



Population genomic assessment of the endangered endemic catfish *Glyptothorax stocki* reveals distinct lineages in Pakistan

Mohsan Raza · Arif Muhammad Khan ·
Chenhong Li 

Received: 11 February 2026 / Revised: 17 June 2026 / Accepted: 21 June 2026
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2026

Abstract The endangered and endemic catfish *Glyptothorax stocki* faces mounting threats in Pakistan, yet its population genetics remain entirely uncharacterized. To address this gap and inform genetically based conservation strategies, we conducted the first genome-scale study of *G. stocki*, analyzing 24 individuals from three river systems: Swat (KPK), Poonch (KK), and upper Neelum (KN). Moreover, mitochondrial genome analysis placed *G.*

stocki within a well-supported clade and provided its phylogenetic context among related Sisoridae taxa. Genome-wide SNP analysis revealed strong population structuring. Pairwise F_{ST} values showed strong divergence between KPK and KN ($F_{ST}=0.175$) and minimal differentiation between KK and KN ($F_{ST}=0.003$). Nucleotide diversity was high in KK (0.0697) and KN (0.0646) but low in KPK (0.0256). KK and KN showed heterozygote deficits ($F_{IS}=0.22, 0.11$), whereas KPK showed excess heterozygosity ($F_{IS}=-0.13$), with genomic inbreeding highest in KK and lowest in KPK. Clustering analyses (PCA, ADMIXTURE, STRUCTURE) and a maximum likelihood tree grouped the populations into two distinct genetic clusters: one comprising KK and KN, and a second, highly divergent cluster represented by KPK. These results point to a long-term isolation of the KPK population, likely driven by geographic barriers. Therefore, we propose managing these clusters as separate management units (MUs). This study provides a genomic baseline for *G. stocki*'s genetically informed conservation and restoration plans.

Handling editor: Christian Sturmbauer

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10750-026-06290-0>.

M. Raza · C. Li
Shanghai Universities Key Laboratory of Marine Animal Taxonomy and Evolution, Shanghai Ocean University, Shanghai 201306, China

M. Raza · C. Li
Engineering Research Center of Environmental DNA and Ecological Water Health Assessment, Shanghai Ocean University, Shanghai 201306, China

M. Raza · C. Li (✉)
College of Fisheries and Life Science, Shanghai Ocean University, 999 Hucheng Huan Road, Pudong New District, Shanghai 201306, China
e-mail: chli@shou.edu.cn

A. M. Khan
Department of Biotechnology, University of Sargodha, Sargodha 40100, Pakistan

Keywords Conservation genomics · Population structure · Endemic fish · Indus River basin · Management units

Introduction

Freshwater biodiversity is declining more rapidly than terrestrial and marine biodiversity (Dudgeon et al., 2006; Butchart et al., 2010; Collen et al., 2014). Many freshwater species have become endangered due to human activities, including population growth near riverine systems, pollution, habitat fragmentation (caused by dams and weirs), and overfishing (Cincotta et al., 2000; Sala et al., 2000; Lakra et al., 2010). Generally, riverine fish species are more susceptible to physical, chemical, and biological factors in aquatic environments. Rapid fluctuations in water quality or habitat can have severe consequences, leading to extinction of species (Poff & Allan, 1995; Gehrke et al., 2002; Meador & Goldstein, 2003). Keeping those in mind, restoration plans have been established to protect endangered species (Moyle et al., 2003; Funk et al., 2012; Harrison et al., 2014). Earlier, the genetic information was overlooked while making the restoration plans. However, this information has been consistently recommended before developing conservation plans for endangered species (Namkoong, 2000; Frankham, 2005). Although the genetic tools have been advanced and are now commonly integrated into conservation management, their overall use remains limited in developing countries, primarily due to a lack of awareness and financial constraints (Pierson et al., 2016; Pavlova et al., 2017; Taylor et al., 2017).

