, U., & Wink, M. (1999). Mitochondrial phylogeography of the European pond turtle, *Emys orbicularis* (Linnaeus 1758). *Molecular ecology*, 8, 1911-1922. https://doi.org/10.1046/j.1365-294x.1999.00791.x
- Li, H., Liu, J., Xiong, L., Zhang, H., Zhou, H., Yin, H., ... & Liu, J. (2017). Phylogenetic relationships and divergence dates of softshell turtles (Testudines: Trionychidae) inferred from complete mitochondrial genomes. *Journal of evolutionary biology*, 30, 1011-1023. https://doi.org/10.1111/jeb.13070
- Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451-1452. https://doi.org/10.1093/bioinformatics/btp187
- Menon, A. G. K. (1951). Further studies regarding Hora's Satpura hypothesis. The role of Eastern Ghats in the distribution of the Malayan Fauna and Flora to peninsular India. In *Proceedings of the National Institute of Science India*, 17, 475-497.
- Mertens, R., & Wermuth, H. (1955). Die rezenten Schildkröten, Krokodile und Brückenechsen. Zoologische Jahrbücher. Abteilung für Allgemeine Systematik, Ökologie und Geographie der Tiere. *Gustav Fischer Verlag, Jena*, 83, 323-440.
- Meylan, P. A. (1987). The phylogenetic relationships of soft-shelled turtles (family Trionychidae). *Bulletin of the AMNH*; v. 186, article 1.
- Molfetti, E., Vilaça, S. T., Georges, J. Y., Plot, V., Delcroix, E., Le Scao, R., ... & De Thoisy, B. (2013). Recent demographic history and present fine-scale structure in the Northwest Atlantic leatherback (*Dermochelys coriacea*) turtle population. *PLoS One*, 8(3). https://doi.org/10.1371/journal.pone.0058061
- Moreau, C. S., Bell, C. D., Vila, R., Archibald, S. B., & Pierce, N. E. (2006). Phylogeny of the ants: diversification in the age of angiosperms. *science*, 312, 101-104. https://doi.org/10.1126/science.1124891

- Praschag, P., Stuckas, H., Paeckert, M., Maran, J., & Fritz, U. (2011). Mitochondrial DNA sequences suggest a revised taxonomy of Asian flapshell turtles (Lissemys Smith, 1931) and the validity of previously unrecognized taxa (Testudines: Trionychidae). *Vertebrate Zoology*, 61, 147-160.
- Rambaut, A. (2012). FigTree v1. 4.0. A graphical viewer of phylogenetic trees. See http://tree. bio. ed. ac. uk/software. figtreetree.
- Reed, D. H., Briscoe, D. A., & Frankham, R. (2002). Inbreeding and extinction: the effect of environmental stress and lineage. *Conservation Genetics*, 3, 301-307. https://doi.org/10.1023/A:1019948130263
- Rhodin, A. G., van Dijk, P. P., Iverson, J. B., & Shaffer, H. B. (2010). Turtles of the world, 2010 update: annotated checklist of taxonomy, synonymy, distribution, and conservation status. *Chelonian Research Monographs*, 5, 000-85.
- Rodgers, W. A., & Panwar, H. S. (1988). Planning a wildlife protected area network in India.
- Rogers, A. R., & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular biology and evolution*, 9, 552-569. https://doi.org/10.1093/oxfordjournals.molbev.a040727
- Siebenrock, F. (1909). Synopsis der rezenten Schildkröten: mit Berücksichtigung der in historischer Zeit ausgestorbenen Arten (No. 3). G. Fischer.
- Simanjuntak, T. (2006). Indonesia–Southeast Asia: climates, settlements, and cultures in Late Pleistocene. *Comptes Rendus Palevol*, 5, 371-379. https://doi.org/10.1016/j.crpv.2005.10.005
- Smith M. A. (1931) Fauna of British India, Including Ceylon and Burma-Reptilia and Amphibia. Vol. 1-Loricata, Testudines. *Taylor and Francis*, London.
- Smith, T., Rose, K. D., & Gingerich, P. D. (2006). Rapid Asia–Europe–North America geographic dispersal of earliest Eocene primate Teilhardina during the Paleocene–Eocene thermal maximum. *Proceedings of the National Academy of Sciences*, 103, 11223-11227. https://doi.org/10.1073/pnas.0511296103
- Spinks, P. Q., Shaffer, H. B., Iverson, J. B., & McCord, W. P. (2004). Phylogenetic hypotheses for the turtle family Geoemydidae. *Molecular phylogenetics and evolution*, 32, 164-182. https://doi.org/10.1016/j.ympev.2003.12.015
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123, 585-595.
- Talukdar, S. K. (1979). *Lissemys punctata punctata* (Bonnaterre)[Testudines: Trionychidae]: An edition to the chelonian fauna of the Brahmaputra drainage, Assam. *Indian J. Zootomy*, 20, 181.
- The World Bank Report No. STEP705 (2011). The National Ganga River Basin Project.
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, 22, 4673-4680. https://doi.org/10.1093/nar/22.22.4673
- Uetz, P., Hošek, J., & Hallermann, J. (2016). The reptile database. 2016. Database accessible at: http://www. reptile-database. org.
- Van Dijk, P. P. (2000). The status of turtles in Asia. *Chelonian Research Monographs*, 2, 15-23.
- Vargas, S. M., Jensen, M. P., Ho, S. Y., Mobaraki, A., Broderick, D., Mortimer, J. A., ... & Hoenner, X. (2016). Phylogeography, genetic diversity, and management units of hawksbill turtles in the Indo-Pacific. *Journal of Heredity*, 107, 199-213. https://doi.org/10.1093/jhered/esv091
- Webb, R. G. (1982). Taxonomic notes concerning the trionychid turtle *Lissemys punctata* (Lacepède). *Amphibia-Reptilia*, 3, 179-184. https://doi.org/10.1163/156853882X00400

- Webb, R. G., & RG, W. (1980). The identity of Testudo punctata Lacepède, 1788 (Testudines, Trionychidae).
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. evolution, 38(6), 1358-1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x
- Wermuth, H., & Mertens, R. (1961). Schildkröten–Krokodile–Brückenechsen. i–xvi, 1–422 pp. G. Fischer, lena.(9) Christopher J. Raxworthy Department of Herpetology, *American Museum of Natural History, Central Park West at 79th Street, New York, NY,* 10024.
- Wermuth, H., & Mertens, R. (1977). Liste der rezenten Amphibien und Reptilien: Testudines, Crocodylia, Rhynchocephalia. Berlin; New York: De Gruyter.
- Willi, Y., & Hoffmann, A. A. (2009). Demographic factors and genetic variation influence population persistence under environmental change. *Journal of evolutionary biology*, 22, 124-133.

5. POPULATION GENETICS OF THE CRITICALLY ENDANGERED THREE-STRIPED TURTLE, *Batagur dhongoka*, FROM THE GANGA RIVER

5.1 Background

Turtles are the most threatened biota on the earth, with 40% of all extant turtles facing extinction risk (Stanford et al., 2020). All six extant *Batagur* species are categorized as 'Critically Endangered' (CR) in the IUCN Red list. The three-striped roofed turtle, *B. dhongoka* is a freshwater turtle confined to northern India (Ganga River basin), Nepal, and Bangladesh (Das et al., 2019; Praschag et al., 2019). Once inhabiting the Ganga and Brahmaputra basin, the *B. dhongoka* populations are now drastically declining throughout its range by over 80%, due to unsustainable human pressure that has deteriorated riverine habitats (Rashid et al., 2000). As a result, the population of *B. dhongoka* has been extirpated from the Gomti River, a tributary of the Ganga, with a major population now restricted to the National Chambal Sanctuary, India (Das et al., 2019).

Consequently, there is a critical need to reevaluate the present distribution of this species (Choudhury et al., 2000). Habitat loss and degradation is the principal threat to *B. dhongoka* survival, as it degrades nesting sites, nursery habitats and basking sites and decreases food availability. Additionally, this species is used for subsistence consumption and is involved in some commercial trade, with recent confiscations reported in Malaysia and China (Das et al., 2019). Due to the increasing decline in its native range and a corresponding rise in international seizures, *B. dhongoka* has gained the highest level of protection within its native regions. In 2022, it was included in Schedule *I* of the Indian Wildlife (Protection) Act 1972 and added to Appendix II of CITES.

It is essential to implement extended monitoring efforts, identify sustainable populations, and execute a scientifically guided reintroduction throughout their native habitat to ensure the long-term conservation of *B. dhongoka*. In 2006, the state forest department and Turtle Survival Alliance initiated the Chambal Conservation Project in the National Chambal Sanctuary to increase and conserve the population of *B. dhongoka* and *B. kachuga* (The Print, 2000). Further, under the National Mission for Clean Ganga project, Wildlife Institute of India, Dehradun, the Turtle Rescue and Rehabilitation Centre (TRRC) at Sarnath, Varanasi was ameliorated and the Ganga Aqualife Rescue and Rehabilitation Centre (GARRC), Narora at Uttar Pradesh was established. Given the challenges and reduction in the population size of freshwater turtles in Ganga River and as part of conservation efforts Turtle Breeding and Rehabilitation Centers play a crucial role in bolstering population numbers and providing valuable support to management initiatives. As part of a conservation project, the Forest and Wildlife Department rescued turtle eggs from the Chambal River and transferred them to the TRRC, Sarnath, for hatching. After reaching a certain size, healthy individuals were released into the suitable habitat of the Ganga River. The TRRC also received rescued individuals seized by enforcement agencies from local areas. The GARRC exclusively received rescued individuals, comprising those either injured or accidentally entangled in fishing nets within the local vicinity of the Ganga River. Analyzing a declining population's genetic structure and diversity holds profound implications for its conservation, especially when individuals are chosen to increase the effectiveness of endeavors aimed at successful breeding, rehabilitation, and reintroduction (Burns et al., 2003). A decrease in population size can affect long-term evolutionary potential by reducing individual fitness and genetic variation in populations (Reed and Frankham, 2003). The eventual strategy to prevent extinction or increase population size is scientifically informed through rescue and rehabilitation, captive breeding, and reintroducing healthy individuals into wild habitats (Choudhury et al., 2000). However, rescue and rehabilitation, captive breeding and reintroduction in the wild have critical complications. Gathering information on the geographic origin of individuals is sometimes difficult and the release of these unknown origin rescue individuals in the wild is often accompanied by high initial losses of released individuals (Villemey et al., 2013). Hence, while a small number of individuals are released locally, genetic monitoring is vital to understand the wider and long-term contemporary genetic diversity for effective management and conservation interventions.

Conservation genetics provides a robust framework for assessing the extent of genetic diversity and population genetic structure of small and declining populations. It plays a vital role in conservation efforts by developing and utilizing comprehensive genetic databases. Until now, only limited information on ecological and genetic aspects of *B. dhongoka* has been available, apart from a comparative mitogenome analysis between *B. kachuga* and *B.* dhongoka (Kumar et al., 2021). Additionally, the utilization of the genomic approach by Çilingir et al. (2017) on the Burmese roofed turtle (*B. trivittata*) in Myanmar, as well as investigation of the southern river terrapin (*B. affinis*) in Cambodia, provided estimations of wild breeder populations and insights into the present genetic structure. Recent study revealed low genetic diversity and notable genetic differences among the *B. affinis* populations in Peninsular Malaysia (Salleh et al., 2023).

In this study, we investigate the genetic diversity and population structure of *B. dhongoka* using three mitochondrial DNA regions and microsatellite DNA markers. Our study focuses on *B. dhongoka* from two Turtle Rescue and Rehabilitation Centers: (1) TRRC at Sarnath, Varanasi, and (2) GARRC at Narora, Uttar Pradesh. Understanding the genetic diversity and population structure of *B. dhongoka* turtles within these Rescue and Rehabilitation centers can have profound impact on rehabilitation and conservation efforts by adopting evidence-based management practices.

5.2 Methodology

5.2.1 Research permits and ethical considerations

The Uttar Pradesh Forest Department vide Letter No. 2036, 3263 and 1854/23-2-12-G, provided necessary permissions for the survey and collection of the biological samples and Institutional Animal Ethics Committee (Letter No. WII/IAEC/2017-18) approved the sampling protocols and has no objection to carrying out the research. The biological samples were collected as part of a sponsored project under Biodiversity Conservation and Ganga Rejuvenation (Grant no. B-02/2015-16/1259/NMCG-WII-proposal) funded by National Mission for Clean Ganga (NMCG), Ministry of Jal Shakti, Government of India.

5.2.2 Sample collection and DNA extraction

A total of 92 tissue samples (egg shell, swabs from hatchlings and remains of dead individuals) were collected from Turtle Breeding and Rehabilitation Centre (TRRC), Sarnath, Varanasi (n=80) and Ganga Aqualife Rescue and Rehabilitation Centre (GARRC), Narora (n=12) (Fig.22). These tissue samples were preserved in 70% ethanol at room temperature. Both the centers, TRRC and GARRC were established near the Ganga River under the aegis of the project Biodiversity Conservation and Ganga Rejuvenation to provide head-start facilities for turtle conservation intended for release into its natural habitat to aid local species recovery. The distance between both centers is approximately 700 km. The TRRC receives turtle eggs from the Chambal River from the Uttar Pradesh forest department and hatched inside the center. Additionally, rescue cases of turtles were also separately maintained in TRRC and GARRC. The genomic DNA (gDNA) was extracted using the DNeasy Blood Tissue Kit (QIAGEN, Germany) in a final elution volume of 100 µl. The gDNA quantification of samples was done in Qubit 4 flurometer (Invitrogen) and diluted to make 50-100 ng/µl for working stock.

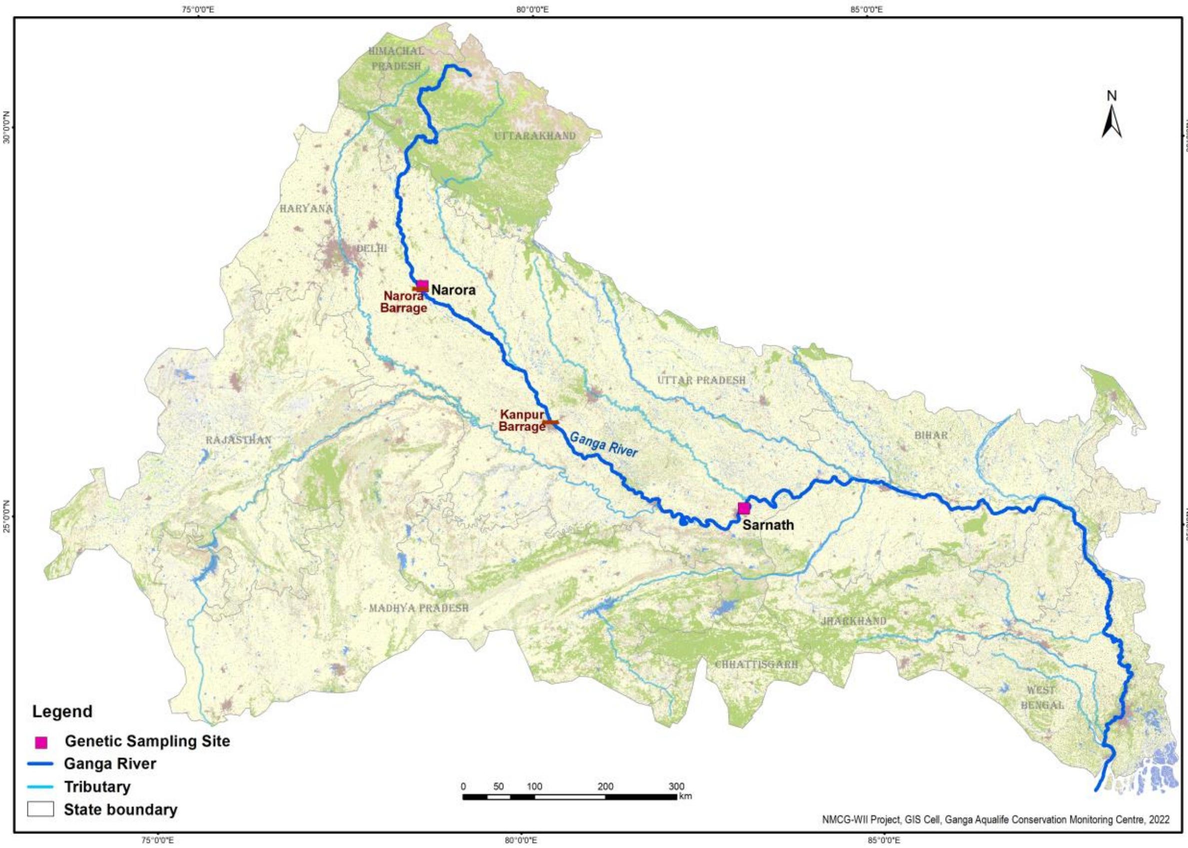


Figure 22: Map of the study area showing sampling sites of *Batagur dhongoka*.

5.2.3 mtDNA amplification and sequencing

Two mitochondrial regions: cytochrome b (Cyt b) gene ~1140 bp and mtDNA control region (CR) ~451bp were amplified using primers: CytbG:5'-AACCATCGTTGTWATCAACTAC-3' and mt-f-na3:5'-AGGGTGGAGTCTTCAGTTTTTGGTTTACAAGACCAATG-3' (Spinks et al., 2004; Lenk et al., 1999) and CRF: 5'-AGCACCGGTCTTGTAAACCA-3' and CRR: 5'-ACAAAACAACCAAAGGCCAG-3' (Gaur et al, 2006). PCR reactions for both regions were performed in 10 µl reaction volumes using 1 × PCR buffer (10mM Tris-HCl, pH8.3, and 50mM KCl), 1.5mM $MgCl_2$, 0.2mM of each dNTPs, 3 pmol of each primer, 5 units of Dream Taq DNA Polymerase (Thermo Scientific) and 1 µl (~50-100ng) of template DNA. The PCR protocol was used as follows: initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 95 °C for 45 sec, annealing at 56°C for Cytb gene and 54°Cfor CR for 60 sec and extension at 72 °C for 1 min 30 sec. The final extension was at 72 °C for 10 min. Negative controls were included in each reaction to check the reliability and contamination of the experiment. All PCR amplification was confirmed by electrophoresis on 2% agarose gel stained with Firefly dye and visualized under a UV transillumінator. The positive PCR products were further purified with exonuclease-I and shrimp alkaline phosphatase (Thermo Scientific Inc.) at 37°C for 20 min for the removal of any remaining primer and dNTPs, which was followed by inactivation of enzymes at 85 °C for 15 min. Amplicons were further purified and concentrated via the ethanol precipitation method. The purified fragments were sequenced in both directions in an Applied Biosystems (ABI) Genetic Analyzer 3500 XL (Applied Biosystems) using BigDye version 3.1 Kit.

5.2.4 Microsatellite genotyping

Genotyping was performed using thirteen microsatellites' markers, including Maucas01, Maucas03, Maucas06, Maucas18, Maucas21, Maucas24; TWT113, TWS190; GP19; GmuD16, GmuB08; Test 21; msEo41 (Vamberger et al., 2001; Perez et al., 2006; Schwartz et al., 2003; Forlani et al., 2005; Rosenbaum et al., 2007). Out of these, three 'Maucas21, Maucas24 and GmuD16' did not amplify consistently in all samples; hence, excluded from the analysis. Therefore, final analysis was performed with ten microsatellite markers. PCR was carried out in 10µl PCR reaction volumes, containing 5µl of QIAGEN Multiplex PCR Buffer Mix (QIAGEN Inc.), 0.4µM primer mix, and 80–100ng of the template DNA. The PCR cycle was performed under the following conditions, initial denaturation at 95°C for 15 min, followed by 8 cycles of denaturation at 95 °C for 45 sec, annealing at 58°C for 60 sec and extension at 72 °C for 75 sec, again followed by 10 cycles of denaturation at 95 °C for 45 sec, annealing at 58 to 50°C for 60 sec and extension at 72 °C for 75 sec which followed by 12 cycles of denaturation at 95 °C for 45 sec, annealing at 52°C for 60 sec and extension at 72 °C for 75 sec with final extension of 60°C for 30 min. Allelic sizes were scored against the size standard GS500 LIZ. We visualized amplicons using an ABI 3500 XL automated sequencer. Genotypes were identified using GeneMarker version 2.7.4 (Applied Biosystems, Foster City, California, USA).

5.3 Data Analysis

5.3.1 Mitochondrial DNA diversity

The forward and reverse sequences were aligned and edited using SEQUENCHER version 4.9 (Gene Codes Corporation, AnnArbor, MI, USA). All the sequences were aligned using CLUSTAL W (Thompson et al., 1994), and checked manually in BioEdit v 7.2.5 software (Hall et al., 1999). DnaSP 5.0 (Librado and Rozas, 2009) was used to analyze the haplotype diversity (h), nucleotide diversity(p), polymorphic sites (s) and mean number of pairwise differences among sequences (K). MEGA X (Tamura et al., 2011) was used to calculate mean within and between population genetic distances (uncorrected p-distances).

5.3.2 Population demography

The spatial distribution of the haplotypes was visualized through a median-joining network, which was created using the PopART software (Leigh & Bryant, 2015). To determine whether the *B. dhongoka* populations carried a signal of demographic expansion or astationary population history, Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997) neutrality tests were performed in Arlequin ver 3.5 (Excoffier and Lischer, 2010). To generate the trends in spatial demography history, mismatch analysis was carried out using the population growth-decline model in DnaSP, and to evaluate fit of the observed distribution, sum of squared deviations (SSD), and raggedness index(r) under the growth-decline model for each population were used in Arlequin ver 3.5. The P-values were obtained from 1000 simulations based on a selective neutrality test. Best-fit nucleotide substitution model was selected based on Akaike information criterion (AIC) and Bayesian information criterion (BIC) using jModel Test v2.1.7 (Darriba et al., 2012). Bayesian skyline analyses available in BEAST v1.7 (Drummond et al., 2012) were used to infer changes in effective population size (Ne) across time, enabling inference of past demographic changes from the current patterns of genetic diversity within a population (Drummond et al., 2005). We performed multiple analyses that were run for 108 iterations with a burn-in of 107 under the TN93+G model. Because no fossil data were available to calibrate the mutation rate, we assume a conventional molecular clock for the turtle mitochondrial gene (1.75 × 10−8 substitutions/ site/year) (Formia et al., 2006; Naro-Maciel et al., 2014). The mean mutation rate was also consistent with the other turtle species mutation rate of 1.2–2.4% / site per Myrs (Encalada, 1996). Length of MCMC chain was 100 million and sampling every 10,000 generations. Results of replicate runs were combined using LogCombiner v1.7.5 with a burn-in of 20 million iterations (20%) for each run. Tracer v1.7.1 (Rambaut et al., 2018) was used to assess the convergence of chains and to reconstruct Bayesian Skyline Plot.

5.3.3 Microsatellite analysis

Microsatellite data was analyzed using the program Micro-Checker 2.2 (Van Oosterhout et al., 2004) to identify genotyping errors, such as null alleles, large allele dropout, or errors in scoring. The number of effective alleles (Ne), alleles per locus (Na), observed heterozygosity (Ho), and expected heterozygosity (HE) were estimated using the GenAlEx v6.5 program (Peakall et al., 2006). The polymorphic information content (PIC) values were calculated using CERVUS v 3.0.7 (Kalinowski et al., 2007). The supporting format of the data input files was prepared using CONVERTv1.31 (Glaubitz, 2004). Allelic richness (Ar) and Inbreeding (FIS) were calculated in Fstat 2.9.3.2 for each population, and its significance was tested, assuming no Hardy-Weinberg equilibrium within the samples and using 1,000 permutations. All the loci were checked for under HWE in GenAlEx v6. To determine the level of distortion from independent segregation of loci, Linkage disequilibrium (LD) was tested using GENEPOP on the web (genepop.curtin.edu.au).

5.3.4 Genetic structure, differentiation and migration

The pairwise FST values (gene flow) among the populations were calculated using GenAlEx v6. Initially, a factorial correspondence analysis (FCA) was performed using GENETIX ver. 4.02 (Belkhir, 2004), which graphically represents the genetic distances between individual multilocus genotypes. Finally, we tested the population genetic structure using the Bayesian method using Structure version 2.3.3 (Pritchard et al., 2000). We assumed an admixture model with correlated allele frequencies with a burn-in period of 60,000 and 6,00,000 MCMC repetitions. Twenty independent replicates were run for K = 1–10, without prior knowledge of sampling locations. We determined the optimal value of K according to the ΔK method (Evanno et al., 2005) using web server of Structure Harvester (Earl, 2012). The clustering results of Structure were visualized using a web server ClumpaK (http://clumpak.tau.ac.il/index.html) that provides a full pipeline for clustering, summarizing, and visualizing Structure results. We used BayesAss v.1.3 program to estimate recent migration rates, m (over the past few generations) among populations based on multilocus genotypes using Markov chain Monte Carlo (Wilson, 2003). We used burn-in iterations 106 followed by 107 iterations and a sampling frequency of 2000. The initial run was performed with the default delta (Δ) value of 0.15 for allele frequencies (A), migration (M), and inbreeding coefficient (F). The final input parameter of ΔM was adjusted at 0.2. As recommended in the manual the changes in these parameters would be accepted between 40 and 60% of the total number of iterations. Four independent runs were also performed to validate the consistency of the results.

5.3.5 Population history: Bottlenecks

We used Garza–Williamson (G-W) index (Garza and Williamson, 2001) and BOTTLENECK v.1.2.0263 programs to determine the signal of past bottlenecks. We estimated G-W index using ARLEQUIN v3.5, which is a mean ratio of the number of observed alleles in a sample divided by the number of alleles expected under the observed size range and can detect population bottlenecks from the past. The value of M ratio below the critical Mc value of 0.68 was considered a sign of a genetic bottleneck. Second, BOTTLENECK program was run under two mutation models: two-phase model (TPM) and the stepwise mutation model (SMM). The proportion of alleles attributed to SMM under the TPM was set at 90%, with a variance of 12. Significance was determined by the sign and Wilcoxon tests with 10,000 replications (Cornuet, 1996). We also checked allele frequency distributions for mode shifts that discriminate populations that have experienced a recent bottleneck from stable populations (Luikart, 1998).

5.4 Results

5.4.1 Genetic diversity

Genetic diversity parameters revealed with mitochondria DNA and microsatellite markers in the *B. dhongoka* from TRRC and GARRC are presented in Table 14 and Table 15. In this study, two mtDNA fragments were obtained from 92 samples; after alignment, 1140bp of cyt *b* and 451bp of the control region were generated, producing a concatenated mtDNA fragment of 1591bp. All generated sequences were deposited in GenBank (OQ378390 – OQ378573). In overall samples, 15 variable sites were observed, thus accounting for only 0.94% of the total sites. All 15 were parsimony-informative sites, comprising 13 transitions and two transversions. We identified 31 haplotypes with an average of 4.59 nucleotide differences (K) in both TRRC and GARRC populations. Among 31 haplotypes, 28 were detected in TRRC with K=4.55 and six were found in GARRC with K=3.39, while four haplotypes (H1-H4) are common in both populations. The haplotypes (h) and nucleotide diversity (π) of TRRC were 0.932 ± 0.013 and 0.002 ± 0.00012, respectively and those of GARRC were 0.845 ± 0.074 and 0.002 ± 0.0002, respectively. The overall Hd and π were 0.936 ± 0.012 and 0.0022 ± 0.00014, respectively (Table 14).

In addition, the selected microsatellite markers showed polymorphic information content with a mean value of 0.476. Loci Maucas18 had the highest allele counts, while Maucas3 showed the lowest. The analysis of microsatellite data did not find evidence of LD. We did not find evidence of a large allele dropout and scoring error in our data. The mean number of alleles (Na) in TRRC and GARRC were 5 and 3.4, respectively, with comparative allelic richness (Ar) in TRRC (Ar = 3.84) and GARRC (Ar = 3.4) with mean Ar = 3.82 (Table 15). The observed heterozygosity (Ho) and expected heterozygosity (HE) in TRRC were Ho: 0.487; HE: 0.493; and in GARRC were Ho: 0.497; HE: 0.50 with mean Ho: 0.495; HE: 0.495. The mean inbreeding coefficient (FIS) value for both the rescued *B. dhongoka* populations was greater than zero mean FIS: 0.028, indicating a heterozygote deficiency (Table 15), which may be attributed to the Wahlund effect and population not being in HWE.

