



OPEN

## Population genetics of the critically endangered three-striped turtle, *Batagur dhongoka*, from the Ganga river system using mitochondrial DNA and microsatellite analysis

Ajit Kumar<sup>1</sup>, Ashish Kumar Panda<sup>1</sup>, Aftab Alam Usmani<sup>1</sup>, Prabhaker Yadav<sup>1,2</sup>, Anshu Panwar<sup>1</sup>, Ruchi Badola<sup>1</sup>, Syed Ainul Hussain<sup>1</sup> & Sandeep Kumar Gupta<sup>1</sup>✉

The three-striped roofed (*Batagur dhongoka*) is a semi-aquatic turtle that belongs to family Geoemydidae. Due to anthropogenic pressure, it has been facing an intense decline of over 80% in its distribution range in the past 50 years. It is considered as 'Critically Endangered' so effective conservation strategies are needed to protect the species by determining their genetic diversity and population genetic structure. This study investigates the genetic diversity, population structure and demographic pattern of *B. dhongoka* from two Turtle Rescue and Rehabilitation Centre established near Ganga river using mitochondrial cytochrome *b* (Cyt *b*: 1140 bp); control region (CR: 451 bp) and ten nuclear microsatellite loci. mtDNA results show low levels of nucleotide diversity ( $\pi = 0.0022$ ) in *B. dhongoka* haplotypes and provide evidence for a low substitution rate. The demographic pattern estimated by the Bayesian skyline plot (BSP) analysis indicates historical stability followed by growth in the effective population size, with a recent reduction in population size from ~2 thousand years ago. The microsatellite findings show a moderate level of observed heterozygosity ( $H_o$ : 0.49). Bayesian-based clustering analysis revealed weak genetic structures in *B. dhongoka* and presence of admixed assignments suggesting close genetic relationships. These findings shed light on *B. dhongoka*'s genetic status and underline the necessity of comprehensive rehabilitation and relocation programs and conservation and management techniques to ensure the species' long-term survival. In order to ensure the effective protection and conservation of *B. dhongoka*, the Government of India has taken a proactive measure by incorporating it into Schedule I of the Wildlife (Protection) Act, 1972, as amended in 2022.

**Keywords** Genetic diversity, Heterozygosity, mtDNA, Microsatellite, Population structure

Turtles are the most threatened biota on the earth, with 40% of all extant turtles facing extinction risk<sup>1,2</sup>. 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<sup>3,4</sup>. 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<sup>5-7</sup>. 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<sup>3</sup>.

Consequently, there is a critical need to reevaluate the present distribution of this species<sup>5,6</sup>. 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<sup>3</sup>. Additionally, this species is used for subsistence consumption and is involved in some commercial trade, with recent confiscations reported in Malaysia and China<sup>3</sup>. Due to the increasing decline in its native range and a corresponding rise in international seizures, *B. dhongoka* has gained

<sup>1</sup>Wildlife Institute of India, P.O. Box # 18, Chandrabani, Dehra Dun, Uttarakhand 248001, India. <sup>2</sup>Present address: Department of System Biology, Centre for Biomedical Research, SGPGIMS Campus, Lucknow, India. ✉email: skg@wii.gov.in

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*<sup>8</sup>. 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 Rescue 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 Department rescued turtle eggs from the Chambal River and transferred them to the TRRC, Sarnath for hatching<sup>9</sup>. 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<sup>10</sup>. A decrease in population size can affect long-term evolutionary potential by reducing individual fitness and genetic variation in populations<sup>11</sup>. 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<sup>6</sup>. 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<sup>12,13</sup>. 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<sup>14</sup>. 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*<sup>15</sup>. Additionally, the utilization of the genomic approach by Çilingir et al.<sup>16</sup> on the Burmese roofed turtle (*B. trivittata*) in Myanmar, as well as investigation of the southern river terrapin (*B. affinis*) in Cambodia<sup>17</sup>, 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<sup>18</sup>.

In this study, we investigate the genetic diversity and population structure of *B. dhongoka* using two mitochondrial DNA (mtDNA) 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.

## Results

### Genetic diversity

Genetic diversity parameters revealed with mtDNA and microsatellite markers in the *B. dhongoka* from TRRC and GARRC are presented in Table 1. In this study, two mtDNA fragments were obtained from 92 samples; after alignment, 1140 bp of *cyt b* and 451 bp of the control region were generated, producing a concatenated fragment of 1591 bp. 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

Population/region	Mitochondrial DNA							Microsatellites				
	n	H	Hd	$\pi$	K	Tajima's D*	Fu's F <sub>s</sub> *	Na	Ar	Ho	HE	FIS
Turtle Rescue and Rehabilitation Centre (TRRC) Sarnath, Uttar Pradesh, India	80	29	0.93	0.003	4.6	2.04	−12.3	5	3.84	0.487	0.493	0.018
Ganga Aqualife Rescue and Rehabilitation Centre (GARRC) Narora, Uttar Pradesh, India	12	6	0.85	0.002	3.4	1.816	0.015	3.4	3.4	0.497	0.5	0.038
Total	92	31	0.94	0.002	4.6	1.56	−1.86	4.1	3.82	0.495	0.494	0.028

**Table 1.** Summary of genetic diversity in *Batagur dhongoka* populations based on concatenated mtDNA (Cyt *b* + CR) and ten microsatellites. n Number of samples, H haplotype, Hd haplotype diversity,  $\pi$  nucleotide diversity, Na number of alleles, Ar allelic richness, HO observed heterozygosity, HE expected heterozygosity, FIS inbreeding coefficient, \*all *P*-values > 0.01 (not significant).

populations. The haplotypes (h) and nucleotide diversity ( $\pi$ ) of TRRC were  $0.932 \pm 0.013$  and  $0.003 \pm 0.00012$ , respectively and those of GARRC were  $0.85 \pm 0.074$  and  $0.002 \pm 0.0002$ , respectively. The overall Hd and  $\pi$  were  $0.94 \pm 0.012$  and  $0.0022 \pm 0.00014$ , respectively (Table 1).