The members of the Sisorid catfishes are well known to survive in swift-flowing riverine systems and streams. Within the family Sisoridae, the genus *Glyptothorax* Blyth, 1860, is a highly diversified genus having over 90 species, out of which 70 are legitimate nominal species (Ng & Hadiaty, 2009; Sayyadzadeh et al., 2022). Members are characterized by a specialized thoracic adhesive apparatus that enables attachment to substrates in high-current habitats (Jiang et al., 2012). Recent mitochondrial genome-based phylogenies have produced conflicting results regarding the monophyly of *Glyptothorax*, likely due to differences in taxon sampling. For example, Iyyappan et al. (2026) recovered *Glyptothorax* as monophyletic but did not include representatives of *Gagata* Bleeker, 1858. Similarly, Sayyadzadeh et al. (2022) did not include *Gagata* taxa, whereas Jiang et al. (2012) included *Gagata gagata* (Hamilton, 1822), but not *Gagata dolichonema* He, 1996. In contrast,

broader phylogenetic analyses including *Gagata* species have suggested that *Glyptothorax* may be polyphyletic (Schedel et al., 2022).

Glyptothorax stocki Mirza & Nijssen, 1978, is an ecologically essential and distinctive species (Mirza & Nijssen, 1978; Thomson & Page, 2006). It is a small freshwater benthopelagic species (14 cm in length and weighing up to 21 g) that lives in the bottom and mid-water columns of these systems, and prefers to stay in microhabitats accompanied by sandy or rocky substrates present in springs or turbulent mountains (Mirza & Nijssen, 1978). It contains up to 76.34% protein and has economic value for local people. It is an endemic species of Pakistan with records from Azad Kashmir, Khyber Pakhtunkhwa (KPK), and Punjab (Javed et al., 2013; Ishaq et al., 2014; Hussain & Dutta, 2016; Rahman et al., 2025; Ullah et al., 2025); however, no recent records have been reported from Punjab. *G. stocki* remains poorly studied, with limited phylogenetic data and no published mitochondrial genome to date. Moreover, it is listed as endangered on the IUCN Red List due to significant population decline and an elevated risk of extinction (IUCN, 2020). Concerning its endangered status, *G. stocki* demands urgent conservation action. To date, research on this species has focused on the discovery of new records and morphological descriptions (Javed et al., 2013; Ishaq et al., 2014; Rahman et al., 2025; Ullah et al., 2025).

Genetic diversity plays a crucial role in the survival and adaptation of a species in relation to its environment (Frankham, 2003; DeWoody et al., 2021). High genetic variability is directly proportional to a species' ability to survive in its environment; conversely, low genetic variability can lead to population decline or extinction (Frankham, 2003; DeWoody et al., 2021). Before creating a conservation plan, it is essential to examine the genetic structure and diversity of a particular species. The world of conservation genetics has been revolutionized by the advent of advanced sequencing technologies, particularly next-generation sequencing (NGS). This high-throughput data is being used to gain insights into genetic structure, gene flow, identification of distinct genetic units, inbreeding coefficient, and other valuable information crucial for developing genetically based conservation strategies (Allendorf et al., 2012).

Despite *G. stocki* being an endangered species, no previous studies have been conducted on its genetic structure and genetic diversity, and phylogenetic placement, which could potentially hinder practical

conservation efforts. Due to the construction of the Mangla dam and the flow of water from two different river systems into the Indus River system, we hypothesize that there might be two distinct populations. The present study is focused on a genome-wide assessment of *G. stocki* in Pakistan using next-generation sequencing. By conducting genetic analysis on *G. stocki*, including population structure, genetic diversity, and inbreeding coefficients, this study provides a scientific baseline for implementing genetically informed conservation strategies across their native range.

Materials and methods

Sample collection and DNA preparation

A total of 30 individuals of *G. stocki* were collected via non-lethal sampling from three river systems across its known distribution in Pakistan: one river from Khyber Pakhtunkhwa (KPK), the Swat River, and two rivers from Azad and Jammu Kashmir (AJK), the upper Neelum River (KN), and the Poonch River (KK), near the Line of Control (India border) (Fig. 1). These sites were selected based on the

species known habitat preferences and previous ecological surveys. Despite multiple attempts, no specimens were obtained from the Chenab River system in Punjab, consistent with the lack of recent records from this region. The habitat picture is provided in the supplementary Fig. 1. The details of the collected specimens are given in Table 1.