Table 14: Summary of genetic diversity in *Batagur dhongoka* populations based on concatenated mtDNA (Cyt *b* + CR) and ten microsatellites

Population/Region	Mitochondrial DNA						
	n	H	Hd	π	K	Tajima's D*	Fu's F_S *
Turtle Rescue & Rehabilitation Centre (TRRC) Sarnath, Uttar Pradesh, India	80	29	0.93	0.003	4.6	2.04	-12.3
Ganga Aqualife Rescue & Rehabilitation Centre (GARRC) Narora, , Uttar Pradesh, India	12	6	0.85	0.002	3.4	1.816	0.015
Total	**92**	**31**	**0.94**	**0.002**	**4.6**	**1.56**	**-1.86**

n number of samples, H haplotype, Hd haplotype diversity, π nucleotide diversity

Table 15: Summary of genetic diversity in *Batagur dhongoka* populations based on ten microsatellites

Population/Region	Na	Ar	Ho	H_E	FIS
Turtle Rescue & Rehabilitation Centre (TRRC) Sarnath, Uttar Pradesh, India	5	3.84	0.487	0.493	0.018
Ganga Aqualife Rescue & Rehabilitation Centre (GARRC) Narora, , Uttar Pradesh, India	3.4	3.4	0.497	0.5	0.038
Total	**4.1**	**3.82**	**0.495**	**0.494**	**0.028**

Na number of alleles, Ar allelic richness, HO observed heterozygosity, HE expected heterozygosity, FIS inbreeding coefficient, *all P-values > 0.01 (not significant)

5.4.2 Genetic structure, differentiation and migration

The median-joining (MJ) network of 31 haplotypes revealed no major branching events among the *B. dhongoka* (Figure 23). The haplotypes genealogy showed four shared haplotypes (H1- H4), which consists of 38 individuals, representing 41% of total studies populations, out of which H3 and H4 represented the core haplotypes. In TRRC, the majority of the haplotype contained a single sequence. We performed the FCA (Figure 24) and Structure analysis (Figure 25) of individual microsatellite genotypes. The sampling plot in the FCA grouped individuals into two suggestive clusters: two groups present in TRRC, whereas GARRC showed a single cluster, and the Bayesian model-based clustering analysis also supported this result. The Bayesian clustering analysis under the admixture model implemented in Structure suggested the existence of two genetic clusters based on high ΔK (mean likelihood of K (Mean LnP(K) = −1787.306). The two identified genetic clusters showed some level of genetic admixture and were not assigned to their populations (cluster with an estimated membership

>0.800). The two genetic signatures observed in TRRC indicates that the individuals of TRRC population shared a similar genetic ancestry with the GARRC population.

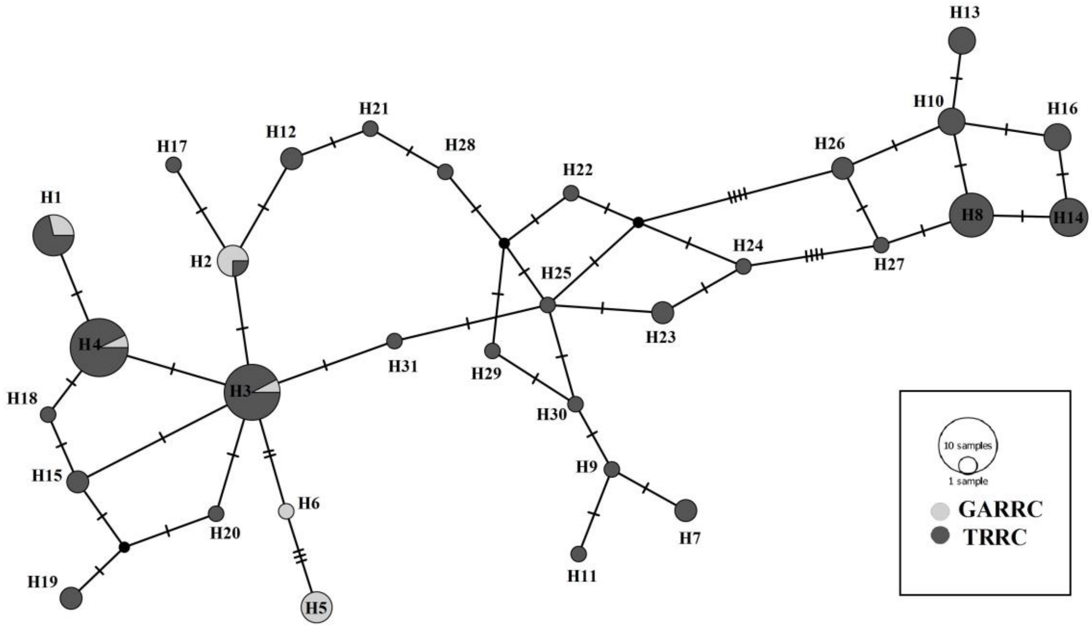


Figure 23: Median-joining network inferred from *Batagur dhongoka* mitochondrial DNA haplotypes. The size of each circle indicates the relative frequency of the corresponding haplotype in the whole dataset. Short tick lines between haplotypes show the number of mutations.

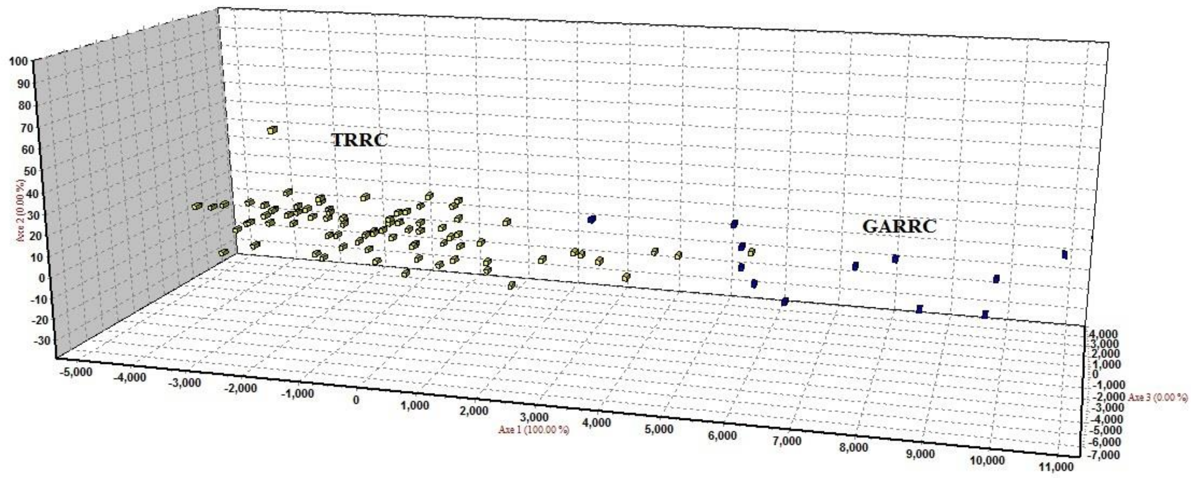


Figure 24: Factorial correspondence analysis (FCA), computed using GENETIX 4.02, shows relationships among the multilocus genotypes of two *Batagur dhongoka* populations (TRRC: Turtle Rescue and Rehabilitation Centre, Sarnath; GARRC: Ganga Aqualife Rescue and Rehabilitation Centre, Narora). Axis 1, Axis 2 and Axis 3 are the first, second and third principal factors of variability.

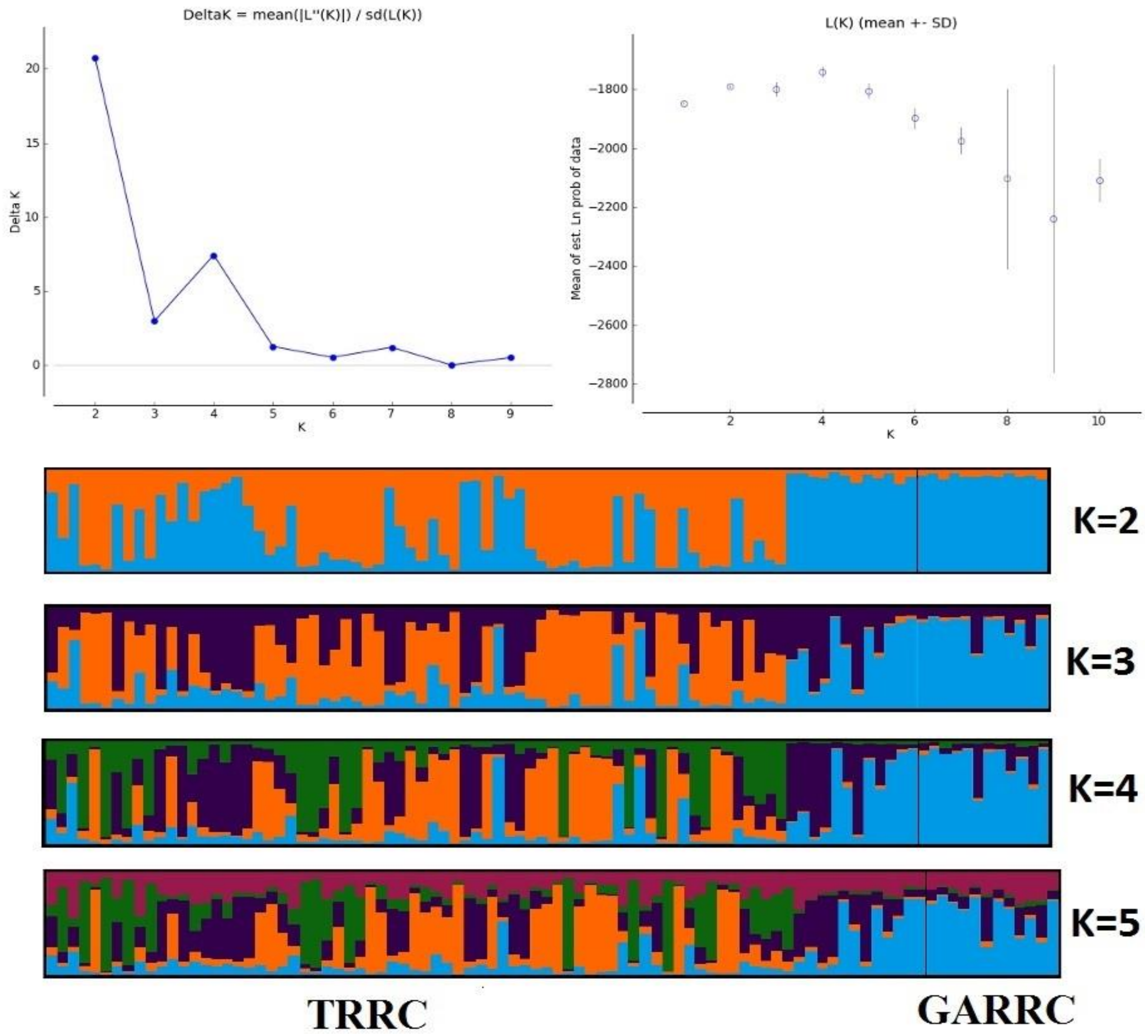


Figure 25: CLUMPK genetic structure plot derived from Bayesian-based cluster analysis in STRUCTURE. The proportion of color in each bar represents the assignment probability of an individual, corresponding to different groups of TRRC: Turtle Rescue and Rehabilitation Centre, Sarnath and GARRC: Ganga Aqualife Rescue and Rehabilitation Centre, Narora

The genetic differentiation level between the studied populations of *B. dhongoka* was lower in mtDNA: F_{ST} = 0.003, whereas a comparatively higher value was observed based on microsatellite: F_{ST} = 0.074. Furthermore, the rate of self-distribution within the GARRC population was 0.9649 (95% CI: 0.8730-0.999), while the TRRC population was 0.9413 (95% CI: 0.8913- 0.9808). The analysis for migration rate (m) was low 0.15 from GARRC to TRRC and 0.035 from TRRC to GARRC.

5.4.3 Population demography

The demographic pattern of studied *B. dhongoka* was inferred from neutrality and mismatch distributions analysis. The mismatch distribution analysis showed a multimodal distribution shaped graph (Figure 26), which might result from a relatively stable population size over the long period.

Results of Tajima's D and Fu's Fs tests were used to explain the population history of both populations (Table 14). In overall samples, the value Tajima's D and Fu's F_S was 1.56 and -1.864, respectively, but non-significant indicated stable population size. Moreover, effective population sizes and demographic trends were also estimated by the Bayesian skyline plot (BSP) analysis. The Bayesian Skyline Plot indicated historical stability followed by growth in the effective population size, with a recent reduction since 2 thousand years (Figure 27).

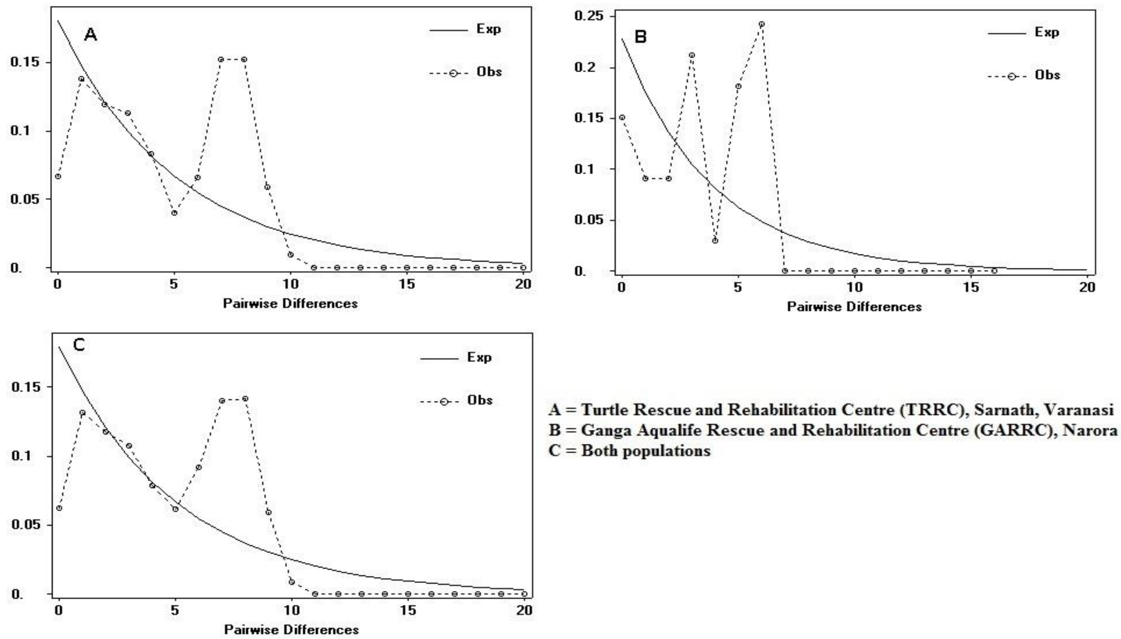


Figure 26: The mismatch distributions graph dashed line showing observed distribution; solid line showing the theoretical expected distribution under a growth-decline model.

5.4.4 Bottleneck

The values of G-W index test (TRRC = 0.31±0.07 and GARRC = 0.32 ±0.11) were lower than the critical value MC = 0.68, indicating the sign of genetic bottleneck in both the population of *B. dhongoka*. In contrast, BOTTLENECK analysis showed that no significant heterozygosity excess was found in both populations under either TPM (TRRC P = 0.72; GARRC P = 0.90) or SMM (TRRC P = 0.21; GARRC P = 0.24). The mode-shift test demonstrated no distortion of the allelic frequency, and a normal L-shaped distribution was observed.

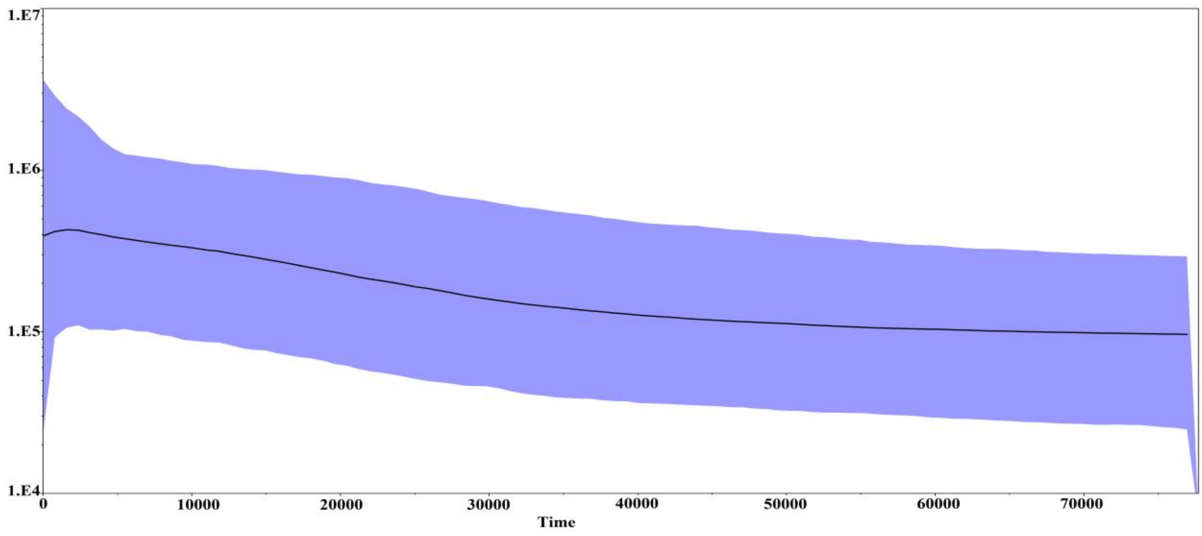


Figure 27: Bayesian skyline plots representing historical demographic trends of *Batagur dhongoka*. X-axis shows time before present in million years ago (MYA). Y-axis indicates effective population size on a logarithmic scale. The thick solid lines represent mean estimates for effective population size, and differentially colored areas reflect the 95% highest posterior density intervals.

5.5 Discussion

To date, genetic studies of *B. dhongoka* have been limited. Our findings reveal new genetic insight about *B. dhongoka* and thus play an important role in freshwater turtle conservation. This study uses mitochondrial and microsatellite markers to investigate the pattern of genetic diversity, population structure and demography of critically endangered *B. dhongoka* from two turtle rescue and rehabilitation centers, TRRC and GARRC. The analysis of mtDNA revealed high haplotype diversity in concurrence with low nucleotide diversity, indicating that haplotypes were closely related. The possible explanation for turtles' low nucleotide diversity is their slow mitochondrial evolutionary rate (King, 2004). It was also evident from the haplotype network, which shows mostly single nucleotide differences between haplotypes. Similarly, high haplotype diversity have been reported in Loggerhead sea turtle (Baltazar-Soares et al., 2020), whereas lower haplotype diversity in turtle species, e.g., in Leatherback turtles, *Dermochelys coriacea* (Wongfu, 2002) and Bog turtle, *Glyptemys muhlenbergii* (Rosenbaum et al., 2007) Our data show that low nucleotide diversity with high haplotype diversity, which can be a signature of population expansion from a small effective population size (Rogers et al., 1992). The clustering of sequences showed that four haplotypes of TRRC was similar to those found in GARRC, indicating close genetic relationships and having common genetic ancestry. Furthermore, microsatellite analysis revealed a moderate level of genetic diversity for both *B. dhongoka* populations. Microsatellite based structure analysis revealed weak genetic clusters between the TRRC and GARRC populations. While the optimal number of clusters for both populations was determined to be two, we extended our analysis to include K values ranging from 2 to 5. Notably, at K higher than three, the TRRC population showed the presence of two genetic groups, while the GARRC population maintained its unity as a single cluster. The genetic clustering observed in TRRC is probably due to a substantial portion of individuals sharing a similar genetic ancestry with the GARRC population. The results support the ongoing conservation practices under the National Mission for Clean Ganga Project, where most of *B. dhongoka* individuals in TRRC belong to the source population of the Chambal River and few rescued individuals from seizures near the Ganga river. Hence, two different stocks are observed in TRRC. In contrast, individuals in GARRC are only the rescue cases recovered from the local areas.

The BayesAss analysis indicated recent migration from GARRC to TRRC (m=0.15), which may be attributed to human interference and the mixing of individuals in TRRC (Ganga and Chambal rivers). The rate of self-assemblage was higher in both GARRC and TRRC, which may be due to limited gene flow and low population size. In the current study, our sample size from GARRC is limited. Therefore, it is important to exercise caution when interpreting the conclusions drawn from the BayesAss analysis. Incorporating a larger number of samples along with additional markers has the potential to yield significantly different estimates of migration (Meirmans, 2014). Moreover, the results of migration rates were consistent with pairwise F_{ST} values obtained from microsatellite and mtDNA markers, indicating weak genetic differentiation between both the *B. dhongoka* populations. Although *B. dhongoka* has experienced a severe reduction in population size due to anthropogenic disturbance, the mtDNA-based demographic history analyses ruled out any significant past contraction in both the studied populations. It was also in accordance with the estimates of the BOTTLENECK analysis with respect to recent bottlenecks.

In contrast, the G-W index test did not give a consistent result with bottlenecks and mtDNA analysis, suggesting a genetic bottleneck in both populations. It may be accounted for by the G-W index's sensitivity in detecting population bottlenecks, in which the number of alleles is typically more reduced than the range due to a decline in population size. As indicated by the Bayesian Skyline Plot, there was a period of historical stability followed by an expansion in the effective population size, but more recently, a slight decline has been observed from ~2 thousand years ago. The rise of anthropogenic activities in the Holocene is one of the major factors influencing global species biodiversity. Additionally, the population of *B. dhongoka* has experienced substantial declines due to entanglement in fishing nets, the construction of major hydrological projects affecting river flow dynamics and nesting beaches, water pollution, as well as the detrimental effects of illegal trade. We believe these are the most important factors in reducing population sizes and shaping the genetic structure of *B. dhongoka*. Our analysis of microsatellite markers indicates that *B. dhongoka* populations have moderate genetic diversity, while mtDNA analysis suggests a low level of nucleotide diversity. The primary factors contributing to reduced genetic diversity are small population sizes, which are intricately connected to the fitness and health of individuals, along with limited population growth capacity (Spielman et al., 2007). Therefore, robust rehabilitation and relocation strategies are needed to support small and declining populations, which include scientifically based releases into natural environments. Finding reliable origin information is often challenging, and releasing rescued individuals of unknown sources into the wild is sometimes associated with the mortality of individuals (Villemey et al., 2013). Additionally, while reintroduction starts with a small population, genetic monitoring is essential to comprehend current genetic diversity and relatedness for management and conservation considerations. As a result of the recent trend of population decline amidst persistent pressures, long-term evidence-based conservation, and management interventions are required for the long-term survival of *B. dhongoka*.

5.6 Conclusion

This study presented a pioneer genetic analysis of *B. dhongoka*. These results expand our genetic knowledge of the Critically Endangered *B. dhongoka* and are significant to support and develop management strategies to manage its genetic diversity in the future. The conservation management of *B. dhongoka* should focus on protecting its core habitats and breeding grounds while tackling illegal poaching and hunting. Its habitat should be monitored and maintained continuously without disturbance and reconnected across its range to maintain a relatively large and healthy effective population size and augment its chances for successful survival and adaptation. If implemented, these measures will conserve not only the *B. dhongoka* but also many threatened sympatric turtle species throughout the Ganga River basin.

5.7 References

- Baltazar-Soares, M., Klein, J.D., Correia, S.M. et al. Distribution of genetic diversity reveals colonization patterns and philopatry of the loggerhead sea turtles across geographic scales. *Sci Rep* **10**, 18001 (2020).
- Belkhir, K. GENETIX 4.05, logiciel sous Windows TM pour la genetique des populations. http://www.genetix.univ-montp2.fr/genetix/genetix.htm.2004.
- Burns, C.E., Ciofi, C., Beheregaray, L.B., Fritts, T.H., Gibbs, J.P., Márquez, C., Milinkovitch, M.C., Powell, J.R. and Caccone, A. The origin of captive Galapagos tortoises based on DNA analysis: implication for the management of natural populations. Anim. Conserv., 6, 329–337 (2003).
- Choudhury, B.C., Bhupathy, S. and Hanfee, F. Status information on the tortoises and freshwater turtles of India. In Asian Turtle Trade: Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia. *Chelonian Res. Monogr.* 2, 86-94 (2000).
- Çilingir, F.G., Rheindt, F.E., Garg, K.M., Platt, K., Platt, S.G., Bickford, D.P. Conservation genomics of the endangered Burmese roofed turtle. *Conserv Biol.* **31**, 1469-1476 (2017).
- Çilingir, F.G., Seah, A., Horne, B.D., Som, S., Bickford, D.P., Rheindt, F.E. Last exit before the brink: Conservation genomics of the Cambodian population of the critically endangered southern river terrapin. *Ecol Evol*. **9**, 9500-9510 (2019).
- Cornuet, J.M., Luikart, G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genet*. **144**, 2001–2014 (1996).
- Darriba, D., Taboada, G.L., Ramon, D. and Posada, D. JModelTest 2: more models, new heuristics and high-performance computing. *Nat Methods* **30**, 772 (2012).
- Das, I., Choudhury, B.C., Praschag, P., Ahmed, M.F. & Singh, S. *Batagur dhongoka* (errata version published in 2019). The IUCN Red List of Threatened Species 2019: e.T10953A152042542.http://dx.doi.org/10.2305/IUCN.UK.2019-1.RLTS.T10953A152042542.en (2019).
- Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol. Evol.* **29**, 1969-1973 (2012).
- Drummond, A.J., Rambaut, A., Shapiro, B.E.T.H. and Pybus, O.G. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol. Evol*, **22**, 1185-1192 (2005).
- Earl, D.A. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conser. Genet. Resour*. **4**, 359–61 (2012).
- Encalada, S.E., Lahanas, P.N., Bjorndal, K.A., Bolten, A.B., Miyamoto, M.M. and Bowen, B.W., Phylogeography and population structure of the Atlantic and Mediterranean green turtle *Chelonia mydas*: a mitochondrial DNA control region sequence assessment. *Mol. Ecol*., 5, 473–483(1996).
- Evanno, G., Regnaut, S., Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*.**14**, 2611–20 (2005).
- Excoffier, L., Estoup, A., Cornuet, J.M. Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. Genet. **169**, 1727–1738 (2005).
- Excoffier, L., Lischer, H.E.L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
- Forlani, A., Crestanello, B., Mantovani, S., Livoreil, B., Zane, L., Bertorelle, G. & Congiu, L. Identification and characterization of microsatellite markers in Hermann's tortoise (Testudo hermanni, Testudinidae). *Mol. Ecol. Notes*, **5**, 228-230 (2005).

- Formia, A., Godley, B.J., Dontaine, J.F., Bruford, M.W. Mitochondrial DNA diversity and phylogeography of endangered green turtle (*Cheloniamydas*) populations in Africa. *Conserv Genet*. **7**, 353– 369 (2006).
- Fu, Y. X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genet.* **147**, 915-925 (1997).
- Garza, J.C. Williamson, E.G. Detection of reduction in population size using data from microsatellite loci. *Mol Ecol*. **10**, 305–318 (2001).
- Gaur, A., Reddy, A., Annapoorni, S., Satyarebala, B. and Shivaji, S. The origin of Indian Star tortoises (*Geochelone elegans*) based on nuclear and mitochondrial DNA analysis: A story of rescue and repatriation. *Conser. Genet*. **7**, 231-240 (2006).
- Ge, C., Cui, Y.N., Jing, P.Y. and Hong, X.Y. An alternative suite of universal primers for genotyping in multiplex PCR. PLoS One **9**, p.e92826 (2014).
- Glaubitz, J.C. Convert: a user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Mol. Ecol. Notes*. **4**,309–10 (2004).
- Goudet, J.F. FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Heredity* **86**:485–6 (1995).
- Hall, T.A. BioEdit software, version 5.0. 9. North Carolina State University, Raleigh, NC (1999).
- IUCN TFTSG and ATTWG (IUCN/SSC Tortoise and Freshwater Turtle Specialist Group and Asian Turtle Trade Working Group). 2000. Recommended changes to 1996 IUCN Red List Status of Asian turtle species. In: P.P. van Dijk, B.L. Stuart and A.G.J. Rhodin (eds), Asian Turtle Trade: Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia. *Chelonian Res. Monogr.* 2, 156-164. Chelonian Research Foundation, Lunenburg, MA, USA.
- Kalinowski, S.T., Taper, M.L., Marshall, T.C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol*.**16**,1099–106 (2007).
- King, T.L. and Julian, S.E., 2004. Conservation of microsatellite DNA flanking sequence across 13 Emydid genera assayed with novel bog turtle (*Glyptemys muhlenbergii*) loci. *Conserv. Genet*. **5**, 719-725 (2004).
- Kumar, A., Yadav, P., Usmani, A., Hussain, S.A. and Gupta, S.K. Comparative Mitogenomics of Two Critically Endangered Turtles, *Batagur kachuga* and *Batagur dhongoka* (Testudines: Geoemydidae): Implications in Phylogenetics of Freshwater Turtles, Research square (2021).
- Kumar, M. In Varanasi, turtles help clean the Ganga (2023, July 08). The Hindu: https://www.thehindu.com/sci-tech/energy-and-environment/a-thousand-turtles-to-be-released-for-a-cleaner-ganga-at-varanasi/article67057945.ece
- Leigh, J. W., & Bryant, D. Popart: full-feature software for haplotype network construction. *Methods Ecol. Evol.* ***6***, 1110-1116 (2015).
- Lenk, P., Fritz, U., Joger, U., & Wink, M. Mitochondrial phylogeography of the European pond turtle, *Emys orbicularis* (Linnaeus 1758). *Mol. Ecol*. **8**, 1911-1922 (1999).
- Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**,1451–1452. (2009).
- Luikart, G., Allendorf, F.W., Cornuet, J.M., Sherwin, W.B. Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered. **89**, 238–247 (1998)
- McCleery, R., Hostetler, J.A. and Oli, M.K. Better off in the wild? Evaluating a captive breeding and release program for the recovery of an endangered rodent. *Biol. Conserv.* 169, 198-205 (2014).
- Meirmans, P. G.. Nonconvergence in Bayesian estimation of migration rates. *Molecular ecology resources*, **14**, 726-733 (2014).