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 ( $N_a$ ) in TRRC and GARRC were 5 and 3.4, respectively, with comparative allelic richness ( $A_r$ ) in TRRC ( $A_r = 3.84$ ) and GARRC ( $A_r = 3.4$ ) with mean  $A_r = 3.82$  (Supplementary table ST1). The observed heterozygosity ( $H_o$ ) and expected heterozygosity ( $H_e$ ) in TRRC were  $H_o: 0.487$ ;  $H_e: 0.493$ ; and in GARRC were  $H_o: 0.497$ ;  $H_e: 0.50$  with mean  $H_o: 0.495$ ;  $H_e: 0.495$ . The mean inbreeding coefficient ( $F_{IS}$ ) value for both the rescued *B. dhongoka* populations was greater than zero mean  $F_{IS}: 0.028$ , indicating a heterozygote deficiency (Table 1), which may be attributed to the Wahlund effect and population not being in HWE.

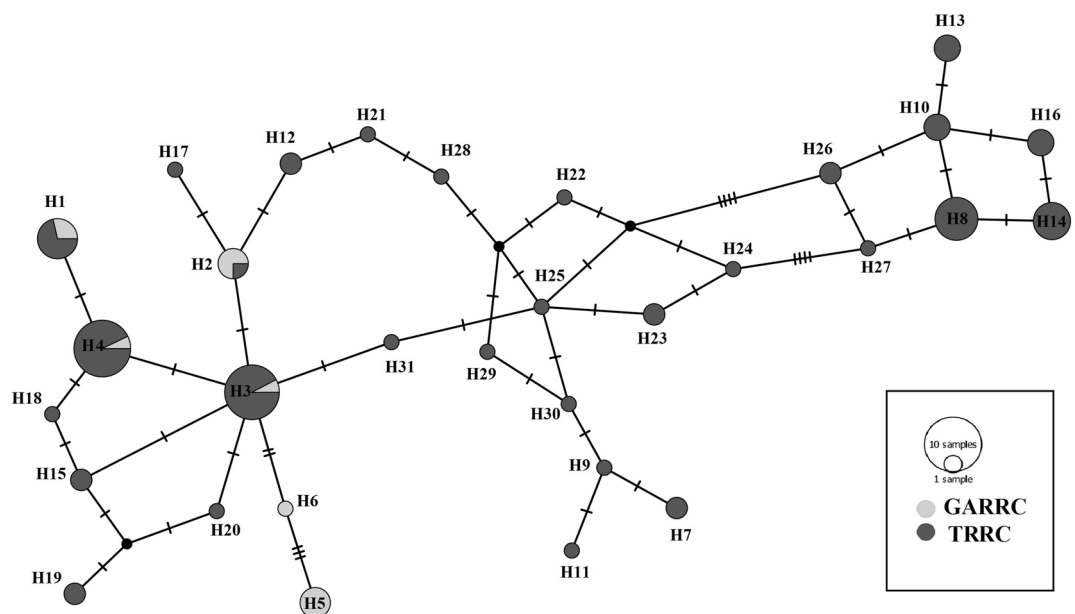
### Genetic structure, differentiation and migration

The median-joining (MJ) network of 31 haplotypes revealed no major branching events among the *B. dhongoka* (Fig. 1). 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 (Fig. 2) and Structure analysis (Fig. 3) 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  $\Delta K$  (mean likelihood of K (Mean  $\ln P(K) = -1787.306$ )). 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 signature observed in TRRC indicates that the individuals of TRRC population shared a similar genetic ancestry with the GARRC population.

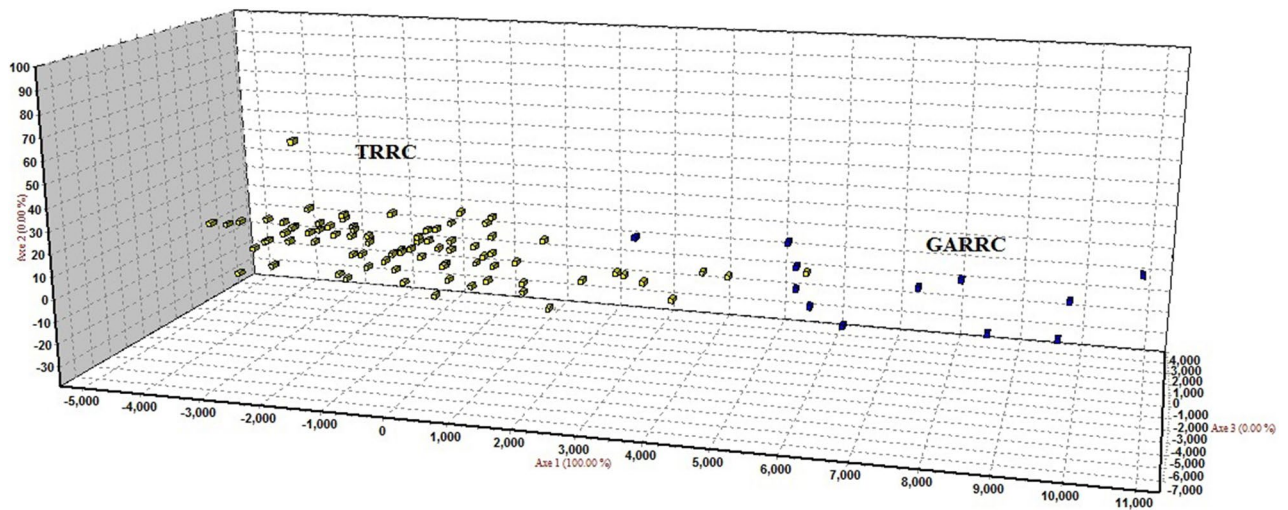
The genetic differentiation level between the studied populations of *B. dhongoka* was lower in mtDNA:  $F_{ST} = 0.003$ , whereas a comparatively higher value was observed based on microsatellite:  $F_{ST} = 0.074$ . Furthermore, the rate of self-distribution within the GARRC population was 0.9649 (95% 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.