During field sampling, a small number of individuals of *Glyptothorax punjabensis* were also encountered alongside *G. stocki* (Mirza & Kashmiri, 1971). No clear variation in external coloration was observed among *G. stocki* individuals across the sampled locations. Due to the politically sensitive nature of parts of Kashmir, fieldwork was conducted in coordination with local authorities under constrained conditions. Following capture, individuals were morphologically identified using established taxonomic keys. A small fin clip was collected from each specimen using sterile scissors and forceps, in accordance with ethical protocols for tissue sampling (Hammer et al., 2013). All fish were handled with care and released immediately at the capture site. The sex of individuals was not identified during sampling because of the noninvasive sampling protocol employed in this study. Fin clips were preserved in 95% ethanol to prevent DNA

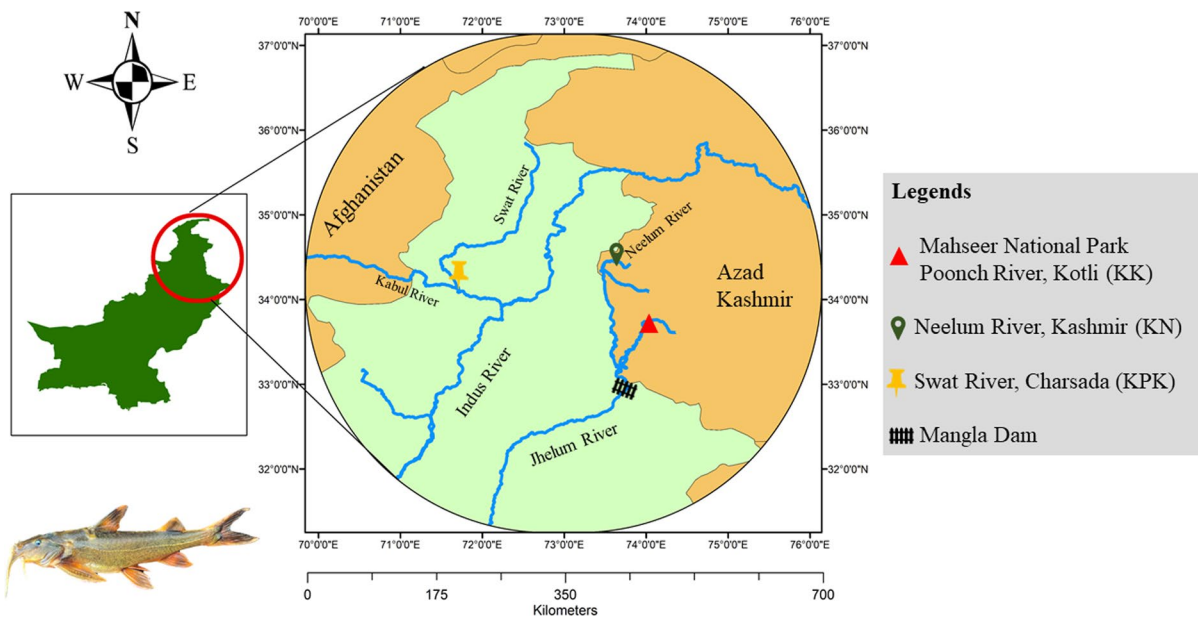


Fig. 1 Sampling locations of *G. stocki* in Pakistan. *G. stocki* samples were collected from three river systems: the Swat River (KPK, yellow), the Poonch River near Kotli (KK, red),

and the upper Neelum River (KN, green). The location of the Mangla Dam is indicated with a barrier. The inset shows the study area within Pakistan

Table 1 Sampling information for the 30 *G. stocki* individuals. Coordinates are reported in the WGS84 geographic coordinate reference system

Group	No. of samples	Samples retained after QC	Mitogenomes assembled	Locality	Latitude	Longitude
KK	10	7	1	Poonch River, Kotli, Kashmir	33.5129° N	73.8830° E
KPK	10	7	0	Swat River, Khyber Pakhtunkhwa	34.165° N	71.591° E
KN	10	10	2	Neelum River, Muzaffarabad, Kashmir	34.448° N	73.5477° E

degradation and stored at -20°C upon return to the laboratory until DNA extraction.