- Naro-Maciel, E., Reid, B.N., Alter, S.E., Amato, G., Bjorndal, K.A., Bolten, A.B., Martin, M., Nairn, C.J., Shamblin, B., Pineda-Catalan, O. From refugia to rookeries: phylogeography of Atlantic green turtles. *J. Exp. Mar. Biol. Ecol.* **461**,306–316 (2014).
- Peakall, R.O., Smouse, P.E. GENALEX 6: genetic analysis in Excel Population genetic software for teaching and research. *Mol. Ecol. Notes* 6, 288–95 (2006).
- Pedall, I., Schäfer, H., Fritz, U. and Wink, M. Isolation of microsatellite markers in the *Emys orbicularis* complex and development of multiplex PCR amplification. *Conserv. Genet.* **10**, 725-727 (2009).
- Peery, M.Z, Kirby. R., Reid, B.N., Stoelting, R., Doucet-Be¨er, E.L., Robinson, S et al. Reliability of genetic bottleneck tests for detecting recent population declines. *Mol Ecol*. **21**, 3403–3418 (2012).
- Perez, M., Bour, R., Lambourdiere, J., Samadi, S. and Boisselier, M.C. Isolation and characterization of eight microsatellite loci for the study of gene flow between Testudo marginata and Testudo weissingeri (Testudines: Testudinidae). *Mol. Ecol. Notes*, **6**, 1096-1098 (2006).
- Praschag, P., Ahmed, M.F., Das, I. & Singh, S. *Batagur kachuga* (errata version published in 2019). The IUCN Red List of Threatened Species 2019: e.T10949A152043133. http://dx.doi.org/10.2305/IUCN.UK.2019-1.RLTS.T10949A152043133.en (2019).
- Pritchard, J.K., Stephens, M., Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics*155, 945–959 (2000).
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G. and Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, **67**, 901-904.
- Rashid, S.M.A., and Khan, S.M.M.H. Trade and Conservation Status of Freshwater Turtles and Tortoises in Bangladesh. In: P.P. van Dijk, B.L. Stuart and A.G.J. Rhodin (eds), Asian Turtle Trade: Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia. *Chelonian Res. Monogr*. 2, 77-85. Chelonian Research Foundation, Lunenburg, MA, USA (2000).
- Reed, D.H. and Frankham, R. Correlation between fitness and genetic diversity. *Conserv. Biol*. **17**, 230-237 (2003).
- Rogers, A.R., and Harpending, H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.***9**, 552-569 (1992).
- Rosenbaum, P.A., Robertson, J.M. and Zamudio, K.R., Unexpectedly low genetic divergences among populations of the threatened bog turtle (*Glyptemys muhlenbergii*). *Conserv.Genet*. **8**, 331-342 (2007).
- Salleh, M.H.M., Esa, Y., Pau, S.S.N. Conservation Genetics of the Critically Endangered Southern River Terrapin (*Batagur affinis*) in Malaysia: Genetic Diversity and Novel Subspecies Distribution Ranges. *Biology* **12**, 520 (2023).
- Schlötterer, C. and Pemberton, J. The use of microsatellites for genetic analysis of natural populations. In Molecular ecology and evolution: Approaches and applications, pp. 203-214. Birkhäuser, Basel (1994).
- Schwartz, T.S., Osentoski, M., Lamb, T. and Karl, S.A. Microsatellite loci for the North American tortoises (genus *Gopherus*) and their applicability to other turtle species. *Mol. Ecol. Notes*, **3**, 283-286 (2003).
- Shaffer, H.B., FitzSimmons, N.N., Georges, A. and Rhodin, A.G. Turtle conservation genetics working group. *Chelonian Res. Monogr*. **4**, 107-123 (2007).
- Spielman, D., Brook, B.W. and Frankham, R. Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences, 101, 15261-15264 (2004).

- Spinks, P. Q., Shaffer, H. B., Iverson, J. B., & McCord, W. P. Phylogenetic hypotheses for the turtle family Geoemydidae. *Mol. Phylogenet. Evol*. **32**, 164-182 (2004).
- Stanford, C.B., Iverson, J.B., Rhodin, A.G., van Dijk, P.P., Mittermeier, R.A., Kuchling, G., Berry, K.H., Bertolero, A., Bjorndal, K.A., Blanck, T.E. and Buhlmann, K.A. 2020. Turtles and tortoises are in trouble. *Curr. Biol.* **30**, 721-735 (2020).
- Tajima H. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genet*. **123**, 585–595 (1989).
- Tamura, K., Stecher, G. and Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol*, **38**, 3022-3027 (2021).
- The Print. World Turtle Day: UP Government Releases Nearly 300Endangered Turtles Into Chambal River (2022). https://theprint.in/india/up-300-turtles-released-into-chambal-on-eve-of-world-turtle-day/966698/
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. **22**, 4673-4680 (1994).
- Vamberger, M., Stuckas, H. and Fritz, U., 2011. Fifteen microsatellite markers for the stripe-necked terrapin *Mauremys caspica* (Testudines: Geoemydidae) and cross-amplification tests in *M. rivulata*. *Conserv. Genet. Res.* **3**, 87-89 (2011).
- Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., and Shipley, P.. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* **4**:535–538 (2004).
- Villemey, A., Besnard, A., Grandadam, J. and Eidenschenck, J. Testing restocking methods for an endangered species: Effects of predator exclusion and vegetation cover on common hamster (*Cricetus cricetus*) survival and reproduction. *Biol. Conserv.* 158, 147-154 (2013).
- Wilson, G.A., Rannala, B. Bayesian inference of recent migration rates using multilocus genotypes. *Genet*. **163**:1177–1191 (2003).
- Wongfu, C., Prasitwiset, W., Poommouang, A., Buddhachat, K., Brown, J.L., Chomdej, S., Kampuansai, J., Kaewmong, P., Kittiwattanawong, K. and Nganvongpanit, K. Genetic Diversity in Leatherback Turtles (*Dermochelys coriacea*) along the Andaman Sea of Thailand. *Diversity* **14**, 764 (2022).

6. GENETIC ASSESSMENT OF BATAGUR BASKA (NORTHERN RIVER TERRAPIN) FOR CONSERVATION PLANNING

6.1 Background of the study

The Sundarbans, one of the largest mangrove forests in the world and also UNESCO World Heritage Site, harbors a rich biodiversity, due to its unique ecosystem. The Northern River Terrapin, also known as *Batagur baska* (Figure 28), is a large iconic riverine species found inestuaries of the Sundarbans region (Figure 29), spanning across Bangladesh and into parts of India (Praschag and Singh, 2019). Historically, the population of *B. baska* extended from Orissa and West Bengal in India, to Bangladesh and Myanmar (lower Ayayarwady, Sittaung, and Thanlwin), and in south of the Kra River on Thailand's Andaman Sea coast (Iverson 1992). However, subsequent research revealed that the populations in southern and eastern regions actually constitute a separate species, known as the southern river terrapin *Batagur affinis* (Praschag et al. 2008). Presently, *B. baska* has been eradicated from most of its former range, and there have been no reports of viable wild populations of *B. baska* in the last 20 years, indicating that the species may have become extinct across much of its range (Mallick et al., 2021). It is now recognized as one of the 25 most critically endangered turtle species globally (Rhodin et al. 2011) and is considered ecologically extinct (Weissenbacher et al. 2015).

Figure 28: Male *Batagur baska,* Sunderban

6.1.1 Habitat, Ecology and Behaviour of Batagur baska

B. baska is primarily found in freshwater and brackish water habitats, including rivers, estuaries, and mangrove swamps. It is most commonly found in the Sundarbans mangrove forest (Figure 29), a UNESCO World Heritage Site shared between India and Bangladesh. Mangroves provide critical shelter and foraging grounds, with their complex root systems offering protection from predators and abundant food sources such as molluscs, crustaceans, and plant matter. The species is known for its unique nesting behavior. Female *B. baska* travel long distances from their feeding grounds to specific nesting sites to lay eggs. The nesting sites are sandy riverbanks or beaches, where the females dig holes and lay their eggs. The nesting season typically occurs during the dry season when water levels are lower. *B. baksa* uses sandy nesting beaches along the sea, frequenting the tidal zones ofestuaries, large rivers and mangroves (Asian TurtleTrade Working Group, 2000). Fishermen's reports indicate that male *B. baska* turtles are commonly captured in estuaries and along the adjacent coastline in the Sundarbans. In contrast, females are often found further upstream, especially during the nesting season (Moll et al., 2009). It is primarily carnivorous, feeding on a variety of aquatic invertebrates, fish, and crustaceans. Their feeding behavior is crucial for maintaining the ecological balance of their habitat.

Figure 29: Mangroves of Sunderban, tiger researve

6.1.2 Threats

Despite its ecological importance, the Northern River Terrapin faces severe threats that have led to a rapid decline in its population. Habitat loss, primarily due to deforestation, conversion of mangrove forests for agricultural purposes, fishing activites (Figure 30) and coastal development projects, has significantly fragmented and degraded its habitat. Furthermore, illegal poaching for the pet trade and consumption has further exacerbated the decline of this endangered species (Praschag and Singh, 2019). The population decline of the Northern River Terrapin in the Sundarbans is alarming and demands urgent conservation attention. Without immediate intervention, this iconic species faces the risk of extinction, which would not only disrupt the delicate ecological balance of the Sundarbans but also lead to the loss of cultural heritage and traditional practices associated with the species among local communities.

Figure 30: Fishing activities in low tide zone in Sunderban

6.1.3 Population status

India

From being abundant in the 19th century, *B. bas*ka population has declined to fewer than 100 mature animals remaining. Less than 50 adult individuals survive in each of the declining natural subpopulations. According to Lucknow Red List Workshop (2005) the estimated population was fewer than 40 individuals, producing about three nests per year with a suggestion to for more intensive surveys to identify more nesting sites. As the record, In August 2008, Turtle surveillance Alliance (TSA) identified and captured 13 *B. baskas* (3 females, 8 males, and 2 juveniles) nearby areas to Sajnekhali Wildlife Sanctuary. Additionally, in the same year, the Madras Crocodile Bank Trust (MCBT) also identified two females that had been rescued from a fish market in the 1970s. So, after this, continuous efforts by TSA and West Bengal Forest department led to the launch the breeding program of *B. baska* in the in the Sajnekhali range of Sunderban Tiger Reserve (Figure 31). These efforts represent critical interventions to conserve this critically endangered species and the number reached to more than 350 by 2019 (Table 16) and recent estimate is 359 individuals in 2021 (Mallick et al., 2021), and it is the largest population of *B. baska* (Figure 32 & Figure 33).

Figure 31: *Batagur baska* pond at Sajnekhali range of Sunderban Tiger Reserve

Figure 32: Hatchery of *B. baska* at Sajnekhali range, Sunderban Tiger Reserve

Bangladesh

In Bangladesh, a long-term monitoring program led to the identification of a few live *Batagur baska* terrapins in fish breeding ponds and markets. Following this discovery, the initial individuals were transferred to the Turtle Island facility in Graz and the Vienna Zoo in Austria, where two juveniles successfully hatched in 2010. Subsequently, a large-scale breeding program, known as 'Project *Batagur*,' was launched in Bangladesh with 14 males and six females in Bhawal National Park. This initiative was a joint effort by the Bangladesh Forest Department, the Turtle Survival Alliance (TSA), the Vienna Zoo, and IUCN Bangladesh (Weissenbacher et al., 2015).

In addition to the conservation station in Bhawal National Park, a backup facility was established six years later in Karamjal, located in the southern Sundarbans, an historic terrapin distribution area. By 2021, there were four clutches at Karamjal and two clutches at Bhawal, resulting in a population of over 400 juveniles.

Table 16: Presence of captive stock from 2008 to 2019 of Northern River Terrapin, *Batagur baska* in Sundarbans.

Year	2008	2012	2013	2014	2016	2017	2019
Hatchlings	0	33	55	57	95	74	50
No of Individuals	12	45	100	157	252	326	376

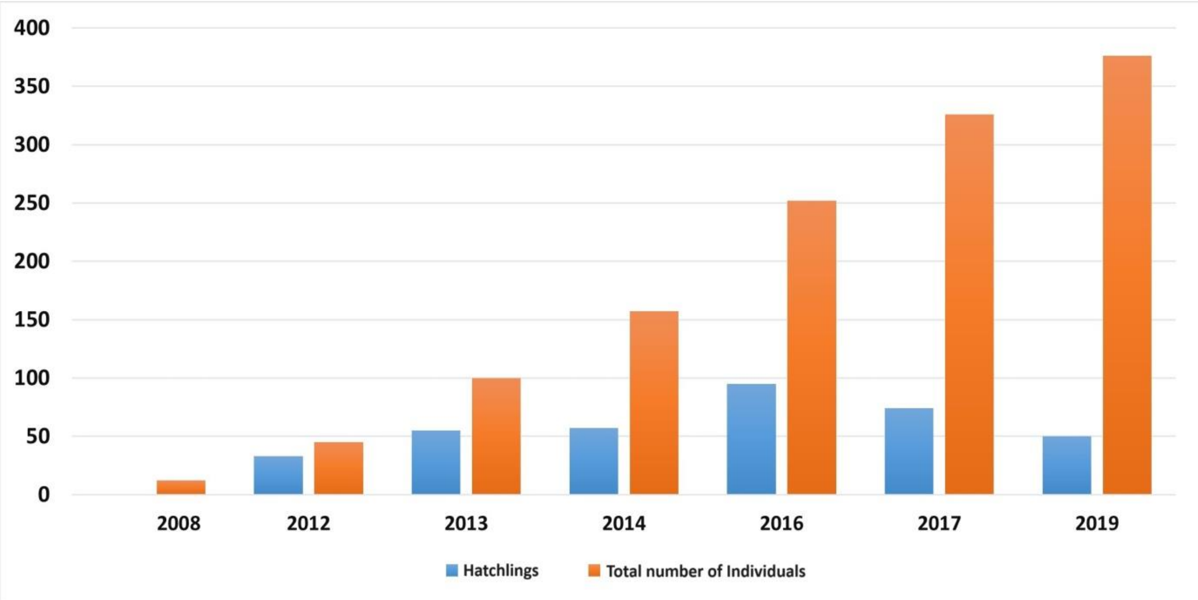


Figure 33: Present stock of hatchlings and total numbers Northern River Terrapin, *Batagur baska* at the breeding centres of Sunderban Tiger Reserves.

Semi adult female Batagur baska in Sundarbans

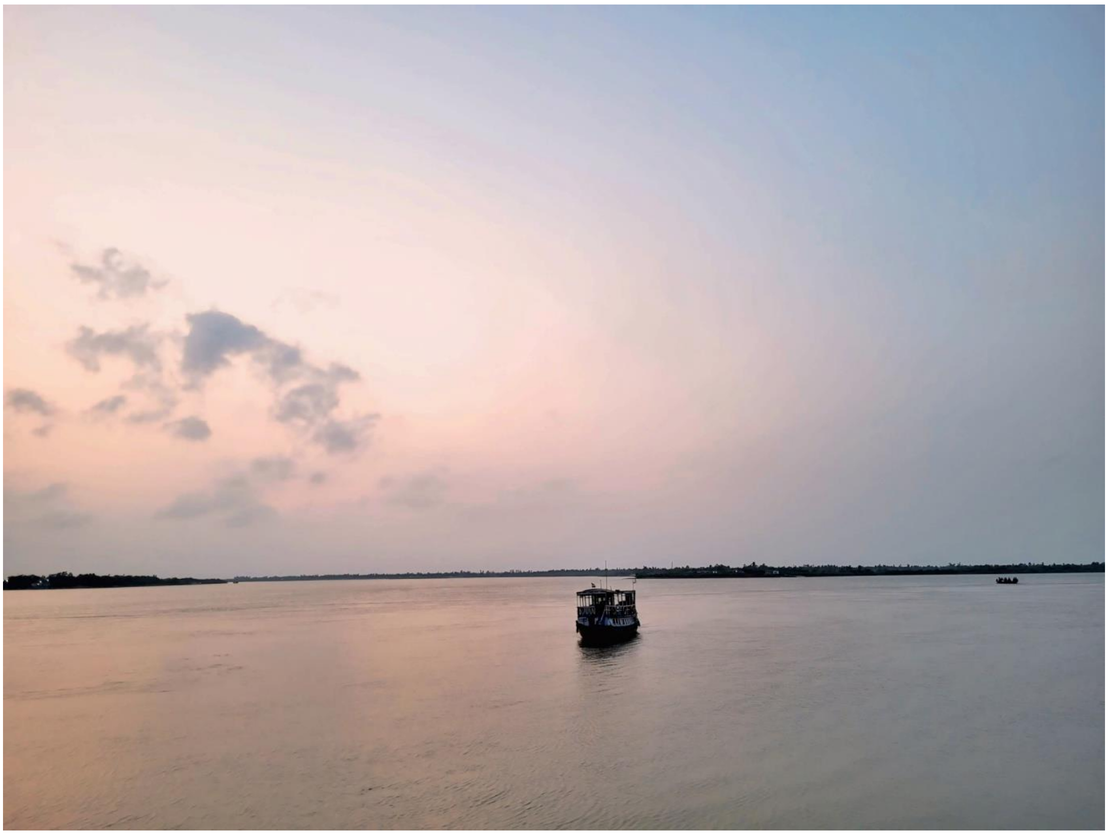


Figure 34: Landscape of Sunderban tiger reserve.

6.1.3 Conservation Actions

Batagur baska, has been listed in CITES Appendix I since 1975, which prohibits all forms of international commercial trade. It is also protected under Schedule I of the Wildlife (Protection) Act, 1972 (amended). In Myanmar, laws controlling egg collection were once in place, but with the near extirpation of the terrapin within the country, these laws are now considered antiquated (Moll 2009). Marine and freshwater turtles are covered under the Myanmar Fishery Law and the Myanmar Protection of Wildlife, Wild Plants, and Conservation of Natural Areas Law, 1994. However, in Myanmar wildlife laws are poorly enforced. In Bangladesh Wildlife Preservation (Amendment) Act (BWPA) 1974, this species is not listed. Efforts to recover the small population of *B. baska* in the Sunderbans have been unintentionally started since 1980's, with ex situ conservation program for the Olive Ridley Turtle, *Lepidochelys olivacea* (Mallick et al., 2021). Nine hatchlings of a different species were found during the 1983 nesting season among the Olive Ridley hatchlings; these were later identified to be *B. baska* (Ghosh & Mandal, 1990). The hatch-and-release program was discontinued in the late 1990s, and the remaining captives were then released in nature (Moll et al. 2009). Considering the international conservation policy and management plan, it is essential to implement extended monitoring efforts, identify sustainable populations, and execute ascientifically guided reintroduction program throughout their native habitat to ensure the long-term conservation of *B. baska*.

In light of these pressing challenges, comprehensive research initiatives are essential to understand the status and dynamics of the *B. baska* population in the Sundarbans. Conservation efforts must focus on mitigating threats to its habitat, combating illegal poaching and trade, and implementing measures to promote recovery and sustainable management. The use of non-invasive genetic sampling techniques, as highlighted earlier, provides a valuable tool for monitoring the population dynamics and genetic health of the *B. baska* without causing further harm to the species or its habitat. Concerted efforts must be made at local, national, and international levels to conserve the *B. baska* and safeguard the ecological integrity of the Sundarbans. Through collaborative research, conservation action, and community engagement, we can work towards ensuring a future where the *B. baska* thrives in its natural habitat, contributing to the resilience and biodiversity of the Sundarbans ecosystem for generations to come.

6.1.4 Molecular Genetics

Molecular genetics has emerged as a powerful tool in Testudine systematics and conservation genetics, enabling a deeper understanding of population genetic structure and evolutionary relationships (Le et al., 2014; Murphy et al., 2013). Genetic diversity is essential for the capacity of a species to adapt to a changingenvironment. Maintaining such diversity is not only critical for the conservationof small populations but is also the basis for the success of any long-term breedingprogramme. Identifying populations that have recently undergone bottlenecks, such as the case with *B. baska*, is paramount due to its implications on genetic diversity. Bottlenecks can escalate inbreeding, diminish genetic variation, and perpetuate harmful allele fixation, thereby compromising a population's adaptability and elevating extinction risks (Luikart and Cornuet, 1998). Yet, detecting recent bottlenecks poses challenges, given the often-elusive historical population sizes and genetic diversity levels. Understanding the genetic makeup of *B. baska* is essential for shaping effective conservation strategies and management plans.

6.1.5 Markers for Genetic study

Mitochondrial DNA Sequencing (mtDNA)

Mitochondrial DNA (mtDNA) stands out as a prominent population genetic marker, due totherapid rate of base substitution, haploid nature, and maternal inheritance. mtDNA selectively provides invaluable insights into the genetic diversity, population structure, and evolutionary history of species. The process of DNA sequencing involves discerning each nucleotide (base) within a specified target region of DNA, commonly known as the genetic marker.

mtDNA control region (CR)

The control region is an essential transcription regulator found in mitochondrial genomes. This region has been reported to have more sequence variation than other regions of mtDNA sequences. Variation in the control region has been used to determine the diversity of many endangered species. Further, mtDNA polymorphisms have remained relevant to studying population structure and intraspecific variation. This utilization of mitochondrial DNA sequencing enhances our understanding of population dynamics and is a powerful tool for species delineation and conservation efforts (Figure 35).

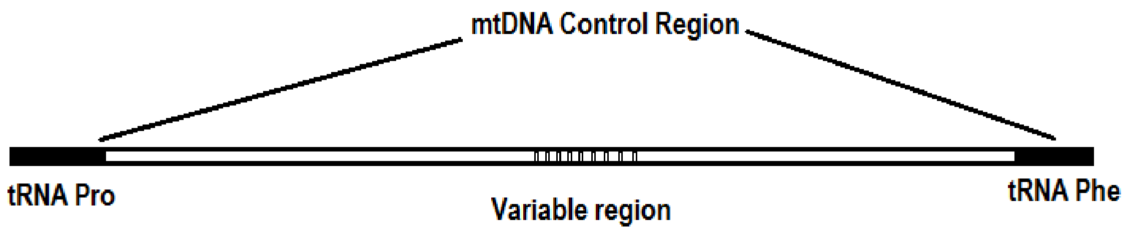


Figure 35: Represenatation of highly varaible portion of control region in mitochondria

6.2 Objective of study

Our study delves into the intricacies of *B. baska*, with a dual focus: firstly, to untangle its phylogenetic and evolutionary connections with other turtle species, and secondly, to probe into its genetic diversity and demographic trends using the mitochondrial marker. Our research pursues two primary objectives: (i) illuminating the extent of genetic diversity and gene flow, and (ii) unraveling historical demographic changes and levels of inbreeding. By addressing these aims, we aim to furnish invaluable insights essential for the conservation and management strategies tailored to safeguard this imperiled turtle species.

6.3 Study Areas

The Sundarbans region in India lies south of the Tropic of Cancer, spanning between 21°32' and 22°40' N latitude and 88°30' N and 89°00' E longitude (Figure 36). Situated in the Bay of Bengal, it forms a delta at the confluence of the Ganga, Brahmaputra, and Meghna rivers. Encompassing approximately 10,000 km^2, with 4,000 km^2 on the Indian side, the Sundarbans feature a complex network of tidal waterways, mudflats, and small islands of salt-tolerant mangroves. The Indian portion includes the Sundarban Tiger Reserve, covering 2,585 km2, with a core area of 1,330 km2 designated as the Sundarbans National Park since 1984. The remaining area serves as a buffer zone, housing the Sajnekhali Wildlife Sanctuary since 1976, spanning 362 km^2. The Sajnekhali Breeding Centre is a vital facility dedicated to the conservation and breeding of endangered species, particularly saltwater crocodiles and turtles. The Dampier-Hodges line marks the boundary separating the Sundarbans from the rest of West Bengal.

6.4 Methodology

6.4.1 Permissions for Biological Sampling

Permit to collect non-invasive biological samples of *B. baska* was obtained from theState Forest Department of West Bengal (Letter Number 2225/WL/2W-797/2017 dated 10/08/2023). Field sampling was carried out as part of an ongoing WII Project on "Planning and Management for Aquatic Species Conservation and Maintenance of Ecosystem Services in the Ganga River Basin for a Clean Ganga" (Figure 37).

6.4.2 Non-Invasive Sampling and DNA Extraction

Non-invasive sampling techniques offer a minimally intrusive method for collecting genetic material from wildlife without causing harm. For instance, skin swabs can extract genetic data for mitochondrial DNA sequencing, microsatellite genotyping, and genomic studies. By employing these techniques, researchers can monitor and conserve critically endangered species like the Northern River Terrapin (*B. baska*) without disturbing their natural habitats. At Sajenakhali, captive individuals are carefully managed by animal keepers who periodically clean the hatcheries. To collect non-invasive samples, we provided gloves to the animal keepers for use while handling these individuals at the hatcheries. Each individual was handled with a single glove, ensuring minimal contamination. In total, 25 gloves were used for collecting mucus layers. Additionally, 15 fresh fecal samples were collected from ponds and hatcheries at Sajnekhali WLS. The gloves used for collection were inverted during removal and filled with 20-30 ml of 70% ethanol. The open end was securely tied with a sterile rubber band to prevent alcohol spillage. After collection, all gloves were vigorously shaken to ensure the mucus layer was completely dissolved in the ethanol solution. The ethanol-containing mucus was then transferred to a 50 ml collection/centrifugation tube and transported to the Wildlife Institute of India (WII) laboratory, where it was stored at -20°C until further analysis. These procedures ensure the integrity of genetic material for subsequent analysis, facilitating research on the *B.baska* population dynamics and genetic diversity.Genomic DNA (gDNA) was extracted from collected samples using the QIAGEN DNeasy

Blood and Tissue Kit with some modifications accordingly and gDNA waschecked by 0.8% agarose gel electrophoresis.

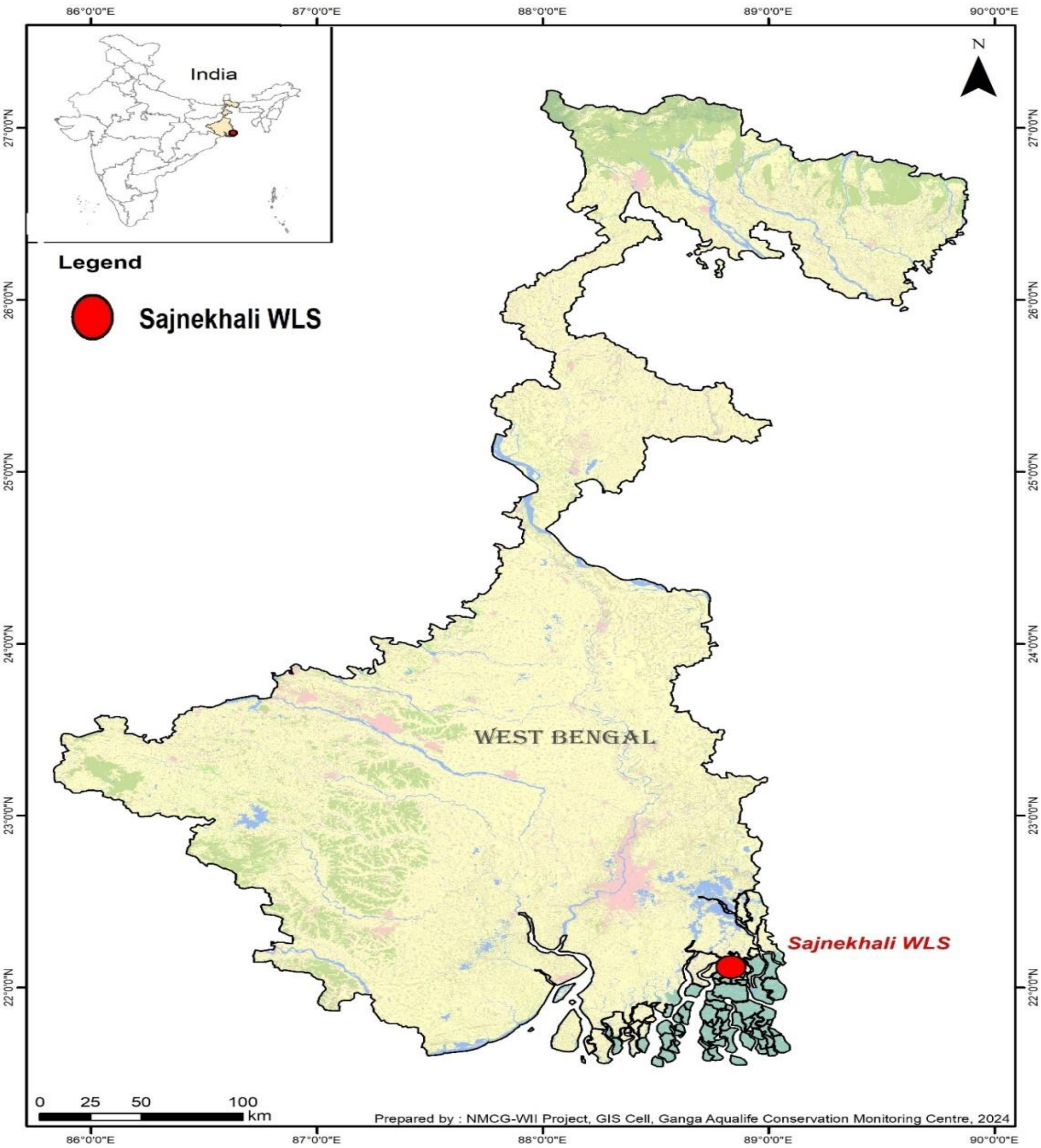


Figure 36: Location of Sajnekhali Wildlife Sanctuary, Sundarban Tiger Reserve

Figure 37: Non-invasive sampling of *B.baska* at Sajnekhali Breeding Centre, Sunderban Tiger Reserve.