### 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 (Fig. 4), which might result from a relatively stable population size over the long period. Results of Tajima's D and Fu's  $F_s$  tests were used to explain the population history of both populations (Table 1). 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



**Figure 1.** 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 2.** 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.

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 (Fig. 5).

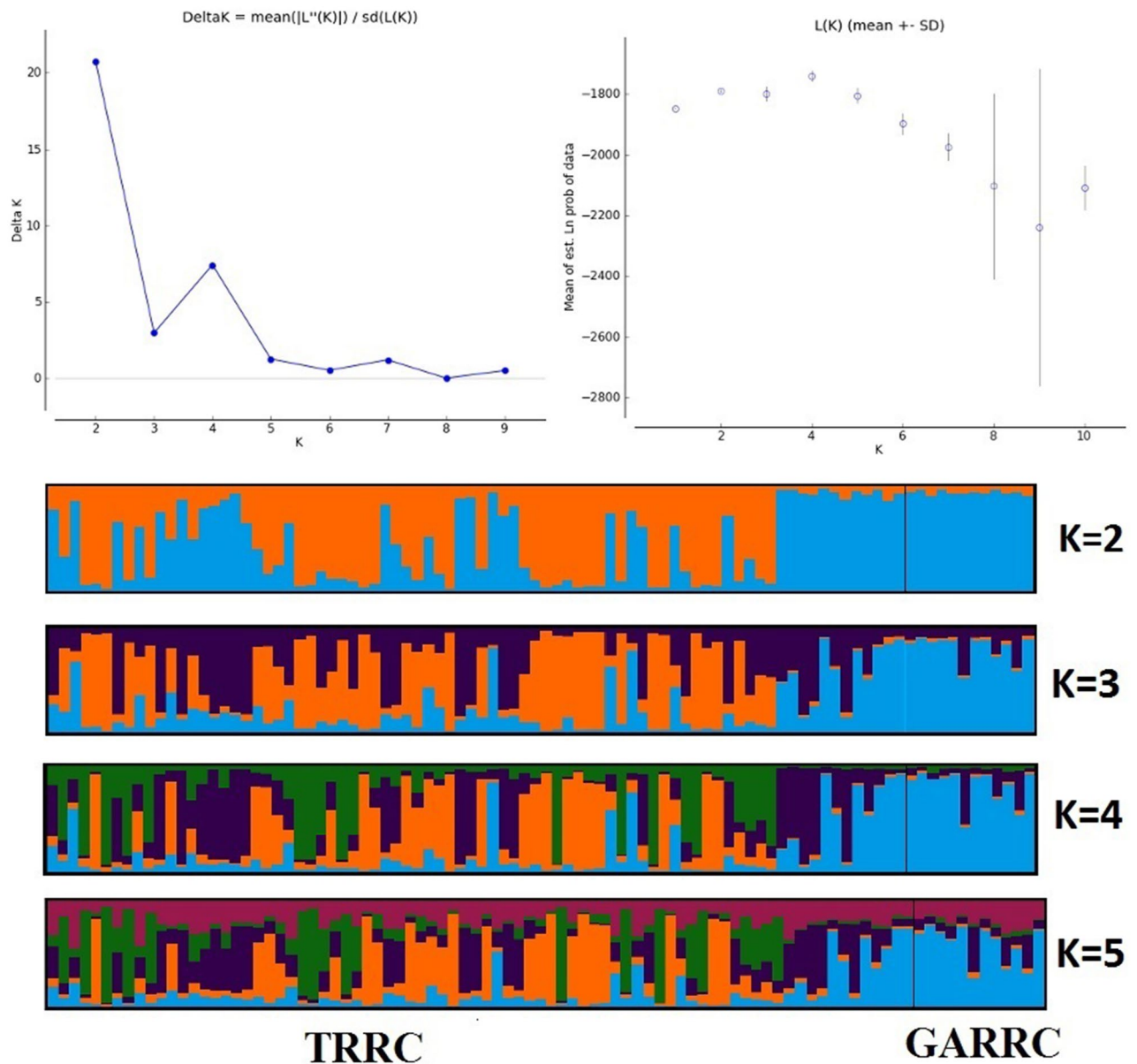
#### Bottleneck

The values of G-W index test (TRRC =  $0.31 \pm 0.07$  and GARRC =  $0.32 \pm 0.11$ ) were lower than the critical value  $M_c = 0.68$ , 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.

## 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<sup>19</sup>. 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<sup>20</sup>, whereas lower haplotype diversity in turtle species, e.g., in Leatherback turtles, *Dermochelys coriacea*<sup>21</sup> and Bog turtle, *Glyptemys muhlenbergii*<sup>22</sup>. Our data show a low nucleotide diversity with high haplotype diversity, which can be a signature of population expansion from a small effective population size<sup>23</sup>. 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.