DNA extraction, library preparation, and sequencing

Genomic DNA was extracted from ethanol-preserved fin clips using the TIANamp Genomic DNA Kit (Tiangen Biotech, China). DNA quality and concentration were assessed spectrophotometrically. Qualified DNA was sheared to an average fragment size of 500 bp using a SCIENTZ18-A Ultrasonic DNA Interrupter (Scientz, Ningbo, China). The DNA fragmentation protocol involves 22 cycles of ultrasonication for 20 s followed by a pause of 20 s, with a processing power of 300 W. Sequencing libraries were constructed using a standard protocol that included end repair and ligation of indexed adapters for sample multiplexing. Solid-phase reversible immobilization (SPRI) beads were used for size selection and purification steps throughout the library preparation process. Final library quality was verified by agarose gel electrophoresis, and quantification was performed using a Qubit fluorometer (Thermo Fisher Scientific, USA). Libraries were pooled in equimolar ratios and subjected to 150-bp paired-end sequencing on an Illumina HiSeq platform at Novogene (Shanghai, China).

Genomic data processing

Because there is no reference genome available for *G. stocki*, *Danio rerio* (Hamilton, 1822) was used as the reference for assembling a set of 4434 single-copy coding loci of *G. stocki* following the Assexon pipeline of Yuan et al. (2019). Each sample was demultiplexed from the pooled sequences based on unique inline indexes specific to each individual. After demultiplexing, samples were subjected to quality

filtering and adapter trimming to ensure the removal of adapters and the presence of quality reads. Furthermore, the low-quality bases (Phred score < 20) were trimmed to keep high-quality reads (Martin, 2011). The resulting cleaned reads were used for the assembly using the Perl script assemble.pl script developed by (Yuan et al., 2019). During the assembly step, paired-end reads were merged, and contigs were constructed for each sample. The PCR duplicates and residuals were filtered out using the delete-unwanted.pl script to ensure the presence of unique, high-confidence sequence information and to minimize bias in the dataset (Yuan et al., 2019). After individual assembly, all of the processed samples were merged into a single file using the Python script "merged.pl" (Yuan et al., 2019). The combined sequences were aligned using the "alignment.pl" script, resulting in an alignment suitable for phylogenetic reconstruction. Consensus sequences were made from the alignments and used as a reference for variant calling across multiple individuals (Yuan et al., 2019).

Variant calling and filtering

The consensus file, carrying sequences for each target locus, was generated using a custom Perl script that extracted the required sequences from assembled contigs (Yuan et al., 2016). The pregenerated trimmed files were aligned to the consensus loci using BWA v0.7.5 (Li & Durbin, 2009). The resulting files (SAM) were converted into Binary Alignment/Map (BAM) format using SAMtools to enable the efficient downstream processing. The BAM files were then genotyped using GATK v4 (Genome Analysis Toolkit) and the GATK Practices workflow (McKenna et al., 2010; DePristo et al., 2011; Van Der Auwera et al., 2013). The HaplotypeCaller, together with basic quality score and duplicate flagging, was performed for variant calling. SNP filtering was conducted using PLINK v1.9,

where loci with a genotype missing rate greater than 20% ($-\text{geno } 0.2$) were excluded. In addition, SNPs with a minor allele frequency (MAF) below 0.05 ($-\text{maf } 0.05$) were removed to reduce the influence of rare variants. The final results of the SNP data were contained in a (.vcf) file. These high-quality final SNPs were then converted into suitable input formats for use in population genetic analysis.