6.5.3 DNA amplification and Sequencing

Out of the 40 samples, 26 were successfully processed to yield genomic DNA (gDNA) and from them 24 amplify specific regions of mitochondrial DNA (mtDNA). These samples were chosen to provide insights into the genetic diversity and demographic patterns of *B. baska*. For genetic diversity analysis, a highly variable region of the mtDNA was targeted for amplification and sequencing. Additionally, the complete region of the Cytochrome *b* gene was amplified in two samples using primers Cyt*b*G and mt-f-na3 (Spinks et al., 2004; Lenk et al., 1999) to elucidate the phylogenetic relationships of *B. baska* with other Geoemydidtaxa.In all samples, the mtDNA control region (CR) were amplified using primers CRF and CRR (Kumar et al. unpublished) (Table 17).

Table 17: Oligonucleotide primer sequences used in this study.

Gene	Primer name: sequence (5'-3')	Source
Cyt b	Cyt*b*G:AACCATCGTTGTWATCAACTAC mt-f-na3:AGGGTGGAGTCTTCAGTTTTTGGTTTACAAGACCAATG	Spinks et al., 2004 Lenk et al., 1999
CR	CRF: CATGAATCGGAGGCCAACCAGTCG CRR: AGTTGCTCTCGGATTTAGGG	Kumar et al., (Unpublished)

The amplified PCR productswere sequenced for both strands using the fluorescence-labelledBigDye Terminator (RR Mix) on theABI 3500 XL automated Genetic Analyzer (Applied Biosystems, Foster City,California, USA).

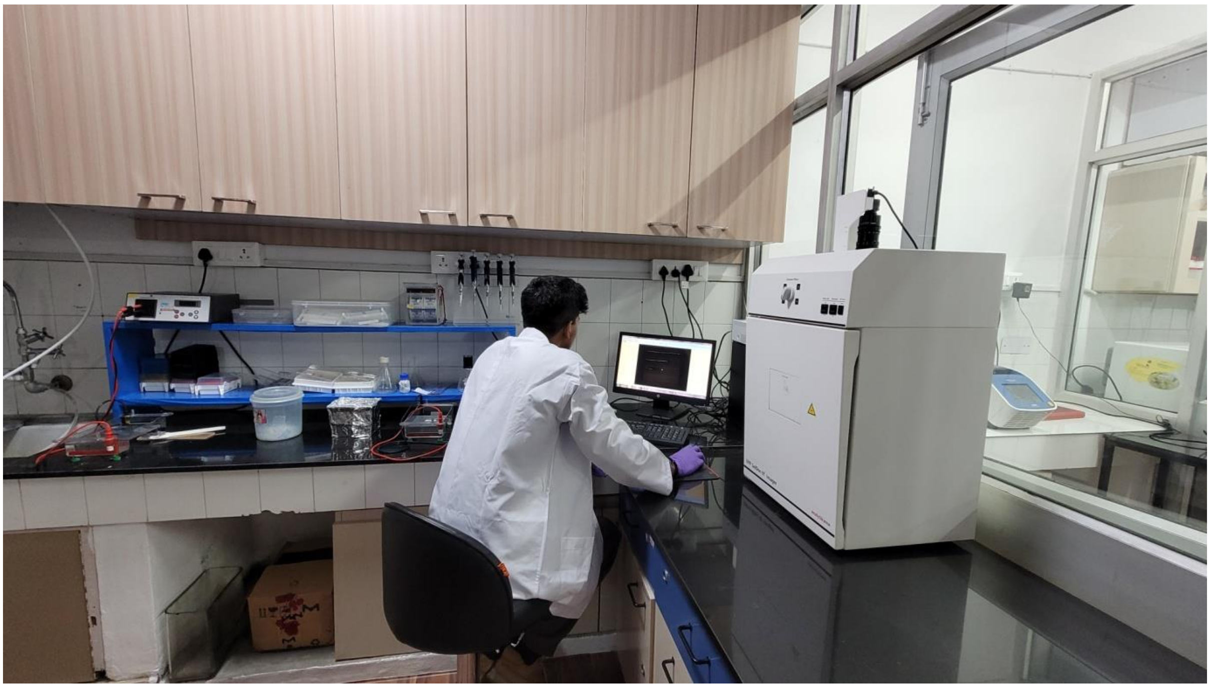


Figure 38: Laboratory analysis

6.5 Data analysis

The sequences obtained from the forward and reverse directions were aligned and edited using SEQUENCHER® version 4.9 (Gene Codes Corporation, Ann Arbor, MI, USA). The analysis of each sequence was performed separately using the CLUSTAL X multiple sequence alignment program (Thompson et al., 1997), and the alignments were examined by visual inspection. DnaSP 5.0 (Librado and Rozas, 2009) was used to analyze the haplotype diversity (h), nucleotide diversity (p), and polymorphic sites (s). The spatial distribution of the haplotypes was visualized through a median-joining network, which was created using the PopART software (Leigh and Bryant, 2015).

To determine whether the *B. baska* populations carried a signal of spatial range expansion or a stationary population history, Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997) neutrality test was performed in DnaSP 5.0 (Librado and Rozas, 2009). A total of 42 species of Testudines were used for the phylogenetic analysis based on their Cyt *b* gene including 6 species from *Batagur* genus. For the phylogenetic relationship, we used the Tamura-3 parameter using a discrete Gamma distribution (TN92+G) with the lowest BIC score value using MEGA 11 (Tamura et al., 2021).

6.6 Results and discussion

The conservation of the last surviving Indian population of *B. baska* in the Sunderban region is one of the key conservation challenges. Inadequate genetic management capacity and a limited understanding of the genetic background and diversity of small population often result in a suboptimal effective population size (Ne). This can lead to adverse outcomes such as inbreeding depression, significantly hindering the success of long-term breeding programs. Ne, which denotes the number of reproductively mature individuals contributing to the next generation, is vital for maintaining genetic diversity across generations. Effective management of Ne is crucial, as the genetic diversity of future generations is directly influenced by the number of reproductive individuals contributing to the gene pool. The current population size of *B. baska* in captivity, which has been recovered from thefounder stock, underscores the critical need for conservation efforts (Mallick et al., 2021). Genetic analysis using 24mtDNA sequences revealed only 5 haplotypes in *B. baska*. The haplotypes pattern suggests that most haplotypes were closely linked and only three mutation siteswere present on location number nt15646, nt16014 and nt16233 (Table 18). The median-joining network showed that most haplotypes appeared to be relatedand differ by only a single mutation site (Figure 39). Overall, only 3 variable sites were observed in the sequences. In at nucleotide position number nt15646, nt16014 and nt16233, accounting very low number of nucleotides changesper sites.

Table 18: Mitochondrial DNA control region haplotypes identified for the *B. baska*

Haplotypes	Number of sequences	Nt15646	Nt16014	Nt16233
H1	5	C	T	C
H2	6	T	.	.
H3	3	T	C	.
H4	2	T	.	T
H5	8	.	.	T

nt: nucleotide numbers showing the variation positions among *B. baksa* haplotypes.

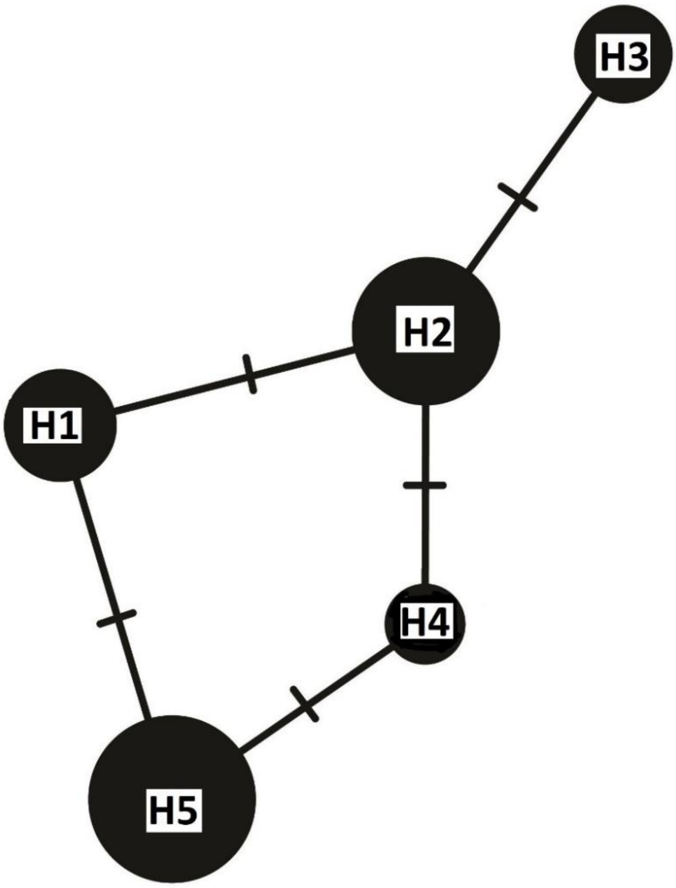


Figure 39: Median-joining network inferred from *Batagur baska* mitochondrial DNA haplotypes. The size of each circle indicates the relative frequency of the corresponding haplotype in the whole dataset. Short tick lines between haplotypes show the number of mutations.

6.6.1 Genetic diversity

The genetic diversity of *B. baska* was assessed using haplotype diversity (hd) and nucleotide diversity (π) calculations. The observed haplotype diversity was moderate 0.782 ± 0.042 within the population. Nucleotide diversity, however, was very low 0.00125 ± 0.0005, suggesting limited nucleotide variation in the population (Table 19). This could be due to factors such as population bottlenecks, genetic drift, or selective pressures that have reduced the overall genetic diversity at the nucleotide level. Understanding these genetic parameters is crucial for the conservation and management of *B. baska* populations, as it provides insights into their evolutionary history, population dynamics, and potential response to environmental changes.

Table 19: Summary of genetic diversity in *Batagur baska* populations based on mtDNA Control Region.

Sample Size	H	Vs	Hd	π	Tajima's D*	Fu's F_S *
24	5	3	0.782	0.0012	1.40	-0.237

Note: H- haplotype, Vs- Variable sites, Hd-haplotype diversity, π- nucleotide diversity, *P-values > 0.01 (not significant).

For comparison, the genetic diversity of other *Batagur* species was analyzed. *B. baska* exhibited 5 haplotypes among the total 24 samples, with hd and π calculated at 0.782 and 0.0012, respectively. In contrast, *B. dhongoka* showed higher genetic diversity, with 92 mtDNA sequences resulting in 31 haplotypes, and hd and π values of 0.94 and 0.002, respectively. Conversely, *B. borneoensis* displayed lower genetic diversity, with 90 mtDNA sequences yielding only four haplotypes, and hd and π values of 0.405 and 0.00076, respectively. Similarly, *B. affinis* showed low genetic diversity, with 120 mtDNA sequences resulting in six haplotypes, and hd and π values of 0.285 and 0.0044, respectively (Table 20).

Table 20-: Comparative of genetic diversity in other species of Batagur based on mtDNA

Species	N	H	Hd	π	Location	References
Batagur baska	24	5	0.782	0.0012	Sunderban Tiger reserve, India	Present study
Batagur dhongoka	92	31	0.94	0.002	Ganga river, India	Kumar et al, 2024
Batagur borneoensis	90	4	0.405	0.00076	Ujung Aceh Tamiang, Indonesia	Guntoro et al., 2020
Batagur affinis	120	6	0.285	0.0044	Peninsula Malaysia	Salleh et al., 2023

The results indicate varying levels of genetic diversity among the *Batagur* species. *B. dhongoka* exhibited the highest genetic diversity, with a large number of haplotypes and high values of haplotype diversity, suggesting a robust genetic structure possibly due to a large population size and/or historical demographic factors. In contrast, *B. borneoensis* and *B. affinis* displayed lower genetic diversity, indicated by fewer haplotypes and lower values of haplotype diversity and nucleotide diversity, possibly due to smaller population sizes, genetic bottlenecks, or founder effects. Considering this, *B. baska* also exhibited low genetic diversity as reported in *B. borneoensis* and *B. affinis* species, suggesting a low level of nucleotide variation among the analysed mtDNA sequences, which could be attributed due to a founder effect.

6.6.2 Demographic history

The demographic history of studied *B. baksa* was inferred from neutrality analysis. Results of Tajima's D and Fu's Fs tests were used to explain the population history of both populations. In overall samples, the value Tajima's D and Fu's FS was 0.676 and -1.041, respectively, but non-significant indicated stable population size. Moreover, effective population sizes and demographic trends were also estimated by the Bayesian skyline plot (BSP) analysis (Figure 40). The Bayesian Skyline Plot indicated decline trend in effective population size of *B. baska* from around 25 kya during the late Pleistocene. Oscillations in tidal zones, alterations in estuarine salinity levels due to climate change, and intensified human activities, including overexploitation, are critical factors leading to the significant population decline of *B. baska* (Mallick et al., 2021).

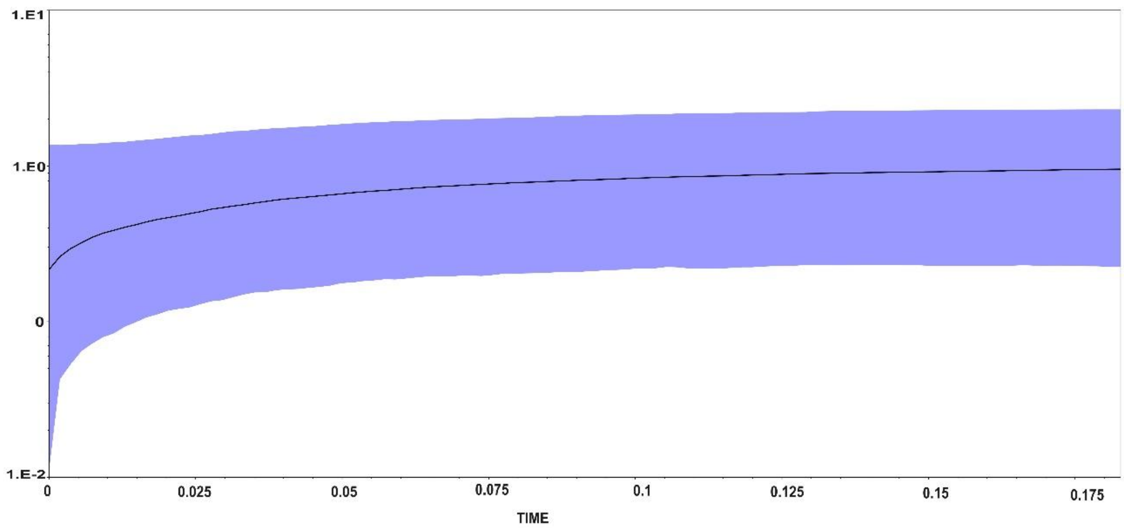


Figure 40: Image of Bayesian Skyline plot to visualize the demographic changes in effective population size (Ne) over time (x-axis) and a logarithmic scale, represents population size (y-axis).

6.6.3 Phylogenetic relationships

A total of 42 species of Testudines were included in our phylogenetic analysis, focusing on the cyt *b* gene. This dataset comprised 6 species from the *Batagur* genus, 32 from the Geoemydinae subfamily, and 4 from other families within Testudines (Figure 41).

In our Maximum Likelihood phylogenetic tree (ML analysis), both the *Batagur* and *Pangshura* genera were positioned basally to all other Geoemydids in the ingroup, with high bootstrap support (>80%). Regarding the *Batagur* genus, our analyses strongly support the riverine *B. kachuga* as basal to a terminal clade that includes the estuarine species *B. baska* and *B. affinis*. This clade is sister to another clade comprising the remaining three *Batagur* species, namely *B. dhongoka, B. borneoensis*, and *B. trivittata.* Our phylogeny, based on the cyt *b* gene, is consistent with the general topology of the phylogeny previously reported by Le et al. (2007) and Praschag et al. (2009). Furthermore, our analysis reveals that the genus *Cuora* is closely related to *Mauremys*, while *Cyclemys, Sacalia, Heosemys, and Notochelys* form a distinct cluster with significant posterior probability support.

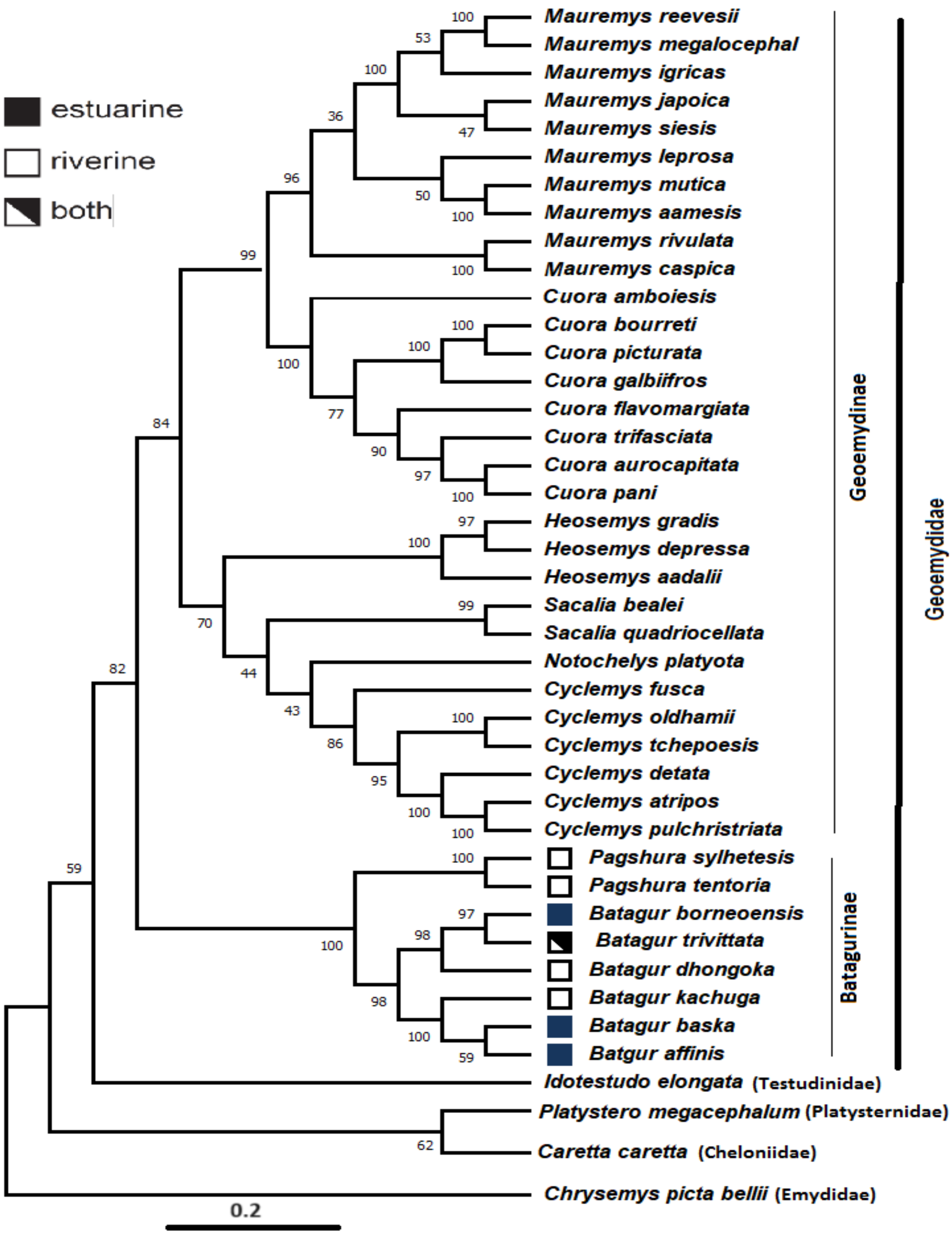


Figure 41: Phylogenetic relationship of *Batagur* inferred from Maximum Likelihood tree (ML using Cytochrome b gene. Node value represents the bootstrap support.

6.7 Further Initiative and Recommendations

6.7.1 Ecological aspects

The elusive reptile*B. baska*, in the Sunderban region, is ecological important because it play a specific role in maintaining the health and balance of the unique habitat. It contributes to the food web dynamics, nutrient cycling, and overall functioning of the mangrove ecosystem. The loss of this species could disrupt these ecological processes and lead to unintended consequences for other species. Efforts to recover the species in captivity from a limited founder stock underscore the urgent need for conservation actions. Restoring the ecology of the last small population of *B. baska* in the Sundarbans for their long-term survival involves a multifaceted approach combining habitat restoration, ex-situ and in-situ breeding program, and community involvement. Firstly, habitat restoration is crucial. This entails the rehabilitation of nesting sites and the protection of critical habitats from human disturbance and encroachment. Ensuring the availability of clean and suitable nesting beaches, along with the restoration of mangroves and estuarine systems, provides the necessary environment for breeding and foraging of *B. baska* (Figure 42).

6.7.2 Conservation aspects

The breeding projects for *B. baska* in India and Bangladesh aims to conserve and increase the genetic diversity of this critically endangered species. Following long-term monitoring programs in Bangladesh and the Sundarbans Tiger Reserve in India, the populations have shown signs of recovery. However, immediate conservation measures are essential to ensure the continued survival and genetic health of *B. baska*. These efforts include maintaining genetic diversity through careful management and implementing strategies to protect and restore their natural habitats.

Genetic diversity within and among species contributes to the evolutionary potential of the ecosystem, enabling it to withstand and adapt to changes over time. Genetic management is essential to maintain genetic diversity and prevent inbreeding depression in such a small population. To enhance genetic diversity and health, it is crucial to conduct genetic studies of *B. baska* populations in India as well as in Bangladesh. Developing a genetic database will be helpful in monitoring and managing the crucial genetic information across populations. Exchange programs of *B. baska* should be promoted through international collaboration and partnerships between the conservation organisations and government agencies of Bangladesh and India to enhance the genetic diversity. Protection and restoration of critical habitats within the Sundarbans and other key areas is vital. This includes implementing measures to reduce habitat destruction, pollution, and human disturbances.

The conservation plan involves collecting tissue samples from existing populations in both countries and conducting comprehensive genetic analyses to examine genetic relatedness, bottlenecks, and assess genetic variability. This genetic data will guide the selection of individuals for exchange, ensuring minimal risk of inbreeding and improving the genetic diversity. Developing a formal collaboration framework between conservation bodies in Bangladesh and India will facilitate the exchange of individuals, which will be monitored to assess their adaptation and integration into new environments. Identifying and prioritizing critical habitats for *B. baska* and implementing conservation measures to protect nesting sites, including creating protected areas and enforcing anti-poaching laws, are crucial steps. These recommendations, if implemented, can improve the chances of survival for this critically endangered species and contribute to overall biodiversity conservation in the region.

Figure 42: Mangroves forest of Sunderban, tiger researve

6.7.3 Community-based conservation

Community-based conservation initiatives involving local communities in the monitoring and protection of *B. baska* and its habitat can foster a sense of ownership and stewardship among stakeholders. Sustainable livelihood options, such as ecotourism or sustainable resource management, can provide alternative income sources and reduce pressure on the species and its habitat. Public awareness campaigns can help increase understanding of the conservation status of *B. baska* and the importance of protecting its habitat.

Regular monitoring and adaptive management strategies should be implemented to assess the effectiveness of conservation efforts and make necessary adjustments. Continued research and collaboration with scientific institutions can further enhance our understanding of *B. baska* and improve conservation strategies. Ultimately, a multi-faceted approach involving government agencies, non-governmental organizations, local communities, and international partners is essential for the successful conservation of *B. baska* in the Sunderbans and ensuring its survival for future generations.

6.8 Conclusion

The conservation status of *B. baska*, the last surviving Indian population in the Sunderban region, is critical, with the species facing extinction across much of its former range. Efforts to recover the species in captivity from a limited founder stock underscore the urgent need for conservation actions. Genetic analysis revealed low nucleotide diversity in *B. baska*, indicating limited nucleotide variations, possibly due to smaller population sizes and founder effects. Comparative analysis with other *Batagur* species showed varying levels of genetic diversity. Effective population sizes and demographic trends estimates by the BSP analysis indicated the declining trend in *B. baska*.

To ensure the survival of *B. baska*, management recommendations include intensive population monitoring, habitat protection measures, genetic management plans, further research, collaboration with conservation organizations, strong legal protection, and public awareness campaigns. Regular monitoring and adaptive management strategies are crucial, along with continued research and collaboration, for the successful conservation of *B. baska* in the Sunderbans.

6.9 References

- Chelonian Research Monographs No. 5: 037.1–037.10.
- Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147,915–925.
- Gaur, A., Reddy, A., Annapoorni, S., Satyarebala, B. and Shivaji, S. (2006) The origin of Indian Star tortoises (*Geochelone elegans*) based on nuclear and mitochondrial DNA analysis: A story of rescue and repatriation. Conser. Genet. 7, 231-240.
- Ghosh, A. & N.R. Mandal (1990). Studies on nesting and artificial hatching of the endangered river terrapin *Batagur baska* (Gray) in the Sundarbans Tiger Reserve, West Bengal. Journal of the Bombay Natural History Society 87: 50–52.
- Guntoro, J.; Riyanto, A. The very low genetic variability on Aceh Tamiang's (Indonesia) population of Painted Terrapin (*Batagur borneoensis*) inferred by cytochrome oxidase I (COI) and D-loop (control region). Biodiversitas J. Biol. Divers. 2020, 21, 2514–2520.
- Kumar, A., Panda, A.K., Usmani, A.A., Yadav, P., Panwar, A., Badola, R., Hussain, S.A. and Gupta, S.K., 2024. Population genetics of the critically endangered three-striped turtle, *Batagurdhongoka*, from the Ganga River system using mitochondrial DNA and microsatellite analysis. Scientific Reports, 14(1), p.5920.
- Le, M., Duong, H.T., Dinh, L.D., Nguyen, T.Q., Pritchard, P.C. and McCormack, T., 2014. A phylogeny of softshell turtles (Testudines: Trionychidae) with reference to the taxonomic status of the critically endangered, giant softshell turtle, *Rafetus swinhoei*. Organisms Diversity & Evolution, 14: 279-293.
- Leigh, J. W., & Bryant, D. (2015). Popart: full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110-1116.
- Lenk, P., Fritz, U., Joger, U., & Wink, M. (1999). Mitochondrial phylogeography of the European pond turtle, Emys orbicularis (Linnaeus 1758). Mol. Ecol. 8, 1911-1922.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.
- Luikart, G. and Cornuet, J.M., 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation biology, 12(1), pp.228-237.
- Mallick, N., S. Singh, D. Chatterjee & S. Sharma (2021). Conservation breeding of Northern River Terrapin *Batagur baska* (Gray, 1830) in Sundarban Tiger Reserve, India. Journal of Threatened Taxa 13(6): 18544–18550.
- Moll, E.O., K. Platt, S.G. Platt, P. Praschag & P.P. Van Dijk (2009). *Batagur baska* (Gray 1830)-northern river terrapin. In: Rhodin, A.G.J., P.C.H. Pritchard, P.P. van Dijk, R.A. Saumure, K.A. Buhlmannn, J.B. Iverson & R.A. Mittermeier (eds.). Conservation Biology of Freshwater Turtles and Tortoises: A Compilation Project of the IUCN/SSC Tortoise and Freshwater Turtle Specialist Group.
- Murphy, M. O., Jones, K. S., Price, S. J., & Weisrock, D. W. (2018). A genomic assessment of population structure and gene flow in an aquatic salamander identifies the roles of spatial scale, barriers, and river architecture. Freshwater Biology, 63(5), 407-419.
- Paul, P., Stuart, L., & Rhodin, G. (2000). Asian Turtle Trade.
- Praschag, P. & Singh, S. 2019. *Batagur baska*. The IUCN Red List of Threatened Species 2019: e.T97358453A2788691. https://dx.doi.org/10.2305/IUCN.UK.2019-1.RLTS.T97358453A2788691.en. Accessed on 04 April 2024.

- Praschag, P., R. Holloway, A. Georges, M. Päckert, A. K. Hundsdörfer & U. Fritz (2009): A new subspecies of *Batagur affinis* (Cantor, 1847), one of the world's most critically endangered chelonians (Testudines: Geoemydidae). – Zootaxa, 2233: 57–68.
- Praschag, P., Sommer, R.S., Mccarthy, C., Gemel, R. and Fritz, U. 2008. Naming one of the world's rarest chelonians, the southern *Batagur*. Zootaxa 1758: 61–68.
- Salleh, M.H.M., Esa, Y. and Pau, S.S.N., (2023). Conservation genetics of the critically endangered Southern River Terrapin (*Batagur affinis*) in Malaysia: genetic diversity and novel subspecies distribution ranges. Biology, 12:520.
- Spinks, P. Q., Shaffer, H. B., Iverson, J. B., & McCord, W. P. (2004). Phylogenetic hypotheses for the turtle family Geoemydidae. Mol. Phylogenet. Evol. 32, 164-182 (2004).
- Tajima, H. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.
- Tamura, K., Stecher, G. & Kumar, S. (2021) MEGA11: Molecular evolutionary genetics analysis version 11. Mol. Biol. Evol 38, 3022–3027.
- Weissenbacher, A., D. Preininger, R. Ghosh, A. G. J. Morshed & P. Praschag (2015): Conservation breeding of the northern river terrapin *Batagur baska* at the Vienna Zoo, Austria, and in Bangladesh. – International Zoo Yearbook, 49: 31–41.