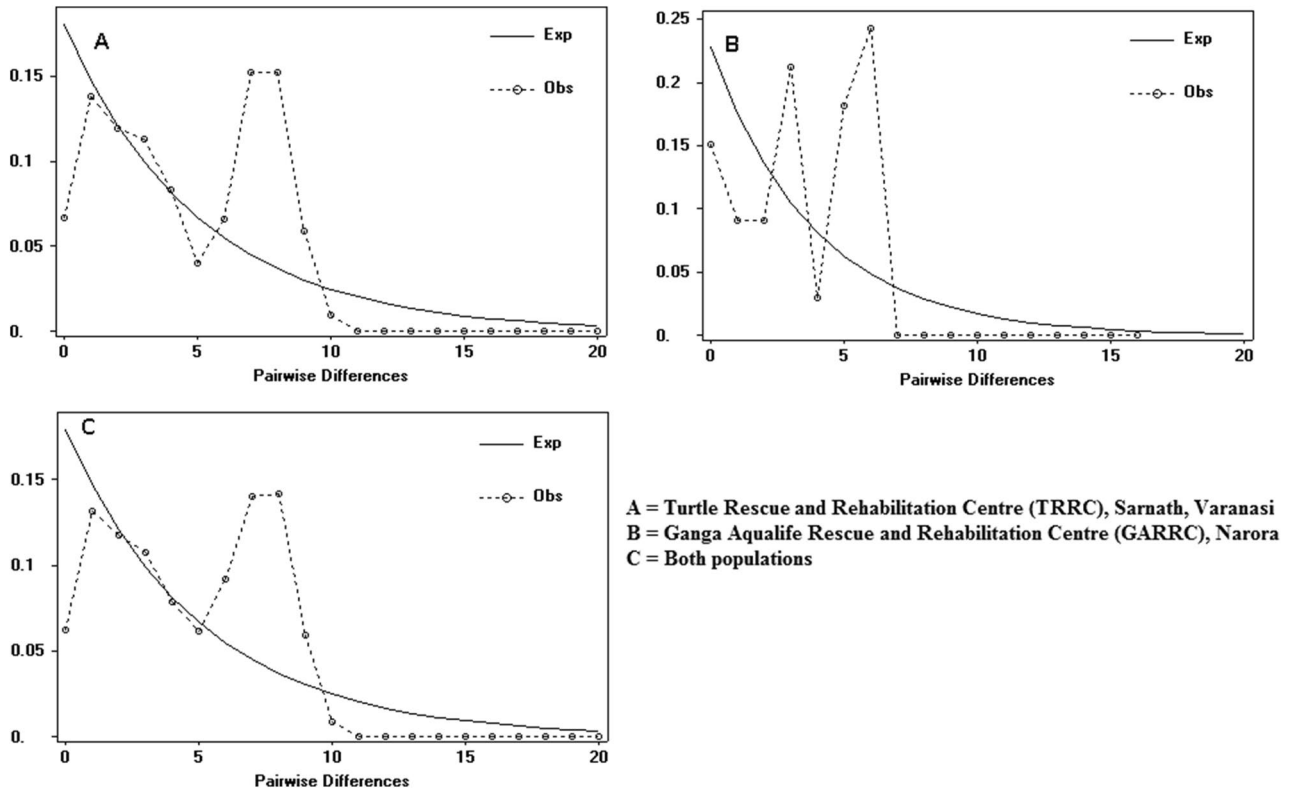
The BayesAss analysis indicated recent migration from GARRC to TRRC ( $m = 0.15$ ), which may be attributed to human interference and the mixing of individuals in TRRC (Ganga and Chambal rivers). The rate of self-assembly was higher in both GARRC and TRRC, which may be due to limited gene flow and low 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<sup>24</sup>. Moreover, the results of migration rates were consistent with pairwise  $F_{ST}$  values obtained from microsatellite and mtDNA



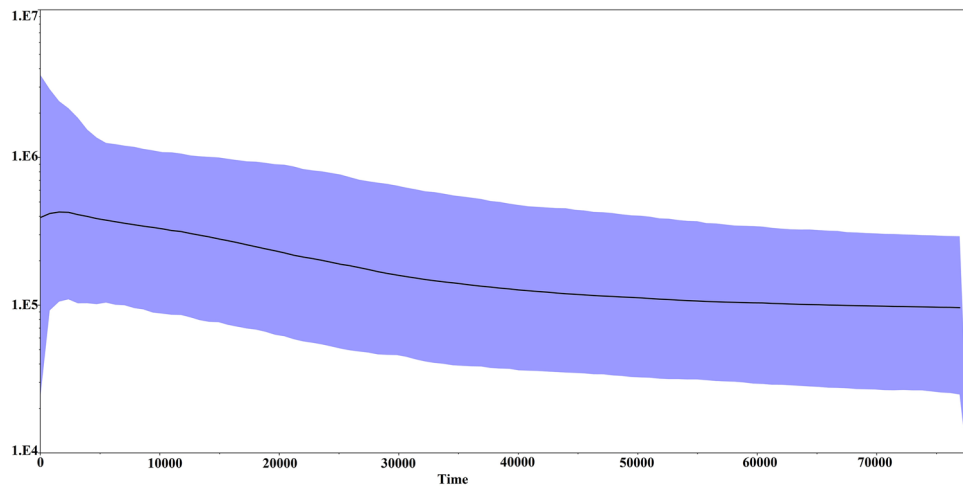
**Figure 3.** CLUMPK genetic structure plot derived from Bayesian-based cluster analysis in STRUCTURE. The proportion of color in each bar represents the assignment probability of an individual, corresponding to different groups of TRRC: Turtle Rescue and Rehabilitation Centre, Sarnath and GARRC: Ganga Aqualife Rescue and Rehabilitation Centre, Narora.