Phylogenetic reconstruction

The phylogenetic tree was constructed among the studied individuals based on a concatenated dataset of aligned DNA sequences created from a dataset file prepared at step 2.3. It employed the maximum likelihood (ML) approach with IQ-TREE v2, along with 1000 ultrafast bootstrap replicates, to assess node support (Minh et al., 2020). The best-fitting model was selected by using the $-\text{spp}$ option within IQ-TREE, which automatically identifies the best-fitting substitution model. The visualization of the resulting tree file (.treefile) was performed using FigTree v1.4.4, which allowed for manual editing of branch labels, rooting, and annotation to achieve the final look of the tree (Rambaut, 2009). Three mitochondrial genomes were assembled from whole-genome resequencing data using NOVOPlasty v4.3.5 (Dierckxsens et al., 2017), a seed-and-extend assembler designed for organelle genome reconstruction from short-read sequencing data. Moreover, to investigate the phylogenetic placement of *G. stocki*, complete mitochondrial genome sequences were compiled, including the newly assembled mitogenomes from this study and publicly available Sisoridae mitochondrial genomes retrieved from GenBank. The details of the samples, including the accession numbers, are given in Supplementary Table 1. Sequences were aligned using MAFFT, and poorly aligned regions were removed using trimAl. A maximum likelihood phylogenetic tree was reconstructed using IQ-TREE with automatic model selection (ModelFinder). Branch support was assessed using 1000 ultrafast bootstrap replicate tests. *D. rerio* was included as an out-group to root the tree.

Nucleotide diversity π

The nucleotide diversity was calculated from the filtered VCF file (Nei & Li, 1979; Nei, 1987; Danecek et al., 2011). The vcfR package in R was

used for the data preparation, reading the VCF file, and indel removal (Knaus & Grünwald, 2017; Paradis & Schliep, 2019); moreover, pegas was used for the calculation of nucleotide diversity (Paradis, 2010); the visualization of the obtained results was done with ggplot2 (Wickham, 2016).

Population genetic analysis

To infer population structure, we employed three complementary methods. First, a Bayesian clustering analysis was performed using STRUCTURE v2.3.4. We used the admixture model with correlated allele frequencies, running 10 iterations for each value of K (genetic clusters) from 1 to 5. Each run included a burn-in period of 100,000 steps followed by 500,000 Markov chain Monte Carlo (MCMC) repetitions. The optimal K value was determined using the ΔK method (Evanno et al., 2005) implemented in STRU CTURE SELECTOR. Second, a principal component analysis (PCA) was conducted using PLINK v1.9. The SNP data were converted to PLINK binary format, and the resulting eigenvectors were visualized in R using ggplot2. Third, we performed maximum likelihood estimation of ancestry proportions using ADMIXTURE software. The analysis was run for $K=1$ to 4, and the results were visualized in Tbttools (Chen et al., 2020).

Genetic diversity

Genetic differentiation among populations was quantified by calculating pairwise F_{ST} values using the Stacks pipeline (Catchen et al., 2013). Additionally, deviations from Hardy–Weinberg equilibrium (F_{IS}) for each population were estimated using the hierfstat package in R, based on the methods of Weir & Cockerham (1984). The individual genomic inbreeding was quantified using SNP-based excess homozygosity (PLINK F) by using the function $-\text{het}$, which estimates individual genomic inbreeding coefficient (F) based on excess homozygosity (Purcell et al., 2007; Kardos et al., 2016).

Population branch statistic (PBS) was used to identify loci showing population-specific signals of differentiation among the three populations (Yelman et al., 2021). Variant data were obtained from a filtered VCF file and converted to a genlight object

in adegenet (R v4.1.1). For each population, allele frequencies were estimated from diploid genotype dosages, excluding missing data. Pairwise genetic differentiation between populations was quantified at each SNP using the Hudson estimator of F_{ST} , which corrects for finite sample sizes. Pairwise F_{ST} values were transformed as $T = -\log(1 - F_{ST})$, and PBS for each population was calculated (Yi et al., 2010). Functional annotation of candidate loci was performed by mapping SNP positions to nearby genes of already annotated genes of all available genomes using biomaRt (Ensembl) (Kinsella et al., 2011). The corresponding Ensembl dataset was queried, and genes located within ± 50 kb of each SNP position were retrieved (Ensembl gene ID, gene symbol, description, genomic coordinates, and strand). The distance from each SNP to each candidate gene interval was calculated, and the nearest gene(s) per SNP were retained.