7. UNDERSTANDING THE GENETIC DIVERSITY WITHIN THE *PANGSHURA* GENUS IN THE GANGA RIVER SYSTEM: IMPLICATIONS FOR CONSERVATION AND SUSTAINABLE MANAGEMENT

7.1 Background of the study

The Geoemydidae family previously known as Bataguridae, encompassing the highly threatened turtle group, plays a pivotal role in freshwater ecosystems (Van Dijk et al., 2000; IUCN, 2021). Comprising primarily freshwater turtles, with some adapted to estuarine and terrestrial habitats, the family Geoemydidae, reflecting diverse ecological adaptations (Van Dijk et al., 2000). With 71 extant species, Geoemydidae is closely related to land tortoises (Family Testudinidae; TTWG, 2017).

The genus *Pangshura*, belongs to family Geoemydidae, shows significant sexual dimorphism and relatively smaller-sized turtles, with maximum shell lengths ranging from 20 to 26.5 cm (Ernst et al., 2000). Despite their ecological importance, *Pangshura* faces severe threats, with four extant species—*P. tecta, P. tentoria, P. smithii, and P. sylhetensis*—distributed across the Southeast Asian landscape. Additionally, a now-extinct species, *P. tatrotia*, existed during the Pliocene epoch in the Siwalik Hills of Pakistan (Walter & Tyler, 2010). Fossil records from the Siwalik Hills and Narmada Valley deposits indicate the historical presence of *Pangshura* in India since at least the Pleistocene epoch (Baruah, Devi, & Sharma, 2016). Further classification reveals multiple existence of subspecies in *Pangshura*, such as *P. tentoria tentoria, P. tentoria circumdata, P. tentoria flaviventer, P. smithii,* and *P. smithii pallidipes*, each displaying distinct morphological characteristics (Praschag et al., 2007).

Presently, the IUCN status designates *P. tecta* as vulnerable, *P. smithii* as near threatened (figure 43), *P. tentoria* as least concern, and *P. sylhetensis* as critically endangered category, emphasizing the urgent need for conservation measures. However, despite their ecological and conservation significance, all species of *Pangshura*, face imminent threats including egg harvesting, overexploitation for food, and habitat alternation. Given their precarious status, comprehensive genetic studies are imperative for effective conservation planning.

The utilization of molecular data has been widely used in Testudine systematics and conservation genetics, contributing to the identification of new species, assessment of population structure, and elucidation of evolutionary and phylogenetic relationships (Murphy et al., 2013; Fritz et al., 2018; Le et al., 2006). In this context, the present study focuses on the *Pangshura* genus, aiming to unravel its genetic diversity using the cytochrome *b* gene, a highly informative mitochondrial marker. The study has two primary objectives: (i) to identify genetic lineages within *Pangshura* in the Ganga River basin, shedding light on the broader distribution of genetic diversity, and (ii) to explore past demographic changes, correlating them with historical climatic fluctuations and the expansion of *Pangshura* populations in the Indian river system. This research endeavors to contribute valuable insights crucial for the conservation and management of this threatened turtle group.

Table 21: Present IUCN Status, distribution and population trend of *Pangshura* (Figure 44)

Species	Common name	IUCN Status	Distribution	Population trend
P.tecta	Indian roofed turtle	VU	Indus, Saharmati, Mahi, Narmada, Ganga and Mahanadi River systems of Pakistan, India, Nepal and Bangladesh	Declining
P.tentoria	Indian tent turtle	LC	Ganga, Brahmaputra, Mahanadi and Godavari River basins	Stable
P.smithii	Brown roofed turtle	NT	Indus, Ganga and Brahmaputra River systems in Pakistan, India, Nepal and Bangladesh	Declining
P.sylhetensis	Assam roofed turtle	CR	Bhutan, India, Myanmar and Bangladesh (possibly extinct)	Declining

Figure 43: Image of *Pangshura smithii*

Pangshura tecta

Pangshura tentoria

Pangshura smithii

Pangshura sylhetensis

Figure 44: Map showing the distribution range of *Pangshura*. Map adapted from Chelonian Research Foundation 2020; IUCN Red List of Threatened Species. Version 2023-1

7.2 Methodology

This study utilized a comprehensive dataset consisting of 117 samples/sequences of *Pangshura*, including 58 sequences of *P. smithii*, 38 sequences of *P. tentoria*, and 21 sequences of *P. tecta* (Table 22). The sampling analysis strategy covered major river systems of Ganga and other rivers. The sequences of all rivers except ganga river were included from NCBI database. Collection procedures employed non-invasive techniques such as mouth swabbing and clipping a small piece of the toe pad, ensuring minimal disturbance to the animals, which were subsequently released back into their natural habitat. Turtles were identified based on morphological traits, and validation from herpetologists was sought when necessary. Whenever possible, samples were also taken from the carcasses and dead remain. The collected samples were preserved in 95% ethanol and stored at 20°C until DNA extraction.

Table 22: Details of sequences from different rivers and regions used in the study

Species	River	Number of Samples/Sequences
Pangshura smithi	Ganga	38
	Ghaghra	5
	Chambal	3
	Brahmaputra	3
	Uttar Pradesh Unknown	9
		58
Pangshura tentoria	Ganga	10
	Gomti	3
	Chambal	2
	Koshi	5
	Ghaghra	2
	Uttar Pradesh Unknown	6
	Bangladesh	3
	Brahmaputra	7
		38
Pangshura tecta	Ganga	6
	Uttar Pradesh Unknown	5
	Chambal	9
	Brahmaputra	1
		21

7.2.1 DNA extraction, amplification and sequencing

Total genomic DNA from biological samples were extracted using DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). The complete cytochrome b (Cyt *b*) gene was amplified using the primers: CytbG:5'-AACCATCGTTGTWATCAACTAC-3' (Spinks and Shaffer 2004) and mt-f-na3:5'-

AGGGTGGAGTCTTCAGTTTTTGGTTTACAAGACCAATG-3’ (Lenk et al. 1999). PCR reactions were performed in total reaction volumes of 20 µl using a PCR buffer (10 mM Tri–HCl, pH 8.3, and 50 mM KCl), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 2 pmol of each primer, 5.0 U of Taq DNA polymerase and 1 µl (~30 ng) of the template DNA. All reactions were run along with negative controls. The PCR conditions were 95°C for 5 min followed by 35 cycles at 95°C for 45 s, annealing 54-56°C for 45 s and extension 72°C for 1 min, with a final extension of 72°C for 10 min. The effectiveness and consistency of the PCR reactions were monitored using positive controls. The amplified PCR amplicons were visualized in UV light on 2% agarose gel stained with ethidium bromide. Exonuclease I (EXO-*I*) and shrimp alkaline phosphatase (SAP) treatments were given to the amplified PCR products (USB, Cleveland, OH) for 15 minutes each at 37°C and 80°C, respectively, to eliminate any residual primer. The amplified PCR products were directly sequenced using the BigDye® Terminator Kit (v3.1) and analysed on an ABI 3500XL Applied Biosystems Genetic Analyzer. All the products were sequenced in both directions.

7.3 Data analysis

7.3.1 Gene diversity estimates

All the sequences were aligned using CLUSTAL W (Thompson et al. 1994), as implemented in the BioEdit v 7.2.5 software (Hall 1999) and manually checked. To estimates the level of genetic diversity number of haplotypes (h), haplotype diversity (hd) and nucleotide diversity (π) within the Ganga population and other known global populations of Lissemys were computed using the software DNASPv5.0 (Librado and Rozas 2009).

7.3.2 Demography estimates

To determine the demography history of *Lissemys* population, we performed different statistical approaches such as Tajima’s *D* (Tajima 1989), Fu’s F*s* test (Fu 1997). Mismatch distribution analysis, a sum of squared deviations (SSD), and the raggedness index (r) were also used to demonstrate the pattern of population stability or population expansion in the past under the sudden demographic expansion models using the ARLEQUIN v3.5 program (Excoffier and Lischer 2010). The *P*-values were obtained from 1000 simulations on the basis of a selective neutrality test. A Bayesian skyline plot (BSP) was constructed using the Monte Carlo Markov Chain (MCMC) method with 100 million generations using BEAST ver 1.7.5 (Drummond et al. 2012). The temporal trends in the effective population size of the *Pangshura* overtime/generations were estimated using a coalescent BSP.

7.3.3 Genetic differentiation

The conventional *F*ST i.e. Wright’s fixation index of population subdivision, was calculated to test for differences in haplotype frequencies (Weir and Cockerham 1984) and ϕ statistics was calculated which incorporate the information of nucleotide differences between haplotypes (Excoffier et al.1992). The significant P values of the statistics were computed using a nonparametric permutation approach with 1000 permutations.

7.3.4 Gene flow pattern estimates

The spatial distribution and relatedness of haplotypes among the *Pangshura* was visualized by a median-joining network, was created with the PopART software (Leigh & Bryant, 2015).

7.4 Results and Discussion

7.4.1 Genetic diversity in *Pangshura*

Cytochrome b sequences (Cyt b~1038bp size) were generated in both forward and reverse directions. Additionally, Cyt b sequences from various localities and rivers were retrieved from the GenBank database at the National Center for Biotechnology Information, USA.

Within *P. smithii*, 58 sequences were organized into 15 haplotypes, while *P. tentoria* exhibited 10 haplotypes among its 38 sequences, and *P. tecta* displayed 9 haplotypes with its 21 sequences. The gene diversity for *P. smithii* was relatively low, with hd=0.592 and π=0.0029, in contrast to *P. tentoria*, which showed higher gene diversity (hd=0.814 and π=0.00323), and *P. tecta*, with a gene diversity of hd=0.852 and π=0.00171, respectively (Table 23).

Table 23: Sample size (N), number of haplotypes (H), haplotype diversity (Hd), nucleotide diversity (Pi), Fu's Fs, and Tajima's D in *Pangshura* populations

Species	N	H	Hd	Pi	Tajima's D*	Fu's Fs
P. smithii	58	15	0.592	0.0029	-2.79*	-3.54
P.tentoria	38	10	0.814	0.00323	-0.015	-0.470
P.tecta	21	9	0.852	0.00171	-1.25	-3.67

7.4.2 Genetic differentiation in *Pangshura*

The analyses unequivocally confirmed significant genetic differentiation among all *Pangshura* species. The within-group mean distance was notably high in *P. smithii* and *P. tentoria*, while it exhibited a lower value in *P. tecta* (Figure 45).

The genetic differentiation percentages further revealed distinctions: 5.9% between *P. smithii* and *P. tentoria*, 6.4% between *P. tecta* and *P. tentoria*, and 7.2% between *P. smithii* and *P. tecta*. Based on these estimates of genetic differentiation, it is evident that *P. smithii* and *P. tentoria* are more closely related than *P. tecta.*

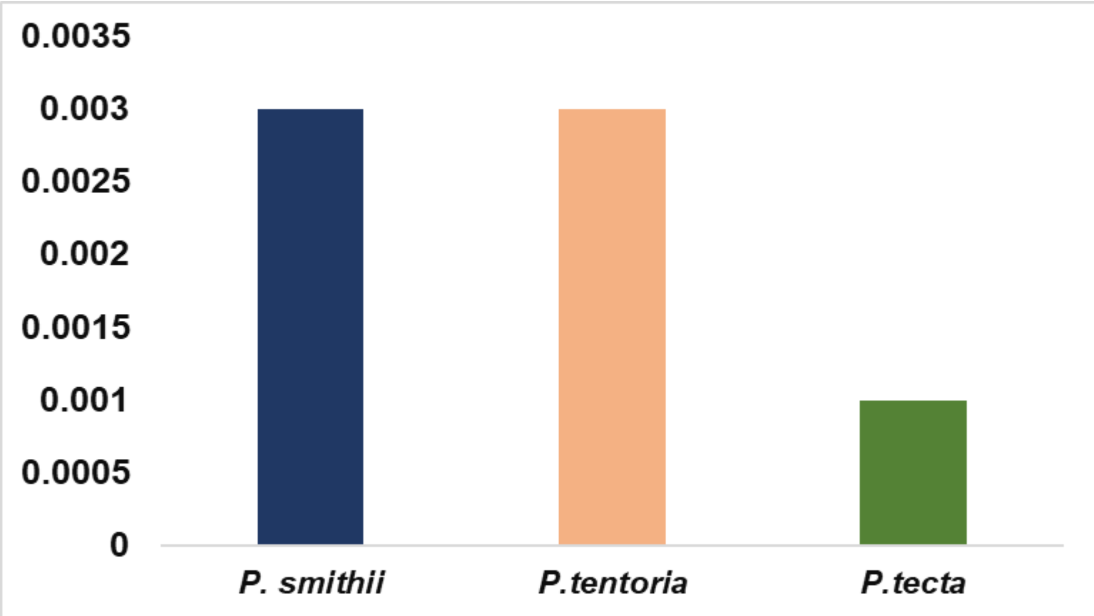


Figure 45: Graph showing the within mean group distance in three species of *Pangshura*

7.4.3 Population demography

The neutrality test analysis, Tajima's *D* and Fu's F*s* tests used employed to infer the demographic history of all *Pangshura* populations (Table 23). The high significant negative Tajima's *D* value was observed in *P. smithii* whereas it was non-significant negative for *P. tentoria* and *P. tecta*. No statistical significance values for Fu's F*s* were observed in all *Pangshura* populations but shows high negative values. The significant negative values observed for both Tajima's D and Fu's Fs in the *Pangshura* population within the Ganga River system suggest the presence of rare nucleotide site variants. These findings serve as evidence of deviation resulting from sudden population growth or selection pressures. In support of this, Bayesian skyline plot (BSP) analyses were conducted, further substantiating the hypothesis of population expansion in *P. smithii* and *P. tentoria* populations after sharp decline. *P. tecta* showed prolonged phase of demographic stability (Figure 46). This anomaly may be attributed to either a limited sample size or a close genetic relationship among individuals of *P. tecta*.

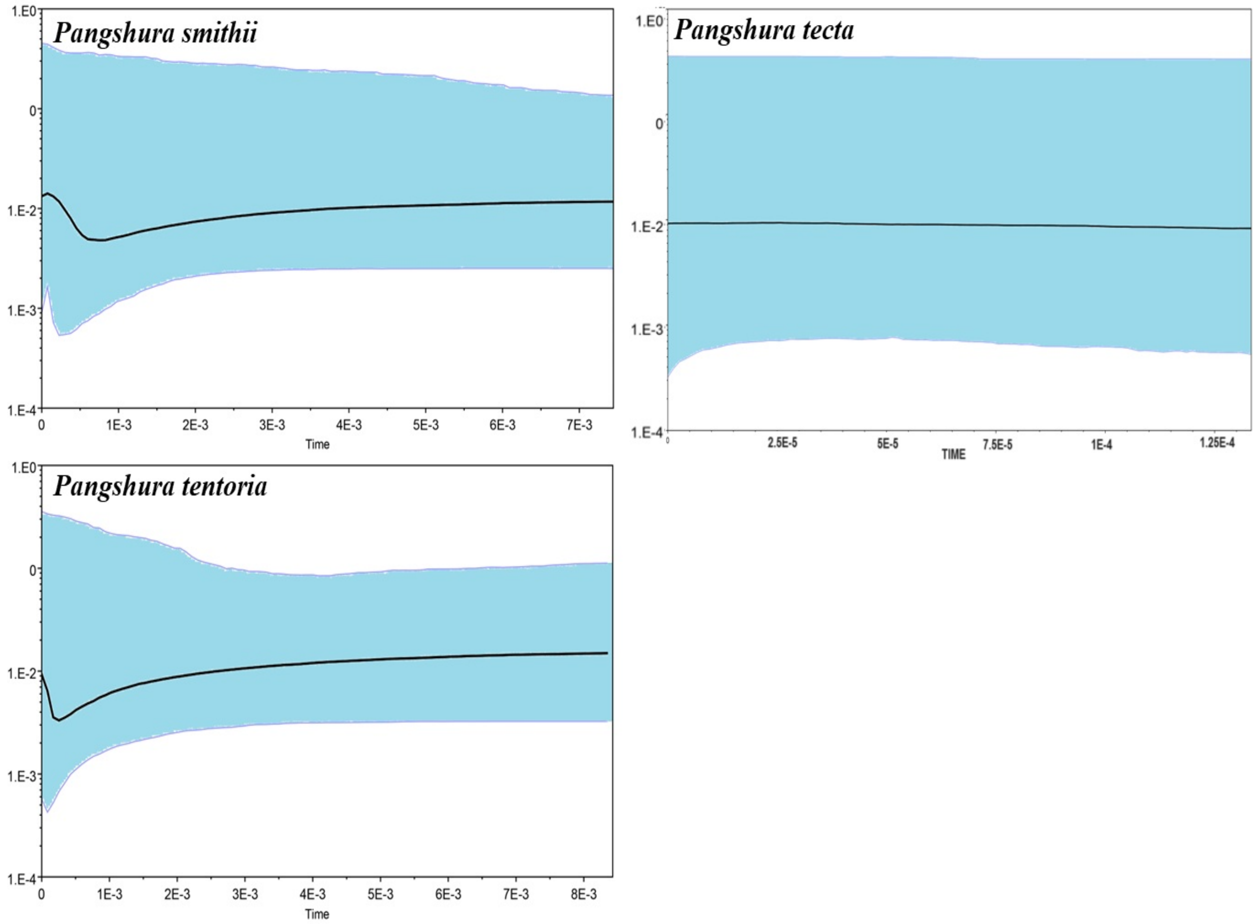


Figure 46: Bayesian skyline plots (BSPs) estimated by BEAST for *Pangshura* populations in the *X*-axis shows time before present (Ma). The *Y*-axis (logarithmic scale) expressed the effective population size (Ne) estimates multiplied by generation time (τ). The solid line indicates the median of population size, and the 95% highest posterior density (HPD) credibility interval is depicted in blue.

7.4.5 Geneflow and haplotypes linkage

A median joining network was constructed to elucidate the genetic relationships among all haplotypes belonging to *Pangshura* populations.

Pangshura smithii: In the case of *P. smithii*, 15 haplotypes were identified, and core haplotypes were notably present in the Ganga River, sharing genes with the Ghaghra and Chambal rivers, as well as samples from unidentified locations in Uttar Pradesh rivers. Additionally, some haplotypes were shared between the Ghaghra and Brahmaputra rivers, as well as between the Ganga and unknown locations in Uttar Pradesh (Figure 47). A star-like network pattern emerged in *P. smithii*, aligning with and supporting the results of the neutrality test and Bayesian skyline plot (BSP) analysis, suggesting population expansion within the Indian river system (Figure 39). The absence of distinct river-wise clustering in the studied population suggests a significant level of maternal gene flow within *P. smithii* across diverse river systems.

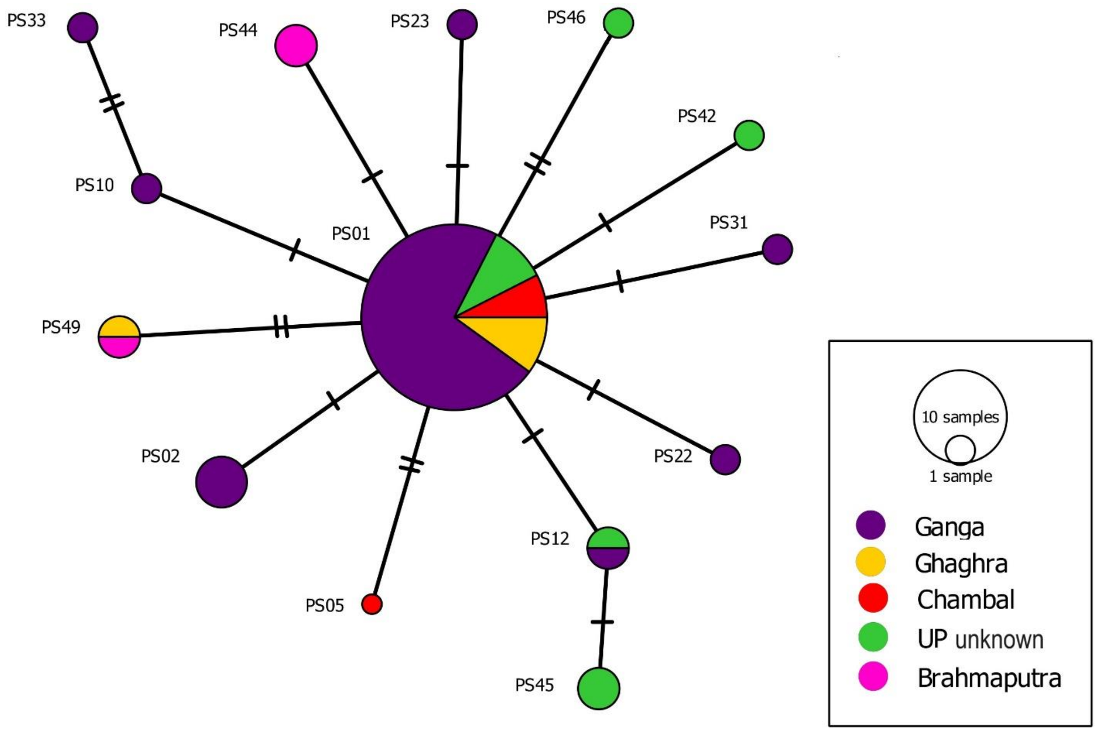


Figure 47: Median-joining network of *Pangshura smithii*

***Pangshura tentoria*:** In the case of *P. tentoria*, 10 haplotypes were identified, and three core haplotypes were observed indicates evolve of different maternal line in tentoria. Sharing of gene were observed in Ganga, Gomti, Chambal, Ghaghra and Brahmaputra River constitute core haplotype. Sharing of haplotypes were also between Gomti, unknown localities of Uttar Pradesh, Chambal and Bangladesh population (Figure 48). The absence of distinct river-wise clustering in the *P. tentoria* suggests a significant level of maternal gene flow across diverse river systems.

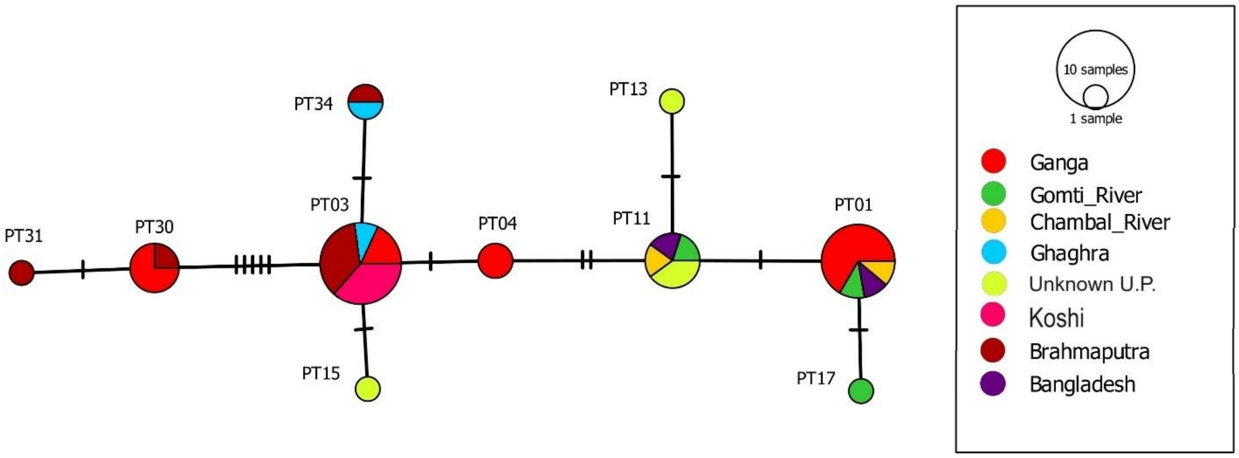


Figure 48: Median-joining network of *Pangshura tentoria.*

***Pangshura tecta*:** In *P. tecta*, 9 haplotypes were identified, and sharing of gene were observed in Ganga, Chambal, and unknown localities of Uttar Pradesh. The absence of distinct river-wise clustering in the *P. tecta* suggests a significant level of maternal gene flow across diverse river systems and having stable population size (Figure 49).

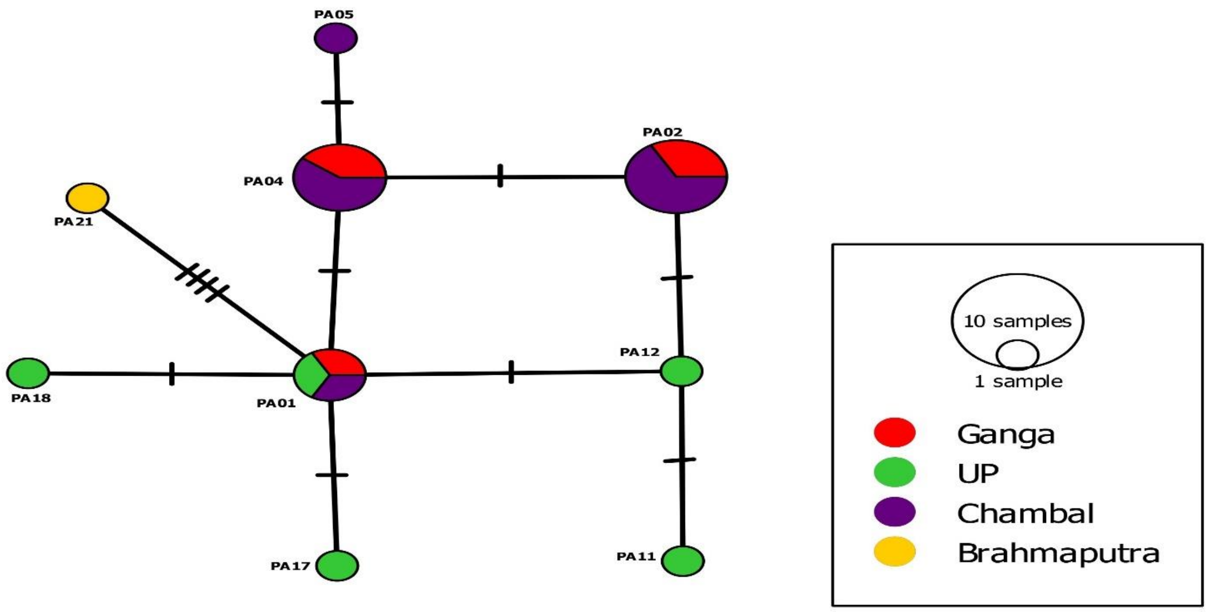


Figure 49: Median-joining network of *Pangshura tecta.*

7.5 Conclusion

Examining the genetic structure and diversity of a species facing significant threats provides crucial insights for its development and effective management (Kumar et al., 2017; Gupta et al., 2018). This study presents, for the first time, detailed genetic data to estimate the level of genetic diversity, gene flow, and demography of all three species of *Pangshura*, namely *P. smithii*, *P. tentoria*, and *P. tecta*, in different river systems of India.

Our data revealed high haplotype diversity (>0.80) in *P. tentoria* and *P. tecta*, whereas it was comparatively low in *P. smithii*. Interestingly, recent population expansion is observed in *P. smithii*, indicating that the population is genetically fit but requires proper habitat management for long-term conservation. Network analysis revealed the presence of highly diverse lineages of *P. tentoria* in the Ganga River system. All *Pangshura* species are maternally well-connected within their respective species and exhibit genetic fitness. Molecular genetic analysis emphasizes the dire need for further sampling in other river tributaries to develop effective conservation strategies. The present data also aid in the restoration of rescued individuals to their natural habitat.