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<sup>3</sup>, 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<sup>25,26</sup>. 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<sup>1</sup>. 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<sup>3,5</sup>. 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



**Figure 4.** The mismatch distributions graph dashed line showing observed distribution; solid line showing the theoretical expected distribution under a growth-decline model.



**Figure 5.** 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.

diversity are small population sizes, which are intricately connected to the fitness and health of individuals, along with limited population growth capacity<sup>27</sup>. 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<sup>12,13</sup>. 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*.

### Conservation implications

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.

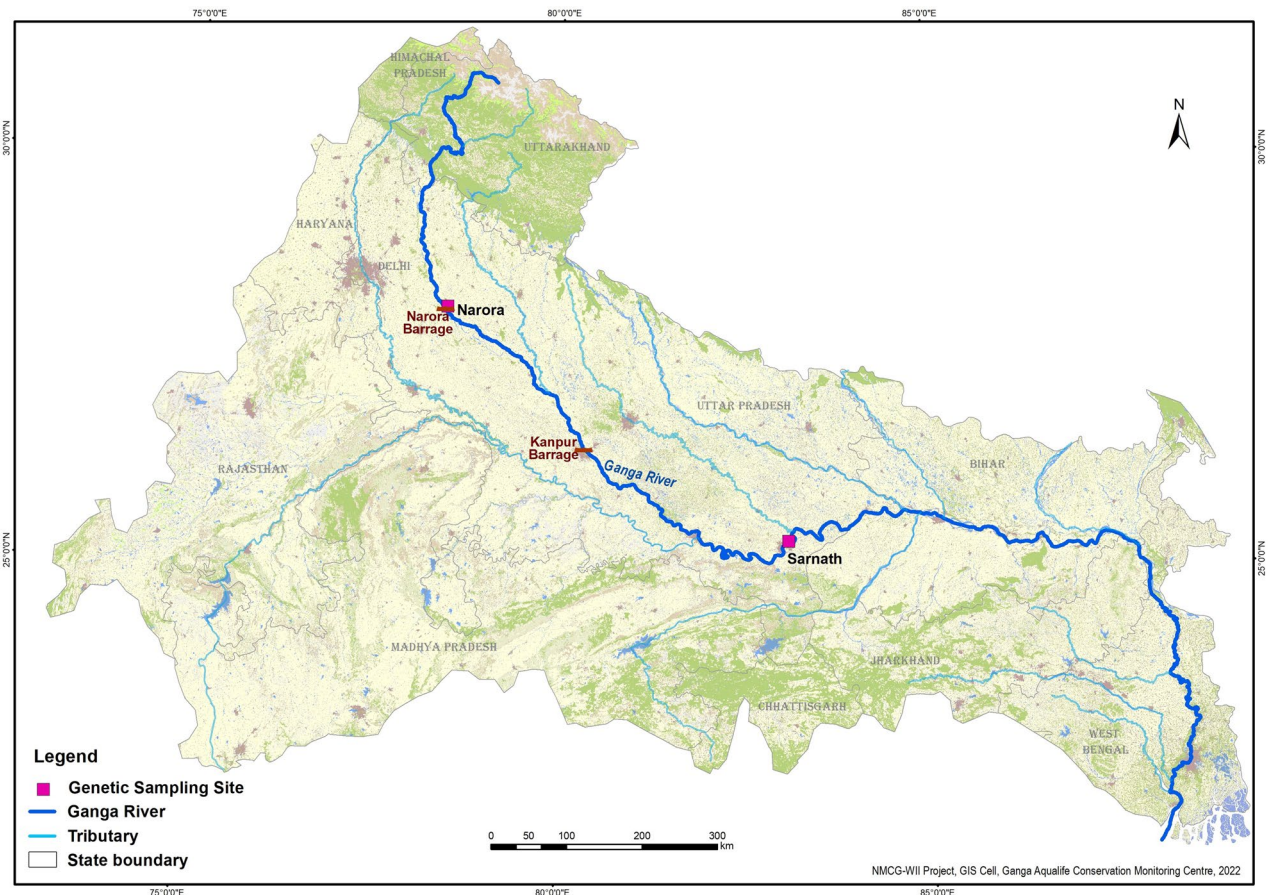
### Materials and methods

#### 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 during 2017 to 2021 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.

#### Sample collection and DNA extraction

A total of 92 tissue samples (eggshell, swabs from hatchlings and remains of dead individuals) were collected from TRRC, Sarnath, Varanasi (n = 80) and GARRC, Narora (n = 12) (Fig. 6). 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



**Figure 6.** Map of the study area showing sampling sites of *Batagur dhongoka*, which was created using ArcGIS Desktop: Release 10.6.1 Redlands, CA: Environmental systems Research Institute.

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  $\mu$ l. The gDNA quantification of samples was done in Qubit 4 fluorometer (Invitrogen) and diluted to make 50–100 ng/ $\mu$ l for working stock.

### mtDNA amplification and sequencing

Two mitochondrial regions: cytochrome *b* (Cyt *b*) gene ~ 1140 bp and mtDNA control region (CR) ~ 451 bp were amplified using primers: CytbG: 5'–AACCATCGTTGTWATCAACTAC–3'<sup>28</sup> and mt-f-na3:5'–AGGGTG GAGTCTTCAGTTTTGGTTTACAAGACCAATG–3'<sup>29</sup> and CRF: 5'–AGCACCGGTCTTGTAACCA–3' and CRR: 5'–ACAAAACAACCAAAGGCCAG–3'<sup>30</sup>. PCR reactions for both regions were performed in 10  $\mu$ l reaction volumes using 1  $\times$  PCR buffer (10 mM Tris–HCl, pH8.3, and 50 mM KCl), 1.5 mM MgCl<sub>2</sub>, 0.2 mM of each dNTPs, 3 pmol of each primer, 5 units of Dream Taq DNA Polymerase (Thermo Scientific) and 1  $\mu$ l (~ 50 to 100 ng) 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 s, annealing at 56 °C for Cytb gene and 54 °C for CR for 60 s and extension at 72 °C for 1 min 30 s. 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 transilluminator. 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.