Results

Data assembly and analysis

Whole-genome resequencing of 30 individuals produced high-quality data for 24 samples after quality filtering. The detailed information, like sequencing depth and number of reads, is given in Supplementary Table 2. After sequence assembly, alignment, and quality filtering, a total of 900–1442 single-copy loci were successfully retrieved per individual. Following stringent quality control, 1000 high-confidence loci were retained for phylogenetic and population genomic analyses. The concatenated alignment of these loci spanned a total of 362,850 base pairs. Variant calling using GATK4 identified 9450 high-quality SNPs across all individuals, which were used for downstream analysis, including genetic structure and diversity estimation.

Phylogenetic analysis

The maximum likelihood phylogenetic analysis based on complete mitochondrial genome sequences recovered *G. stocki* as a well-supported monophyletic clade, with all three newly assembled mitochondrial genomes from the KK and KN populations clustering

together (Fig. 2). This clade was positioned within the broader *Glyptothorax* lineage and formed a sister relationship with *Glyptothorax trilineatus* Blyth, 1860, and *Glyptothorax annandalei* Hora, 1923. The genus *Gagata* (*G. dolichonema*) was recovered within the *Glyptothorax* clade. At higher taxonomic levels, the overall topology was consistent with expected relationships within Sisoridae, with *Bagarius bagarius* (Hamilton, 1822) and *Glyptosternon maculatum* (Regan, 1905) forming successive sister lineages, and more distantly related taxa (*Clarias batrachus* (Linnaeus, 1758), and *D. rerio*) placed basally as out-groups.

Moreover, the evolutionary relationships among *G. stocki* individuals were inferred from a maximum likelihood phylogeny constructed from 1000 high-quality loci. The resulting tree, rooted with *D. rerio* as an out-group, showed a clear bifurcation into two major clades (Fig. 3). One clade exclusively comprised individuals from the KPK population (Swat River), while the second clade contained all individuals from the KK (Poonch River) and KN (Neelum River) populations, demonstrating significant genetic divergence of the KPK group.

Population structure, admixture, PCA, and FST analysis

The pairwise F_{ST} values were calculated to reveal genetic differentiation among three *G. stocki* populations. The values of pairwise F_{ST} were highly variable, representing significant differentiation between populations (Table 2), with the greatest divergence observed between KPK and KN ($F_{ST}=0.175$; $P<0.001$), moderate divergence between KK and KPK ($F_{ST}=0.134$; $P<0.01$), and the least divergence between KK and KN ($F_{ST}=0.003$; $P<0.05$). Very low differentiation was observed between KK and KN, suggesting a high level of genetic similarity between these two populations. These results support the clustering patterns observed in PCA, phylogenetic, and STRUCTURE analysis, indicating substantial genetic divergence of the KPK population from the other two.

Bayesian clustering analysis in STRUCTURE identified two distinct genetic clusters ($K=2$) as the optimal population structure (Fig. 4. A and B). All 24 individuals were assigned to their respective clusters with negligible admixture, indicating strong genetic