7.6 References

- Baruah, C., Devi, P., & Sharma, D. K. (2016). Comparative morphometry and biogeography of the freshwater turtles of genus Pangshura(Testudines: Geoemydidae: Pangshura). International Journal of Pure and Applied Zoology, 4, 107–123.
- Chelonian Research Monographs No. 5: 037.1–037.10.
- Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution, 29(8), 1969-1973.
- Ernst, C. H., Altenburg, R. G. M., & Barbour, R. W. (2000). Turtles of the World. World Biodiversity Database, CD-ROM Series, Windows, Version 1.2. Amsterdam, The Netherlands: Biodiversity Center of ETI.
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources, 10(3), 564-567.
- Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics, 131(2), 479-491.
- Fritz, U., Guicking, D., Auer, M., Sommer, R. S., Wink, M., &Hundsdörfer, A. K. (2008). Diversity of the Southeast Asian leaf turtle genus Cyclemys: How many leaves on its tree of life? Zoologica Scripta, 37, 367–390
- Gupta, S. K., Kumar, A., Angom, S., Singh, B., Ghazi, M. G. U., Tuboi, C., & Hussain, S. A. (2018). Genetic analysis of endangered hog deer (Axis porcinus) reveals two distinct lineages from the Indian subcontinent. Scientific reports, 8(1), 16308.
- Hall, T. A. (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).
- Kumar, A., Ghazi, M. G. U., Hussain, S. A., Bhatt, D., & Gupta, S. K. (2017). Mitochondrial and nuclear DNA based genetic assessment indicated distinct variation and low genetic exchange among the three subspecies of swamp deer (Rucervusduvaucelii). Evolutionary Biology, 44, 31-42.
- Le, M., Raxworthy, C. J., McCord, W. P., & Mertz, L. (2006). A molecular phylogeny of tortoises (Testudines: Testudinidae) based on mitochondrial and nuclear genes. Molecular Phylogenetics and Evolution, 40, 517–531. https://doi.org/10.1016/j.ympev.2006.03.003
- Lenk, P., Fritz, U., Joger, U., & Wink, M. (1999). Mitochondrial phylogeography of the European pond turtle, Emys orbicularis (Linnaeus 1758). Molecular ecology, 8(11), 1911-1922.
- Leigh, J. W., Bryant, D., & Nakagawa, S. (2015). POPART: full-feature software for haplotype network construction. Methods in Ecology & Evolution, 6(9).
- Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25(11), 1451-1452.
- Murphy, R. W., Crawford, A. J., Bauer, A. M., Che, J., Donnellan, S. C., Fritz, U., … Zhang, Y. (2013). Cold Code: The global initiative to DNA barcode amphibians and nonavian reptiles. Molecular Ecology Resources, 13, 161–167. https://doi.org/10.1111/1755-0998.12050

- Praschag, P., Hundsdörfer, A. K., & Fritz, U. (2007). Phylogeny and taxonomy of endangered South and South-east Asian freshwater turtleselucidated by mtDNA sequence variation (Testudines: Geoemydidae:Batagur, Callagur, Hardella, Kachuga, Pangshura). Zoologica Scripta, 36,429–442.
- Spitzweg, C., Praschag, P., DiRuzzo, S., & Fritz, U. (2018). Conservationgenetics of the northern river terrapin (Batagurbaska) breeding project using a microsatellite marker system. Salamandra, 54, 63–70.
- Spinks, P. Q., Shaffer, H. B., Iverson, J. B., & McCord, W. P. (2004).Phylogenetic hypotheses for the turtle family Geoemydidae.Molecular Phylogenetics and Evolution, 32, 164–182. https://doi.org/10.1016/j.ympev.2003.12.015
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids research, 22(22), 4673-4680.
- Van Dijk, P. P., Stuart, B. L. and Rhodin, A.G.J., 2000. Asian Turtle Trade. Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia, Phnom Penh, Cambodia, 1–4 December 1999. Chelonian Res. Monographs 2: 1-164.
- Walter, G. J., & Tyler, R. L. (2010). Pangshuratatrotia, a new species of pond turtle (Testudinoidea) from the Pliocene Siwaliks of Pakistan. Journal of Systematic Palaeontology, 8, 449–458.
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. evolution, 1358-1370.

8. GENETIC ASSESEMENT OF GANGETIC DOLPHIN, (*Platanista gangetica gangetica*)

8.1. Backgroud of the study

The South Asian River dolphins of the genus *Platanista* is the most endangered cetaceans in the world. It is an iconic freshwater species and the only extant representative of the family Platanistidae. Currently, two geographically isolated subspecies of South Asian River dolphin is recognized; the Indus river dolphin (*P. gangetica minor*) restricted to the Indus river system of India and Pakistan, and the Gangetic dolphin (*P. g. gangetica*) inhabiting Ganga (Figure 50), Brahmaputra, Meghna, Karnaphuli–Sangu, Karnali, Sapta Koshi, and Narayani of India, Bangladesh Nepal and potentially Bhutan (Sinha et al. 2000; Smith et al. 2001; Sinha and Kannan 2014). The Ganga is the lifeline and holy river of India, and the basin supports home to more than 600 million peoples (World Bank Report No. STEP705, 2011). Despite its ionic and religious status, today's high anthropogenic pressure is a major threat to its biodiversity. As large mammals of river Ganga, the Gangetic dolphin experiences many risks such as fragmented pollutions, dam, barrages, low water levels, and illegal catching for their oil and meat. The extracted oils are used to catch catfish in the net fishery (Jones 1982). These threats are long-lasting and challenging to manage; therefore, distribution range and population size of dolphin is drastically declining (Sinha and Sharma 2003). Hence, the species is listed under IUCN Red List as "Endangered" (Braulik and Smith 2019) and protected in Schedule *I* of the Indian Wildlife (Protection) Act, 1972. It is included in Appendix *I* of the Convention on International Trade in Endangered Species of Flora and Fauna (CITES), and was recently upgraded to Appendix *I* in the Convention on Migratory Species (CMS). The genus *Platanista* is considered as the most primitive among all other dolphins with diverse evolutionary history; hence it has always been in the debate (de Muizon 1994; Messenger 1994; Árnason1996; Guang and Kaiya1999; Cassens et al. 2000; Hamilton et al. 2001; Nikaido et al. 2001; Yang et al. 2002; Verma et al. 2004; Yan et al. 2005; Xiong et al. 2009).

The genetic relationship between two subspecies of *Platanista* (*minor* and *gangetica*) was investigated using the control region and cytochrome *b* region, from museum specimens (Braulik et al. 2015). High resolution of phylogenetic relationships and genetic studies can be achieved by sequencing the extended coverage of mitogenome (Rokas and Carroll 2005). Till date, information about the mitogenome level characterization of Gangetic dolphin and detailed relationship with subspecies and other river dolphins are lacking that would help to discriminate them. For the effective management plan, it is vital to understand the genetic pattern and adaptive potential of endangered Gangetic dolphins from its distribution range. Therefore, knowledge of complete mitogenome not only provides information about species identification but also offers a wide range of genomic characterization, which are useful for studying evolutionary and demography history of species and populations (Brown et al. 1982; Avise 1989). In the vertebrates, the complete mitogenome is a small and closed-circular molecule, which generally contains 37 genes including 13 protein-coding genes (PCGs), 2 ribosomal RNA genes (rRNAs), and 22 transfer RNA genes (tRNAs) (Boore 1999; Singh et al. 2019; Kumar et al. 2019). In the present study, we generated the sequence of the whole mitogenome of the Gangetic dolphin from different stretches of river Ganga and describe the gene organization, base compositions, codon usage. We also aim to assess the genetic diversity of Gangetic dolphin based on mtDNA control region.

Figure 50: Image of river dolphin (Pic credit: https://en.wikipedia.org/wiki/South_Asian_river_dolphin)

8.2. Methodology

We used tissue samples from six Gangetic dolphins. Of these, two were from Narora, and one each from Balia, Patna, Suphaul, and Araria (Figure 51A). All the samples were collected from dead individuals during the ecological survey of NMCG project, Wildlife Institute of India. The samples were stored in 70% ethanol at room temperature. Total genomic DNA was extracted using the DNeasy Blood Tissue Kit (QIAGEN, Germany) in a final elution volume of 100 µl. The extracted DNA was visualized on 0.8% agarose gel and diluted in a final concentration of 30ng/µl for the Polymerase Chain Reaction (PCR) amplification.

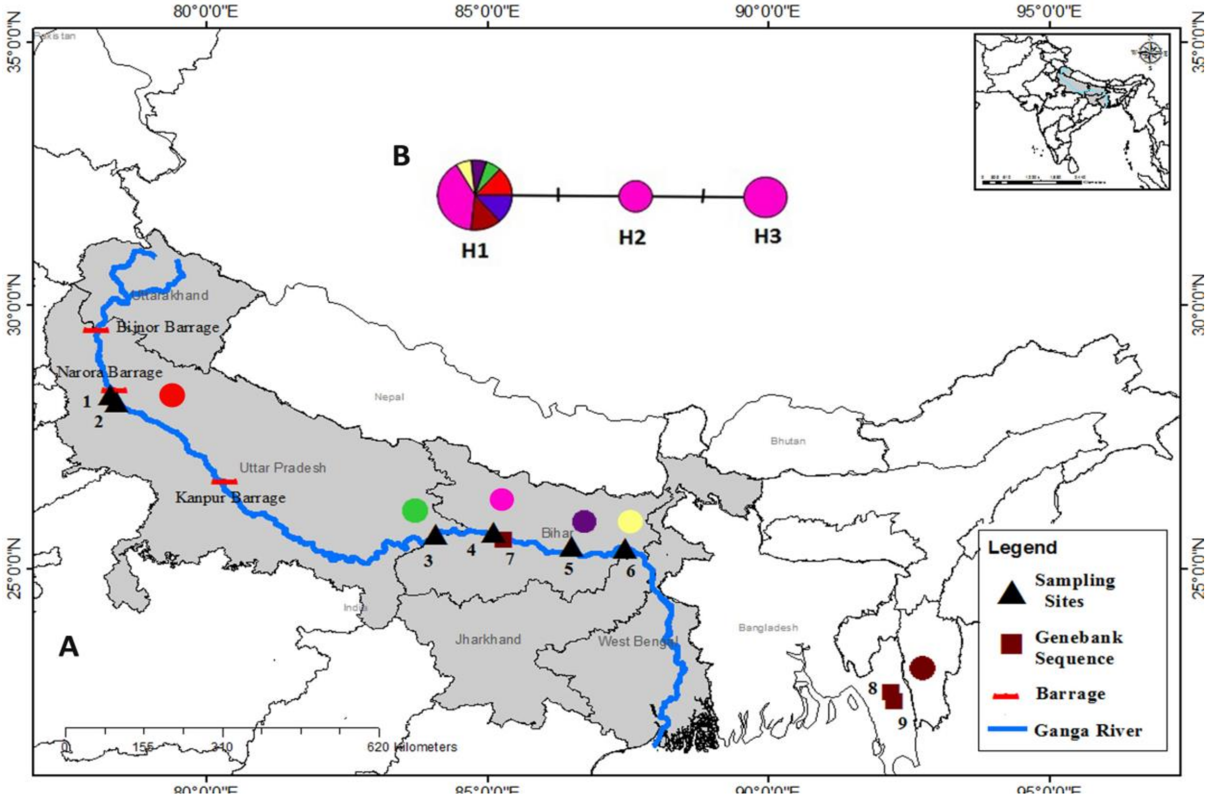


Figure 51: A. The geographic origin of samples used in this study. B. mtDNA control region based haplotype sharing; color of each circle represents the geographic origin.

8.2.1. PCR amplification and sequencing

PCR amplification of the entire mitogenome was performed using the different sets of primers (Hassanin et al., 2009). PCR amplifications were carried out in 20µl reaction volumes using 1 × PCR buffer (10 mM Tris–HCl, pH 8.3, and 50 mM KCl), 1.5 mM $MgCl_2$, 0.2 mM of each dNTPs, 3 pmol of each primer, 0.5 units of AmpliTaq Gold DNA Polymerase (Applied Biosystems) and 1 µl (~ 30 ng) of template DNA. The following PCR conditions were used: initial denaturation at 95 °C for 10 min, followed by 35 cycles of denaturation at 95 °C for 45 sec, annealing at 55 °C for 40 sec and extension at 72 °C for 75 sec. The final extension was performed at 72 °C for 10 min. The amplified products were visualized in a UV transilluminator on using 2% agarose gel electrophoresis. The amplified PCR products were treated with exonuclease-*I* and shrimp alkaline phosphatase (Thermo Scientific Inc.) at 37 °C for 20 min for the removal of any residual primer and dNTPs and followed by inactivation of enzymes at 85 °C for 15 min. The purified fragments were sequenced using both forward and reverse primers in an Applied Biosystems Genetic Analyzer 3500 XL using BigDye v3.1 Kit.

8.3 Data analysis

8.3.1 Mitogenome alignment and annotation

The overlapping fragments of complete mitogenome sequences were assembled and analyzed using Sequencher®version 5.4.6 (Gene Codes Corporation, Ann Arbor, MI, USA). Mitochondrial DNA annotation was done using Mitos WebServer (Bernt et al. 2013) and again confirmed with MitoFish (Iwasaki et al. 2013). We obtained a total length of 16319 bp, and the complete mtDNA gene map of the Gangetic dolphin using CGView Server (http://stothard.afns.ualberta.ca/cgview_server/) (Grant and Stothard 2008). Base compositions and mtDNA genetic code were calculated in MEGA X (Kumar et al. 2018). For estimating the bias in nucleotide composition among the complete mitogenome, protein-coding genes (PCGs), transfer RNA (tRNA), ribosomal RNA (rRNA) and control region (CR), skew analysis was carried out using the following method: AT skew = (A−T)/(A+T), GC skew = (G−C)/(G+C) (Perna and Kocher 1995). We manually estimated the intergenic spacer and overlapping regions between genes of complete mitogenome. We predicted the typical secondary cloverleaf structure of transfer RNA genes (tRNAs) of using tRNAscan-SE server 2.0 (http://lowelab.ucsc.edu/tRNAscan-SE/) (Lowe and Chan 2016) using the vertebrate mitochondrial genetic code under the default mode. The Open Reading Frame Finder (https://www.ncbi.nlm.nih.gov/orffinder/) online web tool was used to check the start and stop codons of PCGs. The comparative analysis of relative synonymous codon usage (RSCU), the relative abundance of amino acids, and codons distribution were calculated using MEGA X (Kumar et al. 2018). The tandem repeats in the control regions (CR) were predicted by the online Tandem Repeats Finder web tool (https://tandem.bu.edu/trf/trf.html) (Benson 1999).

8.3.2 Genetic diversity and demographic analysis

The genetic diversity of the Gangetic dolphin was assessed using the mtDNA control region (CR). In addition to six sequences of Gangetic dolphin generated in the present study, we also included 17 additional GeneBank sequences that consisted of 11 from Patna (Bihar), 2 from Bangladesh, and 4 from unknown localities (Table 24). DnaSP 5.0 (Librado and Rozas 2009) was used to analyze the haplotype diversity (h), nucleotide diversity (p), and polymorphic sites (s).

Table 24: Gangetic dolphin (*Platanista gangetica*) samples collected and mtDNA control region sequences obtained from GenBank that were included in the analyses

Haplotype	Genebank ID	Localities	Reference
Hap 1	MN896020	S1 Narora,Uttar Pradesh	This study
	MN896021	S2 Narora, Uttar Pradesh	This study
	MN896022	S3 Balia, Uttar Pradesh	This study
	MN896023	S4 Patna, Bihar	This study
	MT668915	S5 Suphaul Bihar	This study
	MT668916	S6 Araria Bihar	This study
	AY102527	Pahleza Ghat Patna	Verma et al. 2004
	AY102528	Bangladesh Kaptai Lake	Verma et al. 2004
	AY102529	Digha Patna	Verma et al. 2004
	AY102530	Kurzi Patna	Verma et al. 2004
	AY102531	Durja Patna	Verma et al. 2004
	AY102532	Bangladesh	Verma et al. 2004
	AY102533	Ghagha Ghat Patna	Verma et al. 2004
	KJ629311	Unknown	Braulik et al. 2014
	NC045383	Unknown	Jadhav et al. 2010
Hap 2	AY102534	Patna	Verma et al. 2004
	AY102535	Kurzi Patna	Verma et al. 2004
	KJ629313	Unknown	Braulik et al. 2014
Hap 3	AY102536	Mainpura Digha patna	Verma et al. 2004
	AY102537	Patna	Verma et al. 2004
	AY102538	Patna	Verma et al. 2004
	AY102539	Kasmar Pahleza Patna	Verma et al. 2004
	KJ629312	Unknown	Braulik et al. 2014

The demography history of the Gangetic dolphin was performed by different statistical approaches such as Tajima's *D* (Tajima 1989) and Fu's F*s* test (Fu 1997). Mismatch distribution analysis, a sum of squared deviations (SSD), and the raggedness index (r) were also used to demonstrate the pattern of population stability or population expansion in the past under the sudden demographic expansion models using the ARLEQUIN v3.5 program (Excoffier and Lischer 2010). The *P*-values were obtained from 1000 simulations based on a selective neutrality test. The spatial distribution of the Gangetic dolphin haplotypes was visualized using a median-joining network, which was created using the PopART software package (Leigh and Bryant 2015).

8.4 Results and discussion

8.4.1 Mitogenome feature and organization

We obtained a 16319 bp complete circular mitochondrial genome of the Gangetic dolphin (Figure 52) and deposited in GenBank (acc. no. MN896020-MN896023; MT668915, MT668916). The mitogenome encoded by two rRNA genes, 22 tRNA genes, and one origin of replication (O_L), 13 PCGs, and a non-coding CR. Among these, 28 genes (2rRNA, 14 tRNA and 12 PCGs) were located on the heavy strand (H-strand), except *ND6* gene and eight tRNA genes (*tRNA*^{Gln}, *tRNA*^{Ala}, *tRNA*^{Asn}, *tRNA*^{Cys}, *tRNA*^{Tyr}, *tRNA*^{Ser}, *tRNA*^{Glu}, *tRNA*^{Pro}) and O_L located on the light strand (L -strand). The distribution and arrangement of mitogenes were similar to the other mammalian species (Boore 1999; Singh et al. 2019; Kumar et al. 2019). The nine pairs of the overlapping region in mitogenome were observed among *tRNA*^{Ile}/*tRNA*^{Gln}, *ND2/tRNA*^{Trp}, O_L/tRNA^{Cys}, *ATP8/ATP6*, *ATP6/COIII*, *COIII/tRNA*^{Gly}, *ND4L/ND4*, *ND5/ND6* and *tRNA*^{Thr}/*tRNA*^{Pro}. The overlapping regions ranged from -1 to -39 bp. The longest overlapping was located between ATP8 and *ATP6* (39bp), whereas it was smallest between *ATP6/COIII*, *COIII/tRNA*^{Gly,} and *tRNA*^{Thr}/*tRNA*^{Pro} (1 bp). The O_L sequence was 35 bp in length and was located between the *tRNA*^{Asn} and *tRNA*^{Cys}; this region played an important role in the initiation of DNA synthesis by recognizing the start site of DNA polymerase (Wong and Clayton 1985). The non-coding *CR* region was located between *tRNA*^{Pro} and *tRNA*^{Phe}. In addition, 14 intergenic spacers were observed between the mitochondrial regions ranging from 1 to 32 bp length; the longest intergenic was present between *tRNA*^{Asn} and *tRNA*^{Cys} (Table 25).

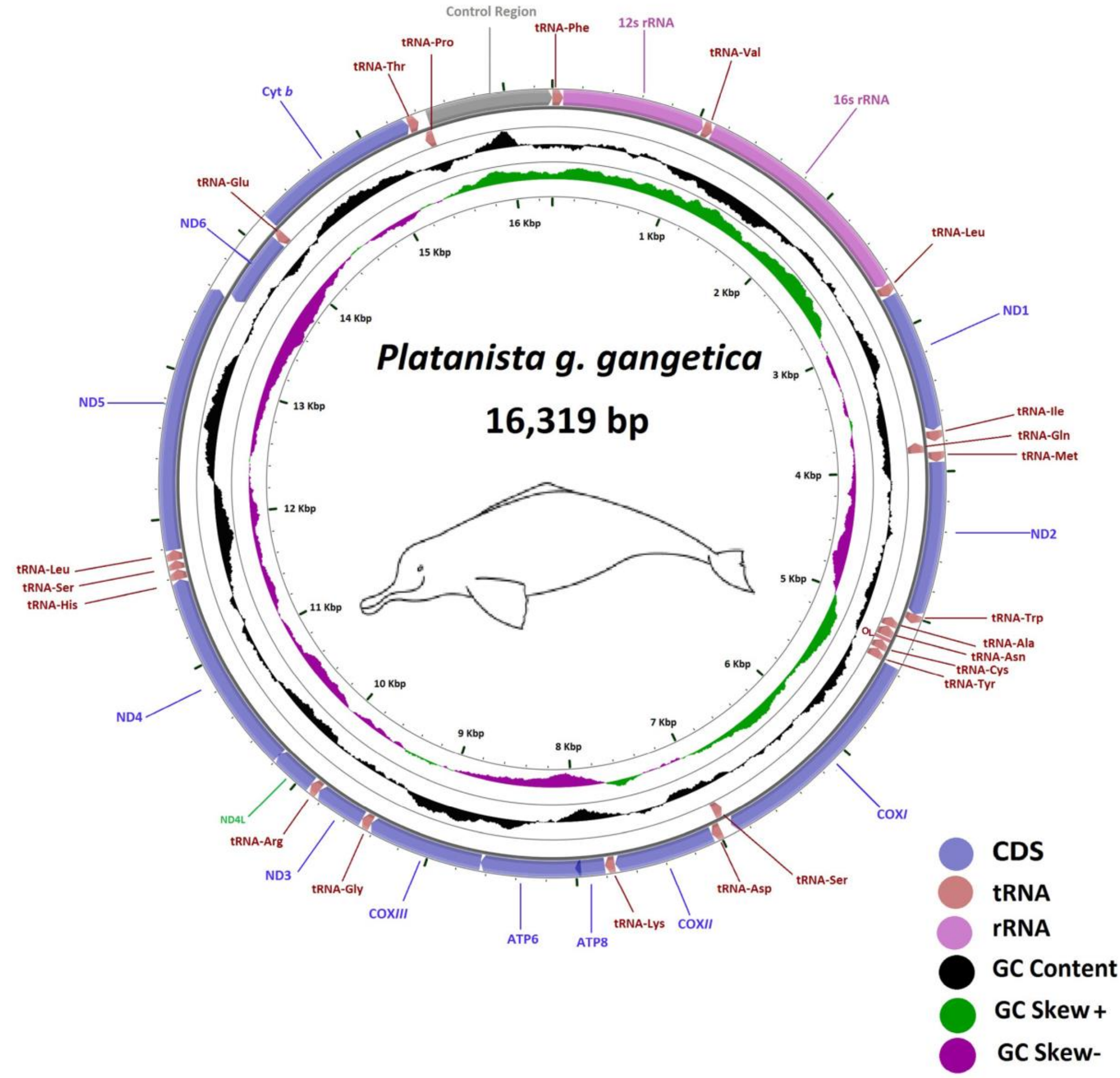


Figure 52: The complete mitochondrial genome organization of the Gangetic dolphin (*P. g, gangetica*). 12S rRNA and 16S rRNA refer to ribosomal RNAs; *COXI*, *COXII* and *COXIII* refer to subunits of cytochrome c oxidase; *Cyt b* denotes to cytochrome b; *ND1-ND6* refer to components of NADH dehydrogenase; *ATPase 6* and *ATPase 8* refers to classes of ATP synthase; Transfer RNAs (tRNA) are labeled with their corresponding amino acids; CDS represent the coding region and are shown in blue.

Table 25: The organization and characterization of the complete mitochondrial genome of the Gangetic dolphin (*P. g. gangetica*).

Gene	Start position	End position	Length (bp)	Initiation codons	Termination codons	Anti-codon	Strand	Space/overlap
tRNA-Phe	1	73	73	-	-	GAA	H	0
12S ribosomal RNA	74	1044	971	-	-	-	H	0
tRNA-Val	1045	1111	67	-	-	TAC	H	0
16S ribosomal RNA	1112	2688	1577	-	-	-	H	0
tRNA-Leu	2689	2763	75	-	-	TAA	H	0
ND1	2766	3722	957	ATG	TAA	-	H	+2
tRNA-Ile	3728	3796	69	-	-	GAT	H	+5
tRNA-Gln	3794	3866	73	-	-	TTG	L	-3
tRNA-Met	3868	3936	69	-	-	CAT	H	+1
ND2	3937	4980	1044	ATA	TAG	-	H	0
tRNA-Trp	4979	5046	68	-	-	TCA	H	-2
tRNA-Ala	5052	5120	69	-	-	TGC	L	+5
tRNA-Asn	5122	5195	74	-	-	GTT	L	+1
rep origin-O_L	5196	5230	35	-	-	-	L	0
tRNA-Cys	5228	5294	67	-	-	GCA	L	-3
tRNA-Tyr	5295	5360	66	-	-	GTA	L	0
COX*I*	5362	6906	1545	ATG	TAA	-	H	+1
tRNA-Ser	6908	6976	69	-	-	TGA	L	+1
tRNA-Asp	6984	7051	68	-	-	GTC	H	+7
COX*II*	7052	7735	684	ATG	TAA	-	H	0
tRNA-Lys	7739	7805	68	-	-	TTT	H	+3
ATP8	7807	8007	201	ATG	TAG	-	H	+4
ATP6	7968	8648	681	ATG	TAA	-	H	-39
COX*III*	8648	9433	786	ATG	TAG	-	H	-1
tRNA-Gly	9433	9501	69	-	-	TCC	H	-1
ND3	9502	9848	346	ATA	TA-	-	H	0
tRNA-Arg	9849	9918	70	-	-	TCG	H	+1
ND4L	9919	10215	297	GTG	TAA	-	H	0
ND4	10209	11586	1378	ATG	T--	-	H	-7
tRNA-His	11587	11655	69	-	-	GTG	H	0
tRNA-Ser	11656	11715	60	-	-	GCT	H	0
tRNA-Leu	11717	11786	70	-	-	TAG	H	+1
ND5	11787	13607	1821	ATA	T--	-	H	0
ND6	13593	14112	520	ATG	T--	-	L	-14
tRNA-Glu	14113	14181	69	-	-	TTC	L	0
CYT *b*	14186	15325	1140	ATG	AGA	-	H	+4
tRNA-Thr	15326	15396	71	-	-	TGT	H	0
tRNA-Pro	15396	15462	67	-	-	TGG	L	-1
Control region	15463	16319	857	-	-	-	H	0

8.4.2 Base composition and AT/GC-skew

The total nucleotide composition of the Gangetic dolphin mitochondrial genome was A (32.9%), T (26.7%), C (27.9%), and G (12.4%), and it was biased toward A+T (59.6%). The total length of all 13 PCGs was 11400 bp; however, 1519 bp for tRNAs, and 2548 bp for rRNA genes. The A+T composition of PCGs, tRNAs, rRNAs and CR were 58.9%, 63.5%, 60.7% and 60.5% respectively (Table 26). The AT skewness was 0.104, and GC skewness was −0.384 in the complete mitogenome of the Gangetic dolphin. The AT-skew was positive for PCGs, tRNAs, and rRNAs, whereas GC-skew was negative except tRNAs in Gangetic dolphin. We estimated base skews between all the freshwater river dolphins for understanding the nucleotide distribution in PCGs (Table 27). The trend of AT-skew and GC-skew values in all 13 PCGs of the Gangetic dolphin is shown in Figure 53. The average AT and GC skews values for the Gangetic dolphin in PCGs were 0.045 and -0.412, respectively. In all studied river dolphins, AT skewness was positive, indicating that the adenines base occurs more frequently than thymine, whereas GC skewness was negative, indicating that the cytosine base is more common than guanine base in PCGs. Relative synonymous codon usage (RSCU) for the 13 protein-coding genes of the Gangetic dolphin revealed that Leucine was the most frequent amino acid, followed by Isoleucine and Threonine, whereas Methionine and Cysteine were less abundant (Figure 54). The total length of tRNA was 1519 bp, overall A+T and G+C content was 63.5% and 36.4%, respectively. The average AT and GC skews values for tRNAs were 0.048 and 0.439, receptively. All the tRNA genes folded into a classic secondary cloverleaf structure, except $tRNA^{ser(GCT)}$ in which dihydrouridine 'DHU' arm did not form a stable structure (Figure 55). The 'DHU' arm of $tRNA^{ser}$ was a large loop instead of a conserved stem-loop structure.

The length of CR in the Gangetic dolphin is 857 bp and was located between $tRNA^{Pro}$ and $tRNA^{Phe}$, which was less than the majority of cetaceans. The AT content was higher than the GC content. In particular, 18 bp repeat consensus (AATACTAATAACAAAAAC) was found within 14062-14101 bp with a copy number 2.2 of CR. However, no repeat sequence was detected in the mitogenome of *P. g. minor*, *Inia geoffrensis*, *Pontoporia blainvillei,* and *Lipotes vexillifer*. Duplicated CRs have also been found in the mitogenomes of deer's (Gupta et al. 2015; Kumar et al. 2017), birds (Abbott et al. 2005), and fishes (Shi et al. 2014).