### Microsatellite genotyping

Genotyping were performed using thirteen microsatellites loci, including Maucas01, Maucas03, Maucas06, Maucas18, Maucas21, Maucas24<sup>31</sup>; TWT113, TWS190<sup>32</sup>; GP19<sup>33</sup>; GmuD16, GmuB08<sup>22</sup>; Test 21<sup>34</sup>; msEo41<sup>35</sup>. Of these, three 'Maucas21, Maucas24 and GmuD16' did not amplify consistently in all samples; hence, excluded from the further analysis. Therefore, final analysis was performed with ten microsatellite markers. The forward primer of each marker was attached with one of the 5' universal sequence tail: T7 (5'-TAATACGACTCACTATAGGG), M13 (5'-TGTAACGACGCGCCAGT) and M13R (5'-CAGGAAACAGCTATGACC)<sup>36</sup>. The universal primers M13, M13R and T7 were labelled with fluorescent dye FAM, HEX, and TET, respectively. All fluorescent-labelled universal primer were dissolved in Tris EDTA buffer (10 mM Tris, 1 mM EDTA, pH 8.0) to produce a 20 mM stock solution and 5 mM locus-specific forward and reverse primer. Primer mix was prepared with universal tail labelled forward primer: reverse primer: dye-labelled universal primer = 1:4:4<sup>36</sup>. Multiplexing of primers was carried out according to fragment size and dye-labelled (Supplementary table ST2). PCR was carried out in 10  $\mu$ l PCR reaction volumes, containing 5  $\mu$ l of QIAGEN Multiplex PCR Buffer Mix (QIAGEN Inc.), 0.4  $\mu$ M primer mix, and 80–100 ng of the template DNA. The PCR cycle was performed under the touchdown conditions, initial denaturation at 95 °C for 15 min, followed by 8 cycles of denaturation at 95 °C for 45 s, annealing at 58 °C for 60 s and extension at 72 °C for 75 s, again followed by 15 cycles of denaturation at 95 °C for 45 s, annealing at 58–50 °C for 60 s and extension at 72 °C for 75 s which followed by 12 cycles of denaturation at 95 °C for 45 s, annealing at 52 °C for 60 s and extension at 72 °C for 75 s 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).

### Data analysis

#### Mitochondrial DNA diversity

The forward and reverse sequences were aligned and edited using SEQUENCHER version 4.9 (Gene Codes Corporation, Ann Arbor, MI, USA). All the sequences were aligned using CLUSTAL W<sup>37</sup>, and checked manually in BioEdit v 7.2.5 software<sup>38</sup>. DnaSP 5.0<sup>39</sup> 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<sup>40</sup> was used to calculate mean within and between population genetic distances (uncorrected p-distances).

#### Population demography

The spatial distribution of the haplotypes was visualized through a median-joining network, which was created using the PopART software<sup>41</sup>. To determine whether the *B. dhongoka* populations carried a signal of demographic expansion or stationary population history, Tajima's D<sup>42</sup> and Fu's Fs<sup>43</sup> neutrality tests were performed in Arlequin ver 3.5<sup>44</sup>. To generate the trends in spatial demography history, mismatch analysis was carried out using the population growth-decline model in DnaSP<sup>39</sup>, 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<sup>44</sup>. 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<sup>45</sup>. Bayesian skyline analyses available in BEAST v1.7<sup>46</sup> 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<sup>47</sup>. We performed multiple analyses that were run for 10<sup>8</sup> iterations with a burn-in of 10<sup>7</sup> 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  $\times$  10<sup>-8</sup> substitutions/site/year<sup>48,49</sup>). The mean mutation rate was also consistent with the other turtle species mutation rate of 1.2–2.4%/site per Myrs<sup>50</sup>. Length of MCMC chain was 100 million and sampling every



10,000 generations. Results of replicate runs were combined using LogCombiner v1.7.5<sup>46</sup> with a burn-in of 20 million iterations (20%) for each run. Tracer v1.7.1<sup>51</sup> was used to assess the convergence of chains and to reconstruct Bayesian Skyline Plot.

#### Microsatellite analysis

Microsatellite data was analyzed using the program Micro-Checker 2.2<sup>52</sup> to identify genotyping errors, such as null alleles, large allele dropout, or errors in scoring. The number of effective alleles ( $N_e$ ), alleles per locus ( $N_a$ ), observed heterozygosity ( $H_o$ ), and expected heterozygosity ( $H_e$ ) were estimated using the GenAlEx v6.5 program<sup>53</sup>. The polymorphic information content (PIC) values were calculated using CERVUSv3.0.7<sup>54</sup>. The supporting format of the data input files was prepared using CONVERTv1.31<sup>55</sup>. Allelic richness ( $A_r$ ) 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 1000 permutations<sup>56</sup>. All the loci were checked for under HWE in GenAlEx v6<sup>53</sup>. 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).