Table 26: Nucleotide composition indices in different regions of mitogenomes of the river dolphins

Species	Accession number	Whole mitogenome		Protein coding genes (PCGs)		Ribosomal RNA (rrnl)	
		Length (bp)	AT%	Length (bp)	AT%	Length (bp)	AT%
P.g.minor	NC_005275	16324	59.7	11406	59.1	2548	60.5
Inia geoffrensis	NC_005276	16588	58.7	11406	57.9	2554	59.5
Pontoporia blainvillei	NC_005277	16593	59.5	11401	58.6	2560	60.8
Lipotes vexillifer	NC_007629	16392	60.8	11406	60.2	2554	60.8

Table 27: Nucleotide composition and skewness in the Gangetic dolphin (*P. g. gangetica*) mitochondrial genome

Platanista gangetica gangetica	Size (bp)	A%	T%	AT-skew	G%	C%	GC-skew
Whole mitogenome	16319	32.9	26.7	0.104	12.4	27.9	−0.384
PCGs	11400	30.8	28.1	0.045	12.1	29.1	−0.412
tRNAs	1519	33.3	30.2	0.048	19.0	17.4	0.439
rRNAs	2548	37.2	23.5	0.225	17.0	22.3	-0.134

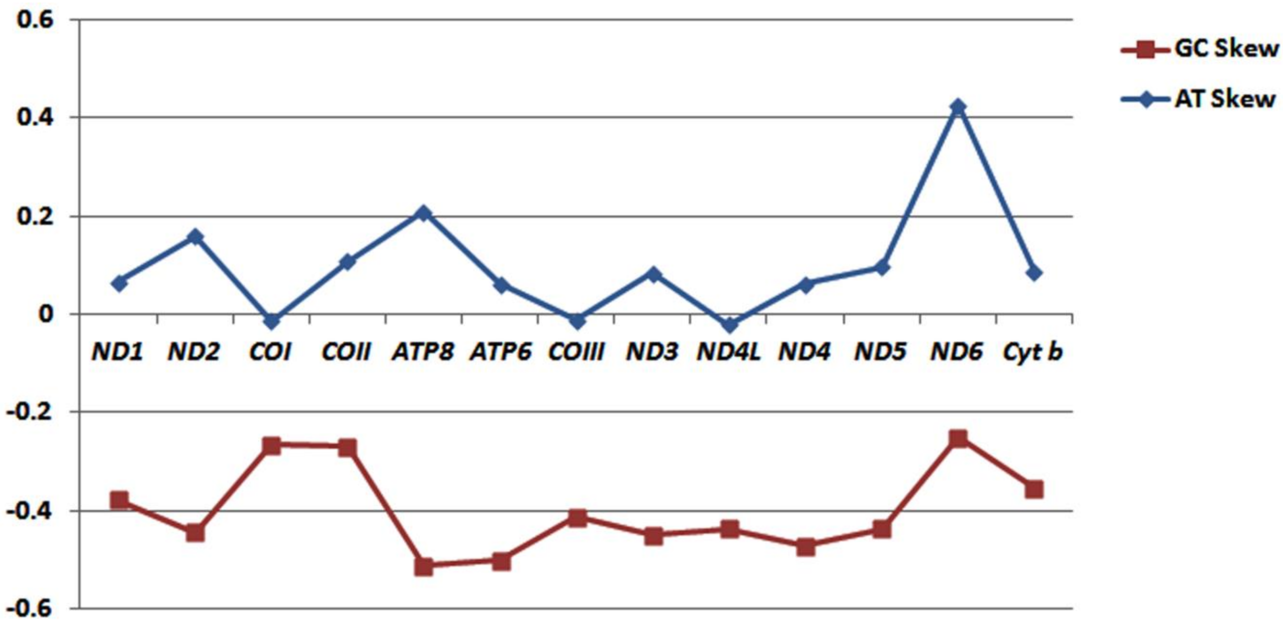


Figure 53: Graphical representation of AT- and GC-skew in all the 13 protein-coding genes of the Gangetic dolphin (*P. g. gangetica*) mitogenome. Values of AT- and GC-skew are plotted on X-axis.

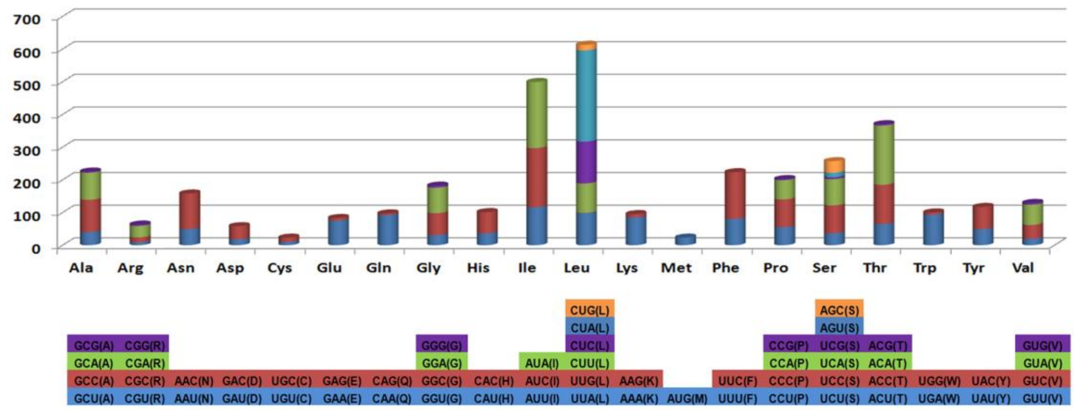


Figure 54: Relative synonymous codon usage (RSCU) of the mitochondrial protein-coding genes of the Gangetic River dolphin. Codon count numbers are provided on the X-axis.

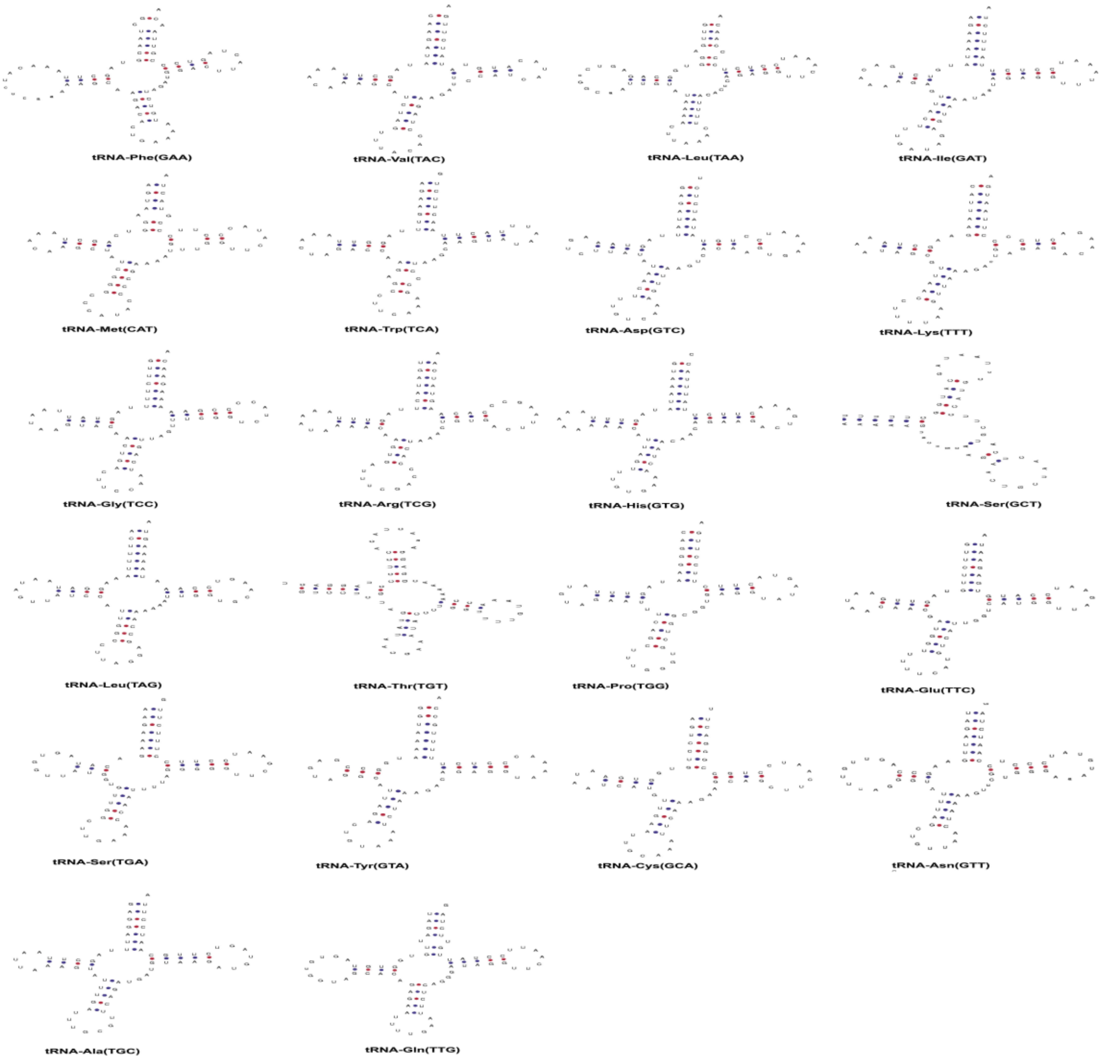


Figure 55: Secondary structures of the 22 tRNA genes of the Gangetic dolphin (*P. g. gangetica*) mitogenome

8.4.3 Genetic diversity, geospatial genetic structuring and demography

We used 23 sequences of mtDNA CR region (857 bp) for the estimation of mtDNA genetic diversity in the Ganga River covering a wide geographic range, i.e., Narora (Uttar Pradesh) in India to Bangladesh (Table 24). Among the 23 CR sequences, we observed only two parsimony informative sites at nucleotide position 15574 and 16136 (corresponding to complete mitogenome) of the Gangetic dolphin (GenBank no. MN896020). These sequences were grouped into 3 haplotypes (hap), where hap1 was core; shared majority sequences from the entire location site from Narora to Bangladesh (Figure 51B). Hap2 and Hap3 consisted of 3 and 5 sequences of Patna (Bihar). The haplotypes (*hd*) and nucleotide diversity (π) of the Gangetic dolphin was *hd*=0.534 +/- 0.097 and π=0.00097 +/- 0.00018. The population of the Gangetic dolphin showed a low level of genetic variations, and the majority of sequences from the Ganga to Karnaphuli river in eastern Bangladesh were clustered in a single haplotype, which indicated a high genetic homogeneity across their distribution ranges. The low genetic diversity may be a result of the founder effect due to historically low population size (Braulik et al., 2014). Low genetic variability indicated that the population of the Gangetic dolphin might have historically suffered from an inbreeding, which could have affected the fitness and local adaptation of species. Remarkably low genetic diversity was observed in common bottleneck dolphins in the southwestern Atlantic Ocean (Fruet et al. 2014); bottlenose dolphins in Panama (Barragán-Barrera et al. 2017), bottlenose dolphins in Western North Pacific (Chen et al. 2017), Australian humpback dolphins (Parra et al. 2018).

We performed Tajima's *D* and Fu's F*s* tests to infer the demographic pattern of the Gangetic dolphin. The non-significant (P>0.10) positive values for Tajima's *D* =1.187 and Fu's F*s* =0.926 were observed. Fu and Li's d and f tests also indicated no significant departure from neutrality (P > 0.10). The sum of square deviation (SSD) and raggedness index (Rd) values under the demographic expansion model was 0.0207; P=0.39 and 0.14395; P=0.54, respectively. The mismatch distribution plot for the Gangetic dolphin was bimodal shaped. The results of the neutrality test and mismatch distribution supported the hypothesis of the demographic stability of Gangetic dolphin populations. It might be because the species survive in restricted habitat with immense human pressure across its range.

8.4.4 Genetic differentiation, and phylogenetic analysis

The mean pairwise genetic distance based on Kimura's 2-parameter model was estimated for 2 rRNA, 13 PCGs, CR, and complete mitogenome between Platanistidae and ten families of cetaceans (Table 28). We found that Platanistidae was genetically closer to the Mysticeti, followed by Ziphiidae. However, few genes such as 12s and *ND1* showed an equal genetic distance from Platanistidae to Kogiidae/Mysticeti (0.126) and Ziphiidae/Mysticeti (0.202), respectively. The *COXIII* showed a low genetic distance with Delphinidae, *ND2* showed low value with Ziphiidae, whereas *ND4l* and *CR* showed a close affinity with Monodontidae. The genetic differentiation results revealed that the genus Platanistidae is close to the Baleen and Beaked whales' group than the other river dolphin, namely *Lipotes*, *Inia,* and *Pontoporia*. The similar genetic character of Platanistidae with whales indicated the South Asian River dolphin followed a different evolutionary line. In this light, the choice of short fragment-based approach for molecular studies within and among cetaceans would affect the accuracy and reliability of the genetic analysis. Hence, long coverage of mtDNA or multigene alignment provides better insight into phylogenetic information than short or single-gene analysis and might give a more accurate taxonomic status (Kumar et al. 2019).

The phylogenetic analysis was performed using ML, and BI trees approach 12 PCGs among 12 families of cetaceans and found similar tree topology (Figure. 56). The ML tree showed that Platanistidae was clustered with the clade of Ziphiidae with bootstrap value (BP) 68%. Within cetaceans topology was (Delphinidae+(Phocoenidae+Monodontidae) + (Lipotidae + (Iniidae+ Pontoporiidae) and (Ziphiidae+Platanistidae)+(Kogiidae+ Physeteridae)+Mysticeti). The two subspecies of *P. gangetica* and *minor* were clustered together with high BP 100%.

Our analysis indicated that river dolphins are polyphyletic, and the placement of *Platanista* is to be more basal and had no affinity with other river dolphins. Based on this, it is clear that *Platanista* had a diverse lineage that has been widely accepted. Based on the partial mitochondrial Cyt *b* and nuclear interphotoreceptor retinoid-binding protein (IRBP) gene sequences, a sister genetic relationship between Platanistidae and Mysticeti was suggested (Verma et al. 2004). This result was consistent with our mean pairwise genetic distance data, where comparatively low distance values were observed between Platanista/Mysticeti than any other group of cetaceans. Since the method used by Verma et al. (2004) was a conventional neighbor-joining analysis, which was a clustering algorithm that clusters groups based on genetic distance. Therefore, to construct the phylogenetic for such a highly diverse species, model-based approaches, ML and BI, and choosing the best fit model provide better insight. The high fragility in the phylogenetic position of *Platanista* within cetaceans may partly because the South Asian River dolphin, sperm, beaked, and baleen whale lineages seems to have been produced through very rapid spit event in the Oligocene-Miocene (Cassens et al. 2000; Nikaido et al. 2001). The previous studies using only cytochrome *b* sequence suggested the split between Indus river dolphin and Gangetic dolphin was around 0.55 MY ago (95% HPD=0.13-1.05) (Braulik et al. 2014), and 0.51 MY ago (95% HPD=0.14-1.02) (McGowen et al. 2009). This split age was possibly associated with the uplift of the Himalayas Mountain driven by tectonic forces, which in turn might have altered the drainage pattern of the western Himalayan Mountain system in the recent geologic past (Clift and Blusztajn

2005). Based on these events, we hypothesize that after the loss of potential connectivity between these rivers, the South Asian River dolphins later colonized and then genetically isolated at around 1.03 MYA (95% HPD = 0.84–1.18).

Table 28: Mean pairwise genetic distance based on Kimura's 2-parameter model, calculated with reference to the Platanistidae.

Gene	Delphinidae	Ziphiidae	Iniidae	Kogiidae	Lipotidae	Monodontidae	Phocoenidae	Pontoporiidae	Physeteridae	Mysticeti
12s rRNA	0.148	0.131	0.176	**0.126**	0.148	0.154	0.169	0.176	0.135	**0.126**
16s rRNA	0.105	0.108	0.130	0.113	0.109	0.109	0.108	0.125	0.131	**0.097**
COX*I*	0.195	0.171	0.209	0.220	0.196	0.207	0.204	0.203	0.206	**0.170**
COX*II*	0.224	0.197	0.243	0.243	0.230	0.209	0.214	0.188	**0.171**	0.177
COX*III*	**0.193**	0.220	0.215	0.240	0.210	0.215	0.215	0.234	0.219	0.207
ND1	0.213	**0.202**	0.241	0.226	0.208	0.210	0.228	0.225	0.224	**0.202**
ND2	0.237	**0.209**	0.281	0.254	0.258	0.218	0.258	0.306	0.250	0.221
ND3	0.253	0.249	0.259	0.282	0.258	0.253	0.244	0.334	0.250	**0.224**
ND4	0.240	0.222	0.244	0.256	0.239	0.235	0.252	0.272	0.221	**0.219**
ND4l	0.321	0.315	0.329	0.343	0.322	**0.298**	0.346	0.329	0.315	0.347
ND5	0.271	0.259	0.294	0.265	0.278	0.273	0.280	0.317	0.264	**0.222**
ND6	0.274	0.255	0.378	0.289	0.355	0.265	0.298	0.351	0.305	**0.245**
Cytb	0.206	0.199	0.228	0.204	0.261	0.198	0.232	0.265	0.215	**0.175**
CR	0.292	0.288	0.393	0.432	0.301	**0.275**	0.288	0.430	0.620	0.307
Overall	**0.227**	**0.216**	**0.259**	**0.249**	**0.241**	**0.223**	**0.238**	**0.268**	**0.252**	**0.210**

Bold, highlighted data indicates minimum genetic distance

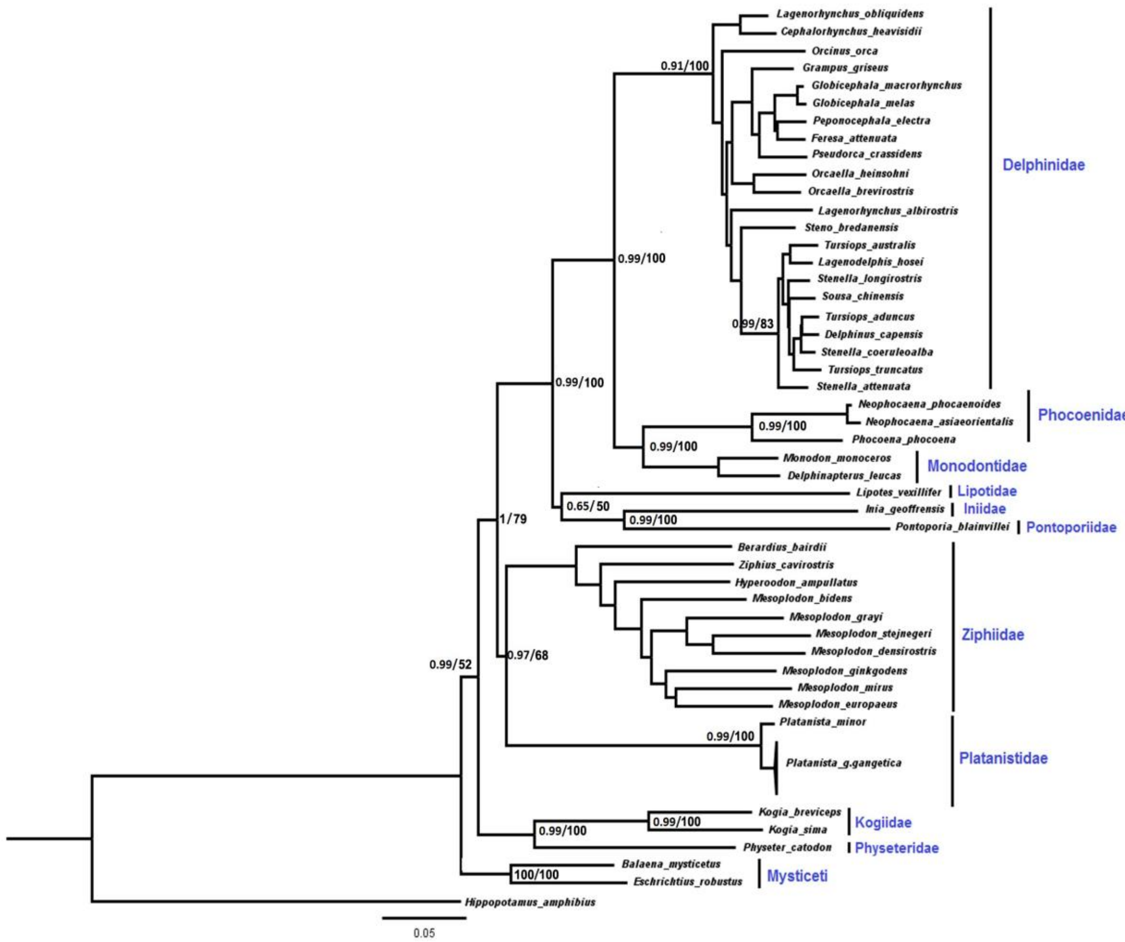


Figure 56: Phylogenetic tree inferred from maximum likelihood (ML) (raxmlGUI 2.0 beta) and Bayesian inference (BI) analyses using 12 concatenated protein-encoding genes. Bootstrap support and Bayesian posterior probability values (≥ 50% BS) are shown at the node of the tree. *Hippopotamus amphibious* (AJ010957) was used as an outgroup.

8.5 Conclusion

Endangered species are typically categorized by low effective population size, which leads to a low level of genetic diversity that may result in inbreeding. The present study indicated a low genetic diversity at the mitochondrial genome with high homogeneity across their distribution range (India and Bangladesh). Although for small and isolated populations, there is a perceived link between a loss of genetic diversity and a high risk of extinction (Gilpin and Soule 1986), low genetic variability in species is not the major challenges it faces, since there is tremendous anthropogenic pressure on its habitat, which affect the population growth. The phylogenetic position of this species holds importance as it is considered to be the evolutionary link with whales group rather than other existing river dolphins (*I. geoffrensis*, *P. blainvillei,* and *L. vexillifer*). Thus, it is a pioneer comparative study of molecular characterization of the complete mitogenome of Gangetic dolphin, which was 16319 bp in length. The conservation and management of the closely related subspecies are challenging to practice due to minimal physical variation. This study provides the baseline information for adaptive evolution, spatial distribution, and evolutionary relationships of *Platanista*. Taking consideration into low genetic diversity and genetic bottleneck, continuous population monitoring and strong conservation efforts are needed to be focus to restrict the mortality rate that causes due to illegal human activities.

8.6 References

- Abbott CL, Double MC, Trueman JW, Robinson A, Cockburn A (2005) An unusual source of apparent mitochondrial heteroplasmy: duplicate mitochondrial control regions in *Thalassarche albatrosses* Mol Ecol 14:3605-3613.
- Árnason Ú, Gullberg A (1996) Cytochrome *b* nucleotide sequences and the identification of five primary lineages of extant cetaceans Mol Biol Evol 13:407-417.
- Avise JC (1989) A role for molecular genetics in the recognition and conservation of endangered species. Trends Ecol Evol 4:279-281.
- Barragán-Barrera DC, May-Collado LJ, Tezanos-Pinto G, Islas-Villanueva V, Correa-Cardenas CA, Caballero S (2017) High genetic structure and low mitochondrial diversity in bottlenose dolphins of the Archipelago of Bocas del Toro, Panama: A population at risk?. PloS one 12:e0189370.
- Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27**:**573-580.
- Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol 69:313-319.
- Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Res 27:1767-1780.
- Braulik GT, Arshad M, Noureen U, Northridge SP (2014) Habitat fragmentation and species extirpation in freshwater ecosystems; causes of range decline of the Indus River Dolphin (*Platanista gangetica minor*). *PloS one* **9**: e101657.
- Braulik GT, Barnett R, Odon V, Islas-Villanueva V, Hoelzel AR, Graves JA (2015) One species or two? Vicariance, lineage divergence and low mtDNA diversity in geographically isolated populations of South Asian River dolphin. J Mamm Evol 22:111-120.
- Braulik GT, Smith BD (2019) *Platanista gangetica* (amended version of 2017 assessment). *The IUCN Red List of Threatened Species.* e.T41758A151913336.
- Brender JR, Zhang Y (2015) Predicting the effect of mutations on protein-protein binding interactions through structure-based interface profiles. PLoS Comput Biol 11:e1004494.
- Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates: tempo and mode of evolution. J Mol Evol 18:225-239.
- Burton RS, Barreto FS (2012) A disproportionate role for mt DNA in D obzhansky–M uller incompatibilities?. Mol Ecol 21:4942-4957.
- Caballero S, Duchêne S, Garavito MF, Slikas B, Baker CS (2015) Initial evidence for adaptive selection on the NADH subunit two of freshwater dolphins by analyses of mitochondrial genomes. PLoS One 10:e0123543.
- Cassens I, Vicario S, Waddell VG, Balchowsky H, Van Belle D, Ding W, Fan C, Mohan RL, Simões-Lopes PC, Bastida R, Meyer A (2000) Independent adaptation to riverine habitats allowed survival of ancient cetacean lineages. Proc Natl Acad Sci 97:11343-11347.
- Chen I, Nishida S, Yang WC, Isobe T, Tajima Y, Hoelzel AR (2017) Genetic diversity of bottlenose dolphin (*Tursiops sp*.) populations in the western North Pacific and the conservation implications. Mar Biol 164:202.
- Chen Z, Xu S, Zhou K, Yang G (2011) Whale phylogeny and rapid radiation events revealed using novel retroposed elements and their flanking sequences. BMC Evol Biol 11:314.
- Clift PD, Blusztajn J (2005) Reorganization of the western Himalayan river system after five million years ago. Nature 438:1001-1003.

- Cozzone AJ (2001) Proteins: Fundamental chemical properties. e LS.
- Da Fonseca RR, Johnson WE, O'Brien SJ, Ramos MJ, Antunes A (2008) The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics* **9**: 119.
- de Muizon C (1994) Are the squalodonts related to the platanistoids. . In Proceedings of the San Diego Soc Nat Hist 29:135-146.
- DeLano WL (2002) The PyMOL molecular graphics system. DeLano Scientific; San Carlos, CA, USA. http://www. pymol. org.
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Bio Evol 29:1969-1973.
- Edler D, Klein J, Antonelli A, Silvestro D (2019) raxmlGUI 2.0 beta: a graphical interface and toolkit for phylogenetic analyses using RAxML. BioRxiv 800912.
- Escalona T, Weadick CJ, Antunes A (2017) Adaptive patterns of mitogenome evolution are associated with the loss of shell scutes in turtles. Mol Biol Evol 34:2522-2536.
- Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564-567.
- Finch TM, Zhao N, Korkin D, Frederick, KH, Eggert LS (2014) Evidence of positive selection in mitochondrial complexes I and V of the African Elephant. PLoS One 9:e92587.
- Fruet PF, Secchi ER, Daura-Jorge F, Vermeulen E, Flores PA, Simoes-Lopes PC, Genoves RC, Laporta P, Di Tullio JC, Freitas TR, Dalla Rosa L (2014) Remarkably low genetic diversity and strong population structure in common bottlenose dolphins (*Tursiops truncatus*) from coastal waters of the Southwestern Atlantic Ocean. Conserv Genet 15:79-895.
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915-925.
- Gilpin ME, Soulé ME (1986) Minimum viable populations: processes of species extinction M.E. Soulé (Eds) *Conservation Biology:* The Science of Scarcity and Diversity Sinauer Associates Sunderland 19-34.
- Grant JR, Stothard P (2008) The CGView Server: a comparative genomics tool for circular genomes. Nucleic Acids Res 36:W181-W184.
- Guang Y, Kaiya Z (1999) A study on the molecular phylogeny of river dolphins. Shou lei xue bao= Acta Theriologica Sinica 19:1-9.
- Gupta SK, Kumar A, Gaur A, Hussain SA (2015) Detection of 40 bp insertion-deletion (INDEL) in mitochondrial control region among sambar (*Rusa unicolor*) populations in India. BMC Res Notes 8:1-7.
- Hamilton H, Caballero S, Collins AG, Brownell Jr RL (2001) Evolution of river dolphins. Pro R Soc Lond B 268:549-556.
- Hassanin A, Ropiquet A, Couloux A, Cruaud C (2009) Evolution of the mitochondrial genome in mammals living at high altitude: new insights from a study of the tribe Caprini (Bovidae, Antilopinae). J Mol Evol 68:293-310.
- Iwasaki W, Fukunaga T, Isagozawa R, Yamada K, Maeda Y, Satoh TP, Sado T, Mabuchi K, Takeshima H, Miya M, Nishida M (2013) MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. Mol Biol Evol 30:2531-2540.
- Jones S (1982) The present status of the Gangetic susu*, Platanista gangetica* (Roxburgh), with comments on the Indus susu, *P. minor* Owen. FAO Fisheries Series 5:97-115.

- Kumar A, Gautam KB, Singh B, Yadav P, Gopi GV, Gupta SK (2019) Sequencing and characterization of the complete mitochondrial genome of Mishmi takin (*Budorcas taxicolor taxicolor*) and comparison with the other Caprinae species. Int J Biol Macromol 137:87-94.
- Kumar A, Ghazi MG, Hussain SA, Bhatt D, Gupta SK (2017) Mitochondrial and nuclear DNA based genetic assessment indicated distinct variation and low genetic exchange among the three subspecies of swamp deer (*Rucervus duvaucelii*). Evol Biol 44:31-42.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547-1549.
- Leigh JW, Bryant D (2015) popart: full-feature software for haplotype network construction. Methods Ecol Evo 6:1110-1116.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452.
- Lowe TM, Chan PP (2016) tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44:W54-W57.
- McGowen MR, Spaulding M, Gatesy J (2009) Divergence date estimation and a comprehensive molecular tree of extant cetaceans. Mol Phylogen Evol 53:891-906.
- Messenger SL (1994) Phylogenetic relationships of platanistoid river dolphins (Odontoceti, Cetacea): assessing the significance of fossil taxa. In Proc San Diego Soc Nat Hist 29:125-133.
- Ngatia JN, Lan TM, Dinh TD, Zhang L, Ahmed AK, Xu YC (2019) Signals of positive selection in mitochondrial protein-coding genes of woolly mammoth: Adaptation to extreme environments?. Ecol Evol 9:6821-6832.
- Nikaido M, Matsuno F, Hamilton H, Brownell RL, Cao Y, Ding W, Zuoyan Z, Shedlock AM, Fordyce RE, Hasegawa M, Okada N (2001) Retroposon analysis of major cetacean lineages: the monophyly of toothed whales and the paraphyly of river dolphins. Proc Natl Acad Sci 98:7384-7389.
- Parra GJ, Cagnazzi D, Jedensjö M, Ackermann C, Frere C, Seddon J, Nikolic N, Krützen M (2018) Low genetic diversity, limited gene flow and widespread genetic bottleneck effects in a threatened dolphin species, the Australian humpback dolphin. Biol Conserv 220:192-200.
- Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol 41:353-358.
- Rambaut A (2012) FigTree v1. 4.0. A graphical viewer of phylogenetic trees. See http://tree. bio. ed. ac. uk/software. figtreetree.
- Rokas A, Carroll SB (2005) More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. Mol Biol Evol 22:1337-1344.
- Romero NA, Nicewicz DA (2016) Organic photoredox catalysis. Chemical reviews 116:10075-10166.
- Shi W, Miao XG, Kong XY (2014) A novel model of double replications and random loss accounts for rearrangements in the Mitogenome of *Samariscus latus* (Teleostei: Pleuronectiformes). BMC Genomics **15**: 352.
- Singh B, Kumar A, Uniyal VP, Gupta SK (2019) Complete mitochondrial genome of northern Indian red muntjac (*Muntiacus vaginalis*) and its phylogenetic analysis. Mol Biol Rep 46:1327-1333.
- Sinha RK, Kannan K (2014) Ganga River dolphin: an overview of biology, ecology, and conservation status in India. Ambio 43:1029-1046.

- Sinha RK, Sharma G (2003) Current status of the Ganga river dolphin, *Platanista gangetica* in the rivers Kosi and Son, Bihar, India. J Bombay Nat Hist Soc 100:27-37.
- Sinha RK, Smith BD, Sharma G, Prasad G, Choudhury BC, Sapkota K, Sharma RK, Behera SK (2000) Status and distribution of the Ganges susu(*Platanista gangetica*) in the Ganges River system of India and Nepal. Occas Pap IUCN Species Survival Comm 23:54-61.
- Smith BD, Ahmed B, Ali ME, Braulik G (2001) Status of the Ganga river dolphin or shushuk *Platanista gangetica* in Kaptai Lake and the southern rivers of Bangladesh. Oryx 35:61-72.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.
- The World Bank Report No. STEP705 (2011) The National Ganga River Basin Project.
- Thewissen JG, Williams EM, Roe LJ, Hussain ST (2001) Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. Nature 413:277-281.
- Verma SK, Sinha RK, Singh L (2004) Phylogenetic position of *Platanista gangetica*: insights from the mitochondrial cytochrome *b* and nuclear interphotoreceptor retinoid-binding protein gene sequences. Mol Phylogenetics Evol 33:280-288.
- Wallace DC (2007) Why do we still have a maternally inherited mitochondrial DNA? Insights from evolutionary medicine. Annu Rev Biochem 76:781-821.
- Wong TW, Clayton DA (1985) In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis. Cell 42:951-958.
- Woolley S, Johnson J, Smith MJ, Crandall KA, McClellan DA (2003) TreeSAAP: selection on amino acid properties using phylogenetic trees. Bioinformatics 19:671-672.
- Xia X (2017) DAMBE6: new tools for microbial genomics, phylogenetics, and molecular evolution. J Hered 108:431-437.
- Xiong YE, Brandley MC, Xu S, Zhou K, Yang G (2009) Seven new dolphin mitochondrial genomes and a time-calibrated phylogeny of whales. BMC Evol Biol 9:20.
- Yan J, Zhou K, Yang G, (2005) Molecular phylogenetics of ‘river dolphins’ and the baiji mitochondrial genome. Mol Phylogenetics Evol 37:743-750.
- Yang G, Zhou K, Ren W, Ji G, Liu S, Bastida R, Rivero L (2002) Molecular systematics of river dolphins inferred from complete mitochondrial cytochrome *b* gene sequences. Mar Mammal Sci 18:20-29.
- Yang XG (2009) Bayesian inference of cetacean phylogeny based on mitochondrial genomes. Biologia 64:811-818.
- Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591.
- Yang Z, Bielawski JP (2000) Statistical methods for detecting molecular adaptation. Trends Ecol Evol 15:496-503.
- Yang Z, Wong WS, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol 22:1107-1118.

9. GENETIC CHARACTERIZATION OF MANGUR, *Clarias gariepinus*, IN THE SELECTED STRETCH OF RIVER YAMUNA

9.1 Background of the study

The river Yamuna, the largest tributary of the Ganga, holds a rich ecosystem of both local and exotic fish species. Unfortunately, increasing anthropogenic pressures in recent times have disrupted the habitat structures, posing a threat to indigenous fish species and facilitating the invasion of alien species, such as the African catfish (*Clarias gariepinus*), into the natural water (Tyor & Pahwa, 2017). The African catfish, widely distributed in freshwater habitats globally, a massive species weighing up to 60 kg and measuring 1.7 m in length, has established a prominent presence in various stretches of the river Yamuna (Khan et al., 2021). This species also known as "mangur" in India, belongs to the family Claridae and coexists with another variety, the Thai magur (Singh et al., 2018). The African catfish has been observed in major rivers like the Ganga, Yamuna, Sutlej, Godavari, Periyar, and lakes such as Vembanad (Kumar et al., 2011). Concerns arise from the rapid increase in the biomass of the African catfish, leading to a decline in native species within the river. This species, known for its adaptability to harsh environments, acts as a proficient scavenger, consuming slaughterhouse and fish wastes (Singh et al., 2012). Its uncontrollable feeding habits, preying on native fish and aquatic animals, raise alarms about its impact on the biodiversity of the River Yamuna.

The Delhi region of the river Yamuna faces significant pollution challenges from industrial, mining, agricultural, and wastewater activities (Kansal et al., 2007). The introduction of the voracious predator, *Clarias gariepinus*, outside its natural range exacerbates concerns about the sustainability of local fish diversity (Khan et al., 2021). Despite the Indian government's ban on the cultivation of *C. gariepinus*, its presence persists in natural water systems, including the River Yamuna. Moreover, the African catfish carries approximately 18 species of infectious bacteria, making it a bioindicator for monitoring aquatic pollution. Its large size, extensive life cycle, and relevance to human consumption underscore the need for urgent attention to assess its genetic status. As per the IUCN Red list category, this fish comes under the Least Concern category (International Union for Conservation of Nature).

Figure 57 : Image indicating the size and morphology of *Clarias gariepinus*

The safe human consumption of this fish imposes increasing concerns from polluted, freshwater impoundments (Wagenaar et al., 2015). Therefore, it is important to understanding the genetic status of differentiation between a population or subpopulation at urgent attention today. However, despite the elemental importance of studying naturally observed genetic variation and the availability of diversified methods of describing patterns of genetic variation (Pritchard et al., 2000; Gao et al., 2007; Jombart et al., 2009; Engelhardt & Stephens, 2010), still very few studies have tried to investigate the relative benefactress affecting genetic differentiation across a species range. Previous studies strongly suggest that genetic differentiation is strongly influenced by two processes: its isolation by distance and its differential adaptation locally (Wright 1931, 1943; Slatkin 1987; Nosil et al., 2005, 2008).

Controlling the alien species and designing innovative options for their management is only possible through clear understanding of its breeding cycle in such studies. Khedkar et al., 2014 suggested that *Clarias batrachus* also known as Magur, has declined from natural habitats in India during the last few decades. This fish is in high demand and is preferred by the Indian consumers and also has high market demand. As a result, the traders often substitute *C. batrachus* with a morphologically similar but supposedly banned exotic catfish, *C. gariepinus*, in India. This study uses various morphological comparisons confirmed by DNA barcoding analysis to examine the level of substitution of *C. batracus* by *C. gariepinus* in India. This study indicates that up to 99% (in most cases) of the market samples sold as Magur or *C. batrachus* were observed to be *C. gariepinus*. John et al.,2015 studied the Mitochondrial cytochrome C Oxidase I (COI) sequence variation among the other clariid fishes of India such as *C. batrachus*, *C.*

dussumieri and *C. gariepinus* and also their relationship with other representative clariids using molecular genetic techniques. It was observed that the morphological similarity of *C. dussumieri* and *C. batrachus* was not replicated in the genetic level studies. *C. dussumieri* was observed to be more close to African catfish *C. gariepinus* thus indicating the utility of COI phylogeny to identify the African-Asian relationships within the cat fish family. The molecular study results also suggested that *C. magur* and *C. batrachus* are genetically distinct from each other.

Understanding the genetic diversity is important for the persistence of a species in an environment (Lande, 1988). It is also observed that the loss of genetic diversity can be rapid in a small population, where genetic drift can destroy the genetic variation in a few generations (Frankharm et al., 2002). Hence, a clear understanding on the genetic diversity of this fish will help in further analyzing its negative impacts on other native fishes in the river. Further studies must be implied to understand the false effects of these fishes on the native environment. To contribute towards a better understanding of the invasion of this species, this study will help in estimation of genetic variations and the phylogenetic status of *Clarias gariepinus* distributed among the selected stretch of River Yamuna.

9.1.1 Morphological differentiation

Clarias batrachus: Also referred to as the walking catfish, this species has an elongated body structure. Adult specimens typically attain a length of approximately 30-40 cm (12-16 inches). The head, which is relatively large in proportion to the body, can be measured from the tip of the snout to the posterior edge of the operculum. The eye's diameter is determined as the distance between the anterior and posterior edges of the eyeball. The dorsal fin extends from the head to the tail along the back. Notably, the pectoral fins play a crucial role in locomotion, particularly when the catfish exhibits its unique ability to "walk" on land.

Clarias gariepinus: Commonly recognized as the African sharp-tooth catfish, this species stands out as a relatively large catfish, capable of reaching impressive lengths of up to 1.5 meters (5 feet). Its head, notable for its relatively large size compared to the body, is harmonized by large eyes that maintain a proportional relationship with the overall body structure. The dorsal fin, spanning from the head to the tail along the back, adds to the distinctive features of this catfish. Meanwhile, the pectoral fins serve a dual purpose, contributing to both locomotion and stability. On the ventral side, the anal fin plays a crucial role in providing stability and facilitating precise maneuvering

9.2 Methodology

A comprehensive study involved the collection of a total of 26 tissue samples from local fisherman shops strategically located along the banks of the river Yamuna at Kalindi Kunj, New Delhi, Okhla, and ITO. To ensure genetic analysis, samples were meticulously obtained from the fin parts of the fish. Subsequently, all collected samples were carefully preserved in 70% ethanol within 2.0 Eppendorf tubes, creating an optimal environment for DNA extraction and subsequent analyses. Total genomic DNA was extracted from the tissue samples using the phenol-chloroform extraction protocols with a final elution volume of 100 µl. The extracted DNA was checked on 0.8% agarose gel and diluted in a final concentration of 50ng/µl for PCR amplification.

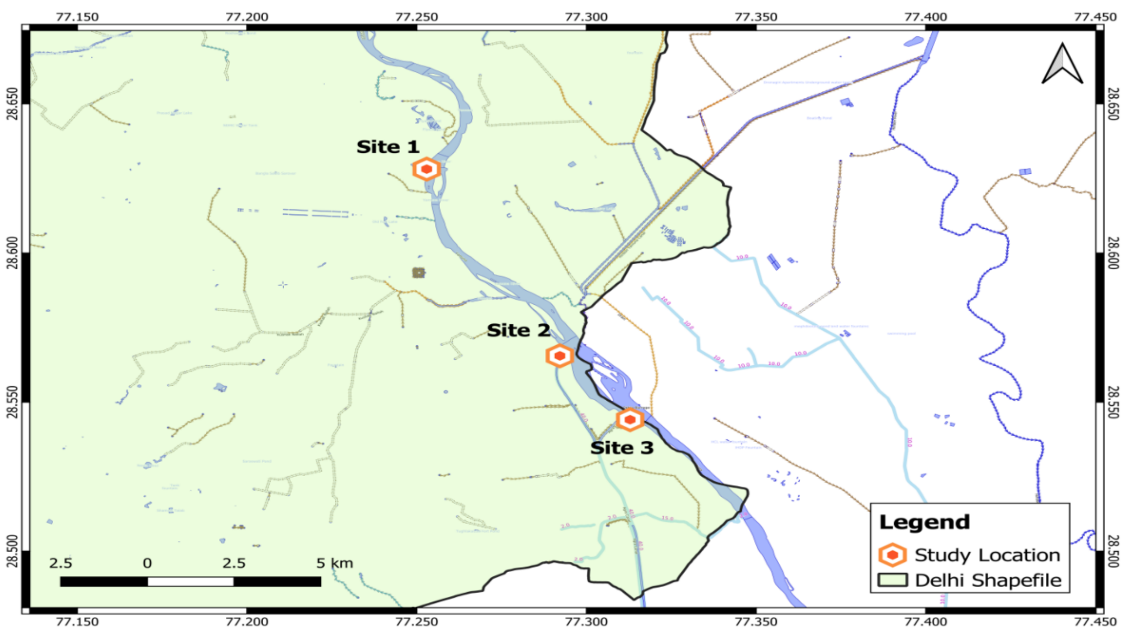


Figure 58: Map indicating the sampling site in Yamuna river.

9.2.1 mtDNA amplification and sequencing

Polymerase Chain Reaction (PCR) was performed for all collected samples to amplify the mtDNA Cytochrome b gene (Cyt *b*) using primers L14724 and H15915 (He and Chen, 2007), Cytochrome c oxidase (COI) using primers: Fish F1 and Fish R1 (Ward et al., 2005) and partial CR using primers (Mitchell et al., 1993). PCR was carried out in a final volume of 10µl reaction using a PCR buffer (10 mM Tri–HCl, pH 8.3, and 50 mM KCl), 1.5 mM MgCl2, 0.2 mM of each dNTP, 0.25 mM of each primer, 5U of Taq polymerase and 1 μl of the template DNA (approximately 50ng). PCR thermal conditions were as follows: an initial denaturation at 95 C for 5 min, followed by 35 cycles of 95 C for 35 s, annealing temperature at 56˚C for 45 seconds, and extension at 72˚C for 75 seconds. The final extension was at 72˚C for 10 minutes. To monitor the effectiveness and consistency of the PCR reactions, positive controls were included. The amplified PCR products were run on a 2% agarose gel stained with ethidium bromide and visualized under UV light. To eliminate any residual primer, the amplified PCR products treated with Exonuclease I (EXO-I) and shrimp alkaline phosphatase (SAP) for 15 minutes each at 37˚C and 80˚C, respectively. Subsequently, the amplified PCR products were directly sequenced using the BigDye Terminator Kit (v3.1) and analyzed on an ABI 3500XL Applied Biosystems Genetic Analyzer. Both forward and reverse sequences were obtained for all products.

9.3 Data analysis

The sequences obtained from the forward and reverse directions were aligned and edited using SEQUENCHER® version 4.9 (Gene Codes Corporation, Ann Arbor, MI, USA). The analysis of each sequence was performed separately using the CLUSTAL X multiple sequence alignment program (Thompson et al., 1997), and the alignments were examined by visual inspection. The sequences of *Clarias* were downloaded from NCBI database for the comaparative purpose. To estimates the level of genetic diversity number of haplotypes (h), haplotype diversity (hd) and nucleotide diversity (π) within the yamuna population of *Clarias* were computed using the software DNASPv5.0 (Librado and Rozas 2009). For the phylogenetic relationship, we used the Tamura-3 parameter using a discrete Gamma distribution with the lowest BIC score value using MEGA software (Tamura et al., 2021).

9.3 Results

All 26 samples were amplified using Cytochrome c oxidase I gene. We used all 26 individuals *Clarias gariepinus* fish samples to evaluate the genetic variation, differentiation and phylogenetic relationship. We also used sequences of fishes from different geographical locations from Asia, Africa and middle east regions available in public domain. All these sequences were used for the genetic assessment of *Clarias gariepinus* in Yamuna River patch in New Delhi regions.

In total three haplotype in the *Clarias gariepinus* samples were observed in the present study. The sequences were clustered with the previously reported haplotypes in different river stream of the Indian rivers. The genetic relationships of different haplotypes is shown in figure 59. Estimated molecular diversity indices, haplotype (hd) and nucleotide (π) diversity for the species was 0.644 and 0.0082 respectively.

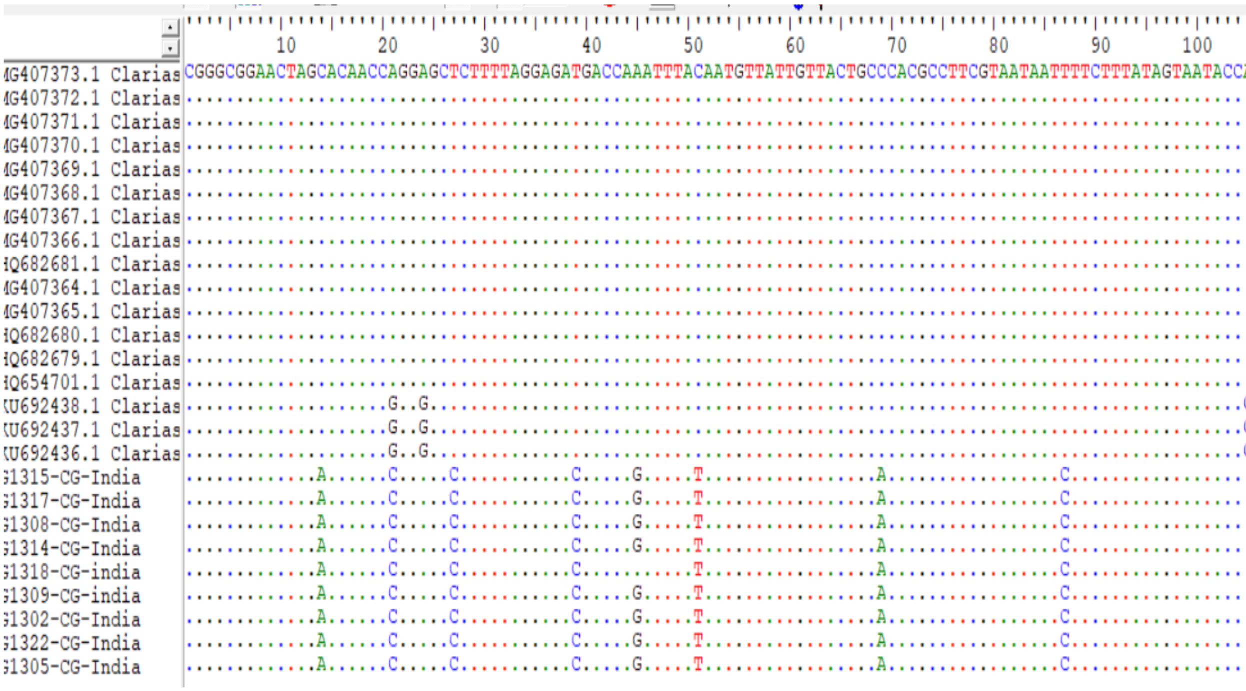


Figure 59: Sequence differentiation between *C. batrachus* and *C. gariepinus*.

Based on similarity an analysis, it was evident that the collected samples did not match with any of the species belonging to *Clarias batrachus*.

9.3.1 Phylogenetic analysis

Phylogenetic analysis was conducted for the population of Mangur from River Yamuna in Delhi in comparison with other populations from different countries. For the comparison, data of this particular species available in NCBI. It was found that the collected samples were forming two groups. The Mangur haplotypes from the present study forms cluster 1 with various other countries while the other group is seen to have similar ancestral origin. The clade of Mangur forms an early branch of the all the mangur community and diverges genetically from haplotypes from neighboring areas (Figure 60).

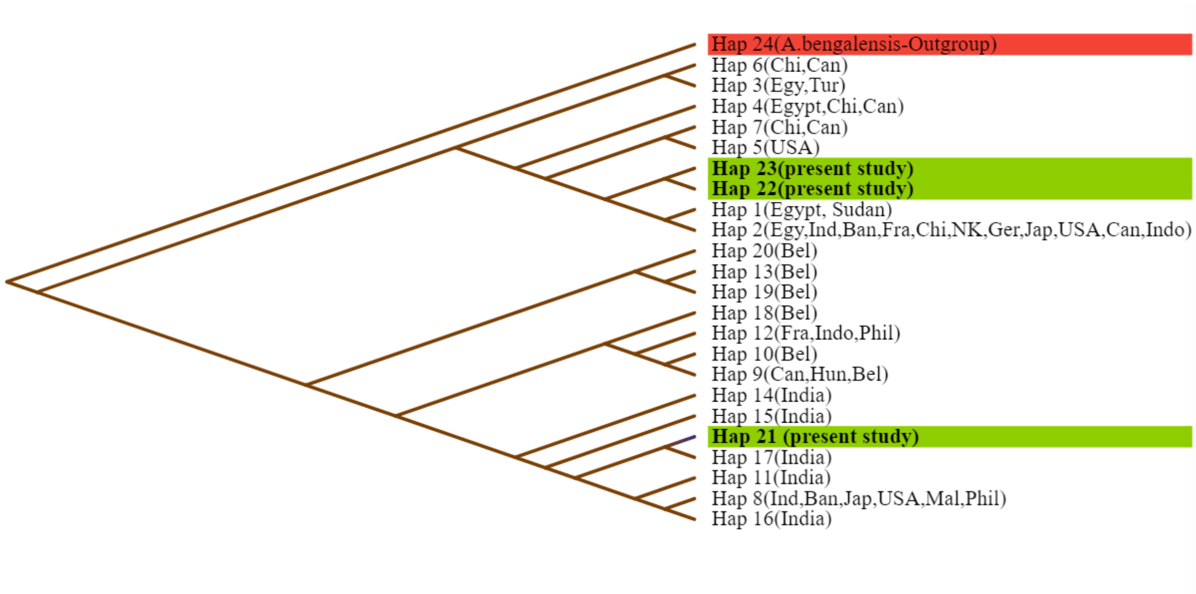


Fig 60: Tree showing phylogenetic relationship of *C. gariepinus* population in India with other populations across different countries.

.

9.4 Discussion

The phylogenetic analysis reveals that the present study samples are seen to be distinguished in 2 groups. The Mangur haplotypes from the present study in group 1 forms cluster with various other countries while the other group is seen to have similar ancestral origin. The clade of Mangur forms an early branch of the all the mangur community and diverges genetically from haplotypes from neighboring areas. Haplotype 24 (*Anguilla bengalensis*) is seen to be divergent from all the other haplotypes as it is an outgroup. It is observed that haplotype 23 and 24 forms cluster with various other countries while Haplotype 21 is seen to be clustering among the Indian population. Various authors have published researches on the identification problems between *Clarias gariepinus* and the closely related and partially sympatric, *C. anguillaris* belonging to the family *Anguilla bengalensis*. The only reliable difference between these species is the number of gill rakers on the first branchial arch (Agnese et al.,1996). Also, *A. bengalensis* share similar morphology and is one of the oldest clades observed when compared to *C. gariepinus*. This is the reason why we took *Anguilla bengalensis* into consideration as an outgroup in our study. This clearly indicates that though both species have similar morphology, they vary genetically with different ancestral origin.

Genetic diversity estimation suggests that a standard deviation of 0.101 is observed among the haplotype diversity. The number of haplotype diversity (Hd) was 3. The observed nucleotide diversity is 0.00827 along with the observed haplotype diversity of 0.644. Other data obtained from public domain for countries like Egypt and other parts of India shows a haplotype diversity of 0.522, nucleotide diversity of 0.00104 and 0.733, 0.00566, respectively (Elberri et al., 2020). It is observed that the present study samples and the previously studied samples from India did not vary much significantly in terms of diversity. This suggests that in the present study, both nucleotide diversity and haplotype diversity is distinct from the others which indicates that the genetic diversity of the fish in Egypt has declined as it could probably be a result of the small base population, inbreeding, and genetic drift (Navarro et al., 2000). If the genetic distance between populations is greater, breeding would be lesser and they might be more isolated from one another. Lower genetic distance between populations indicates that more breeding is present between them and are less isolated from one another. Hence, the populations are termed to be fixed. It means that they do not share any alleles with one another, i.e. they do not breed with one another and are completely different populations which are completely isolated from one another.

9.5 Conclusion

Analysis of the genetic characterization and genetic diversity of this species revealed two clearly different population group and their respective population structures. Our results from the genetic data suggests that the Indian population has good genetic diversity and apart from the Indian haplotypes, we also observed similar haplotypes in countries like Sudan, Egypt etc. Since this is a preliminary study, further studies should be carried out with larger sample size and more upcoming works must be conducted on microsatellite loci to provide more information about the population structure and conservation genetics of *Clarias gariepinus*. The baseline information on genetic characterization and genetic diversity and the information on the observed haplotypes generated from this study would be useful for planning effective strategies to understand if Thai mangur is present in the river Yamuna in Delhi NCR region.

9.6 References

- Barasa, J. E., Abila, R., Grobler, J. P., Agaba, M., Chemoiwa, E. J., & Kaunda-Arara, B. (2016). High genetic diversity and population differentiation in Clarias gariepinus of Yala Swamp: evidence from mitochondrial DNA sequences. *Journal of fish biology*, *89*(6), 2557-2570.
- Bardonnet, A., Bolliet, V., & Belon, V. (2005). Recruitment abundance estimation: role of glass eel (Anguilla anguilla L.) response to light. *Journal of Experimental Marine Biology and Ecology*, *321*(2), 181-190.
- Barnhoorn, I. E., Van Dyk, J. C., Genthe, B., Harding, W. R., Wagenaar, G. M., & Bornman, M. S. (2015). Organochlorine pesticide levels in Clarias gariepinus from polluted freshwater impoundments in South Africa and associated human health risks. *Chemosphere*, *120*, 391-397.
- Chris, U. O., Singh, N. B., & Agarwal, A. (2018). Nanoparticles as feed supplement on Growth behaviour of Cultured Catfish (Clarias gariepinus) fingerlings. *Materials Today: Proceedings*, *5*(3), 9076-9081.
- Devassy, A., Kumar, R., Shajitha, P. P., John, R., Padmakumar, K. G., Basheer, V. S & Mathew, L. (2016). Genetic identification and phylogenetic relationships of Indian clariids based on mitochondrial COI sequences. *Mitochondrial DNA Part A*, *27*(5), 3777-3780.
- Duong, T. Y., Scribner, K. T., Kanefsky, J., & Na-Nakorn, U. (2017). Lack of introgressive hybridization by North African catfish (Clarias gariepinus) in native Vietnamese bighead catfish (Clarias macrocephalus) populations as revealed by novel nuclear and mitochondrial markers. *Aquaculture*, *473*, 468-477.
- Gavryushkina, A., Welch, D., Stadler, T., & Drummond, A. J. (2014). Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS computational biology*, *10*(12), e1003919.
- Goda, A. M., El-Haroun, E. R., & Kabir Chowdhury, M. A. (2007). Effect of totally or partially replacing fish meal by alternative protein sources on growth of African catfish Clarias gariepinus (Burchell, 1822) reared in concrete tanks. *Aquaculture Research*, *38*(3), 279-287.
- Kassegne, A. B., Okonkwo, J. O., Esho, T. B., Mekonnen, K. N., Malehase, T., & Asfaw, S. L. (2018). Distribution of total mercury in surface sediments and African catfish (*Clarias gariepinus*) from Akaki River catchment and Aba Samuel Reservoir, downstream to the mega-city Addis Ababa, Ethiopia. *Emerging Contaminants*, *4*(1), 32-39.
- Khan, M. F., Panikkar, P., Salim, S. M., Leela, R. V., Sarkar, U. K., Das, B. K., & Eregowda, V. M. (2021). Modeling impacts of invasive sharp tooth African catfish Clarias gariepinus (Burchell 1822) and Mozambique tilapia Oreochromis mossambicus (Peters, 1852) on the ecosystem of a tropical reservoir ecosystem in India. *Environmental Science and Pollution Research*, *28*(41), 58310-58321.
- Khedkar, G. D., Tiknaik, A. D., Shinde, R. N., Kalyankar, A. D., Ron, T. B., & Haymer, D. (2016). High rates of substitution of the native catfish Clarias batrachus by Clarias gariepinus in India. *Mitochondrial DNA Part A*, *27*(1), 569-574.
- Lloyd, G. C. (2018). *Physiological responses and bioaccumulation of metals in Clarias gariepinus from the Nyl River system, Limpopo Province, South Africa* (Doctoral dissertation, University of Johannesburg (South Africa).
- Mohammed-Geba, K., Sheir, S. K., Hamed, E. A. E. A., & Galal-Khallaf, A. (2020). Molecular and morphological signatures for extreme environmental adaptability of the invasive mussel Brachidontes pharaonis (Fischer, 1870). *Molecular and Cellular Probes*, *53*, 101594.

- Pasch, J., Appelbaum, S., Palm, H. W., & Knaus, U. (2021). Growth of Basil (Ocimum basilicum) in Aeroponics, DRF, and Raft Systems with Effluents of African Catfish (Clarias gariepinus) in Decoupled Aquaponics (ss). *AgriEngineering*, *3*(3), 559-574.
- Tenesa, A., Navarro, P., Hayes, B. J., Duffy, D. L., Clarke, G. M., Goddard, M. E., & Visscher, P. M. (2007). Recent human effective population size estimated from linkage disequilibrium. *Genome research*, *17*(4), 520-526.
- Tyor, A. K., & Pahwa, K. (2017). Research article ovarian development of African sharptooth catfish (Clarias gariepinus, Burchell 1822) from Delhi segment of River Yamuna. *J. Fish. Aquatic. Sci*, *112*(3), 117-126.
- Van Steenberge, M. W., Vanhove, M. P., Chocha Manda, A., Larmuseau, M. H., Swart, B. L., Khang‟Mate, F., ... & Volckaert, F. A. (2020). Unravelling the evolution of Africa‟s drainage basins through a widespread freshwater fish, the African sharptooth catfish Clarias gariepinus. *Journal of Biogeography*, *47*(8), 1739-1754.
- Zhu, L., Li, W., Zha, J., Wang, M., Yuan, L., & Wang, Z. (2014). Butachlor causes disruption of HPG and HPT axes in adult female rare minnow (Gobiocypris rarus). *Chemico-Biological Interactions*, *221*, 119-126.