#### Genetic structure, differentiation and migration

The pairwise  $F_{ST}$  values (gene flow) among the populations were calculated using GenAlEx v6<sup>53</sup>. Initially, a factorial correspondence analysis (FCA) was performed using GENETIX ver. 4.02<sup>57</sup>, 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<sup>58</sup>. 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  $\Delta K$  method<sup>59</sup> using web server of Structure Harvester<sup>60</sup>. 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<sup>61</sup>. We used burn-in iterations  $10^6$  followed by  $10^7$  iterations and a sampling frequency of 2000. The initial run was performed with the default delta ( $\Delta$ ) value of 0.15 for allele frequencies ( $A$ ), migration ( $M$ ), and inbreeding coefficient ( $F$ ). The final input parameter of  $\Delta M$  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.

#### Population history: bottlenecks

We used Garza–Williamson (G–W) index<sup>62</sup> and BOTTLENECK v.1.2.02<sup>63</sup> programs to determine the signal of past bottlenecks. We estimated G–W index using ARLEQUIN v3.5<sup>44</sup>, 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<sup>62</sup>. The value of  $M$  ratio below the critical  $M_c$  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<sup>63</sup>. We also checked allele frequency distributions for mode shifts that discriminate populations that have experienced a recent bottleneck from stable populations<sup>64</sup>.

#### Ethical approval and informed consent

This study did not involve human clinical data; hence no informed consent was required for this work. Samples were mostly collected from the dead remain and necessary permission has been obtained from the forest department. All methods were carried out in accordance with relevant guidelines and regulations. This study was carried out in compliance with the ARRIVE guidelines. The experiments were conducted after obtaining necessary approval from the Institutional Animal Ethics Committee vide letter no. No. WII/IAEC/2017-18 dated 25/06/2020.

#### Data availability

All the relevant sequencing data has been uploaded on NCBI, GenBank (OQ378390–OQ378573).

Received: 28 February 2023; Accepted: 16 February 2024

Published online: 11 March 2024

#### References

- Stanford, C. B. *et al.* Turtles and tortoises are in trouble. *Curr. Biol.* **30**, 721–735 (2020).
- Shaffer, H. B., FitzSimmons, N. N., Georges, A. & Rhodin, A. G. Turtle conservation genetics working group. *Chelonian Res. Monogr.* **4**, 107–123 (2007).
- 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. <https://doi.org/10.2305/IUCN.UK.2019-1.RLTS.T10953A152042542.en> (2019).
- 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. <https://doi.org/10.2305/IUCN.UK.2019-1.RLTS.T10949A152043133.en> (2019).
- Rashid, S. M. A. & Khan, S. M. M. H. Trade and conservation status of freshwater turtles and tortoises in Bangladesh. In *Asian Turtle Trade: Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia*. *Chelonian Research Monographs* Vol. 2 (eds van Dijk, P. P. *et al.*) 77–85 (Lunenburg, Chelonian Research Foundation, 2000).

6. 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 Research Monographs* Vol. 2, 86–94 (2000).
7. IUCN TFTSG and ATTWG (IUCN/SSC Tortoise and Freshwater Turtle Specialist Group and Asian Turtle Trade Working Group). Recommended changes to 1996 IUCN Red List Status of Asian turtle species. in *Asian Turtle Trade: Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia*. *Chelonian Research Monographs* (eds van Dijk, P. P., Stuart, B. L. & Rhodin, A. G. J.) Vol. 2, 156–164 (Chelonian Research Foundation, Lunenburg, 2000).
8. The Print. World Turtle Day: UP Government Releases Nearly 300 Endangered Turtles Into Chambal River (2022). <https://theprint.in/india/up-300-turtles-released-into-chambal-on-eve-of-world-turtle-day/966698/>
9. 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>
10. Burns, C. E. *et al.* The origin of captive Galapagos tortoises based on DNA analysis: Implication for the management of natural populations. *Anim. Conserv.* **6**, 329–337 (2003).
11. Reed, D. H. & Frankham, R. Correlation between fitness and genetic diversity. *Conserv. Biol.* **17**, 230–237 (2003).
12. Villemey, A., Besnard, A., Grandadam, J. & 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).
13. McCleery, R., Hostetler, J. A. & 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).
14. Schlötterer, C. & Pemberton, J. The use of microsatellites for genetic analysis of natural populations. In *Molecular Ecology and Evolution: Approaches and Applications* 203–214 (Birkhäuser, Basel, 1994).
15. Kumar, A., Yadav, P., Usmani, A., Hussain, S. A. & 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).
16. Çilingir, F. G. *et al.* Conservation genomics of the endangered Burmese roofed turtle. *Conserv. Biol.* **31**, 1469–1476 (2017).
17. Çilingir, F. G. *et al.* Last exit before the brink: Conservation genomics of the Cambodian population of the critically endangered southern river terrapin. *Ecol. Evol.* **9**, 9500–9510 (2019).
18. 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).
19. King, T. L. & Julian, S. E. Conservation of microsatellite DNA flanking sequence across 13 Emydid genera assayed with novel bog turtle (*Glyptemys muhlenbergii*) loci. *Conserv. Genet.* **5**, 719–725 (2004).
20. Baltazar-Soares, 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).
21. Wongfu, C. *et al.* Genetic diversity in leatherback turtles (*Dermochelys coriacea*) along the Andaman sea of Thailand. *Diversity* **14**, 764 (2022).
22. Rosenbaum, P. A., Robertson, J. M. & Zamudio, K. R. Unexpectedly low genetic divergences among populations of the threatened bog turtle (*Glyptemys muhlenbergii*). *Conserv. Genet.* **8**, 331–342 (2007).
23. Rogers, A. R. & Harpending, H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**, 552–569 (1992).
24. Meirmans, P. G. Nonconvergence in Bayesian estimation of migration rates. *Mol. Ecol. Resour.* **14**, 726–733 (2014).
25. Excoffier, L., Estoup, A. & Cornuet, J. M. Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics* **169**, 1727–1738 (2005).
26. Peery, M. Z. *et al.* Reliability of genetic bottleneck tests for detecting recent population declines. *Mol. Ecol.* **21**, 3403–3418 (2012).
27. Spielman, D., Brook, B. W. & Frankham, R. Most species are not driven to extinction before genetic factors impact them. *Proc. Natl. Acad. Sci.* **101**, 15261–15264 (2004).
28. 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).
29. 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).
30. Gaur, A., Reddy, A., Annapoorni, S., Satyarebala, B. & Shivaji, S. The origin of Indian Star tortoises (*Geochelone elegans*) based on nuclear and mitochondrial DNA analysis: A story of rescue and repatriation. *Conserv. Genet.* **7**, 231–240 (2006).
31. Vamberger, M., Stuckas, H. & Fritz, U. 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).
32. Perez, M., Bour, R., Lambourdiere, J., Samadi, S. & 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).
33. Schwartz, T. S., Osentoski, M., Lamb, T. & 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).
34. Forlani, A. *et al.* Identification and characterization of microsatellite markers in Hermann's tortoise (*Testudo hermanni*, Testudinidae). *Mol. Ecol. Notes* **5**, 228–230 (2005).
35. Pedall, I., Schäfer, H., Fritz, U. & Wink, M. Isolation of microsatellite markers in the *Emys orbicularis* complex and development of multiplex PCR amplification. *Conserv. Genet.* **10**, 725–727 (2009).
36. Ge, C., Cui, Y. N., Jing, P. Y. & Hong, X. Y. An alternative suite of universal primers for genotyping in multiplex PCR. *PLOS ONE* **9**, e92826 (2014).
37. 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).
38. Hall, T. A. BioEdit software, version 5.0.9. North Carolina State University, Raleigh, NC (1999)
39. Librado, P. & Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
40. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
41. Leigh, J. W. & Bryant, D. Popart: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
42. Tajima, H. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
43. Fu, Y. X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925 (1997).
44. 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).
45. Darrriba, D., Taboada, G. L., Ramon, D. & Posada, D. JModelTest 2: More models, new heuristics and high-performance computing. *Nat. Methods* **30**, 772 (2012).
46. 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).

47. Drummond, A. J., Rambaut, A., Shapiro, B. E. T. H. & Pybus, O. G. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**, 1185–1192 (2005).
48. Formia, A., Godley, B. J., Dontaine, J. F. & Bruford, M. W. Mitochondrial DNA diversity and phylogeography of endangered green turtle (*Cheloniemydas*) populations in Africa. *Conserv. Genet.* **7**, 353–369 (2006).
49. Naro-Maciel, E. *et al.* From refugia to rookeries: Phylogeography of Atlantic green turtles. *J. Exp. Mar. Biol. Ecol.* **461**, 306–316 (2014).
50. Encalada, S. E. *et al.* 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).
51. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).
52. Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**, 535–538 (2004).
53. Peakall, R. O. & Smouse, P. E. GENALEX 6: Genetic analysis in Excel Population genetic software for teaching and research. *Mol. Ecol. Notes* **6**, 288–295 (2006).
54. 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**, 1099–1106 (2007).
55. Glaubitz, J. C. Convert: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Mol. Ecol. Notes* **4**, 309–310 (2004).
56. Goudet, J. F. FSTAT (version 1.2): A computer program to calculate F-statistics. *J. Hered.* **86**, 485–6 (1995).
57. Belkhir, K. GENETIX 4.05, logiciel sous Windows TM pour la genetique des populations. <http://www.genetix.univ-montp2.fr/genetix/genetix.htm>.2004.
58. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
59. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
60. Earl, D. A. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conser. Genet. Resour.* **4**, 359–361 (2012).
61. Wilson, G. A. & Rannala, B. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177–1191 (2003).
62. Garza, J. C. & Williamson, E. G. Detection of reduction in population size using data from microsatellite loci. *Mol. Ecol.* **10**, 305–318 (2001).
63. Cornuet, J. M. & Luikart, G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001–2014 (1996).
64. 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).

## Acknowledgements

We gratefully acknowledge the help and support received from the Director General, National Mission for Clean Ganga, Ministry of Jal Shakti, Government of India for funding support. We take this opportunity to thank the Director and Dean and the entire project team of NMCG-WII for their support. We acknowledge the Forest Departments of Uttar Pradesh for granting permission to carry out the research work. We thank Ms. Aishwarya Ramachandran, Ms. Maneesha Bhatt and Mr. A. Madhanraj for their support.

## Author contributions

S.A.H. and R.B. conceived the study, raised funds, supervised the study, review and editing of the manuscript. S.K.G. & A.K. designed the methodology for molecular study. A.K.P. and A.A.U. collected samples from the field. A.K., P.Y. and A.P. performed the laboratory work and generated the data. A.K. performed data analysis and wrote the manuscript. S.K.G. supervised and facilitated the validation and interpretation of the data and did review and editing of the manuscript. All authors agreed and approved the final manuscript.

## Funding

This study was funded by the National Mission for Clean Ganga (NMCG), Ministry of Jal Shakti, Government of India sponsored project ‘Biodiversity Conservation and Ganga Rejuvenation (Grant no. B-02/2015-16/1259/NMCG-WII)’ and ‘Planning and Management for Aquatic Species Conservation and Maintenance of Ecosystem Services in the Ganga River Basin for a Clean Ganga’. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-54816-0>.

**Correspondence** and requests for materials should be addressed to S.K.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024