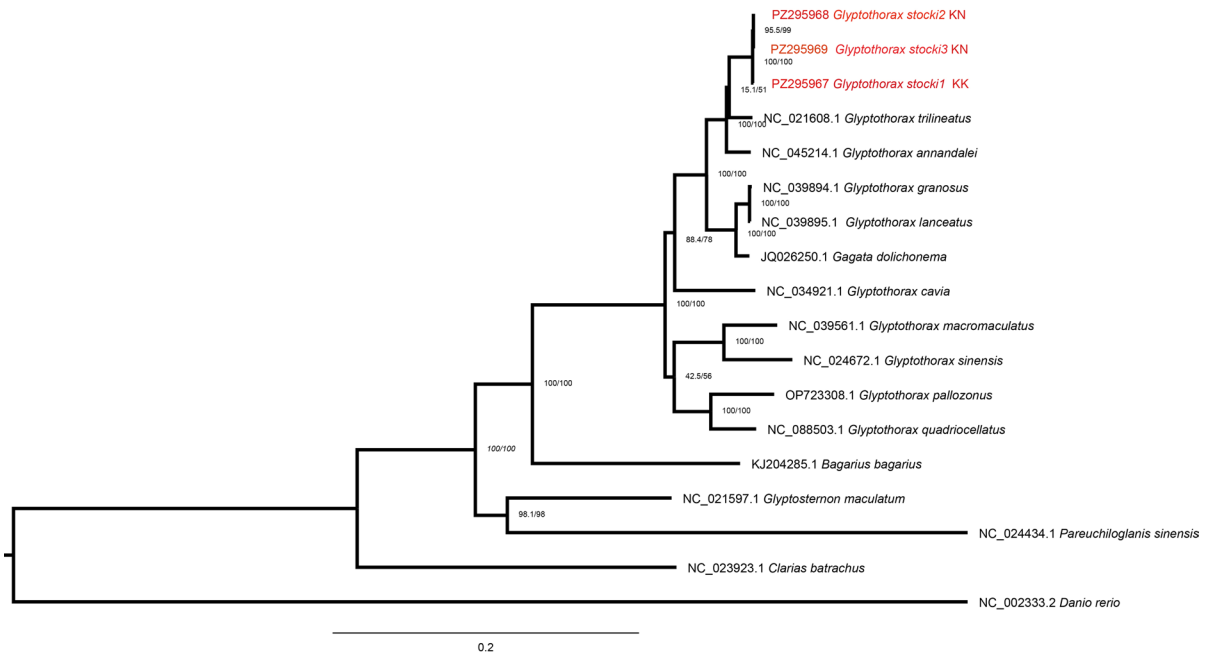


Fig. 2 Maximum likelihood phylogenetic tree based on complete mitochondrial genome sequences of *G. stocki* and related Sisoridae taxa. The *G. stocki* individuals are highlighted in red, with corresponding GenBank accession numbers and sampling localities (KN: Neelum River, Kashmir; KK: Poonch River, Kashmir) indicated beside the sample names. The tree

was inferred using IQ-TREE under the best-fitting substitution model selected by ModelFinder. Branch support values are shown at nodes, bootstrap (UFboot) percentages based on 1000 replicates. *D. rerio* was used as an out-group. Scale bar indicates the number of substitutions per site

differentiation. One cluster consisted exclusively of individuals from the KPK population, while the second cluster contained all individuals from the KK and KN populations. This result is fully congruent with the phylogenetic and PCA analysis.

This bifurcation was further supported by ADMIXTURE analysis (Fig. 5A). The most biologically meaningful clustering occurred at $K=2$, which cleanly separated the KPK population from a unified KK/KN genetic group. At $K=3$, the model failed to distinguish KK from KN, instead showing admixture between them. This suggests very recent divergence or ongoing gene flow between these two sites. The consistent results from PCA, STRUCTURE, phylogenetics, and ADMIXTURE strongly support the existence of two primary genetic populations: an isolated KPK population and a second population comprising the closely related KK and KN individuals. Principal component analysis (PCA) revealed clear genetic structuring among the sampled populations (Fig. 5B). The first two principal components (PC1 and PC2) explained 13.08% and 7.21% of the total genetic

variance, respectively. The plot showed a distinct separation of the KPK individuals from a tight cluster formed by the KK and KN individuals, corroborating the patterns observed in the STRUCTURE and phylogenetic analysis.

Nucleotide diversity

Estimates of nucleotide diversity (π) revealed significant differences in genetic variation among the populations. Diversity was significantly lower in the KPK population ($\pi=0.0256$) than in the KK ($\pi=0.0697$) and KN ($\pi=0.0646$) populations, which showed similar, high levels of diversity. The high overall nucleotide diversity ($\pi=0.1196$) reflects the substantial genetic divergence between the two identified clusters.

Inbreeding coefficient

Population-level deviations from the Hardy–Weinberg equilibrium varied among river systems. The

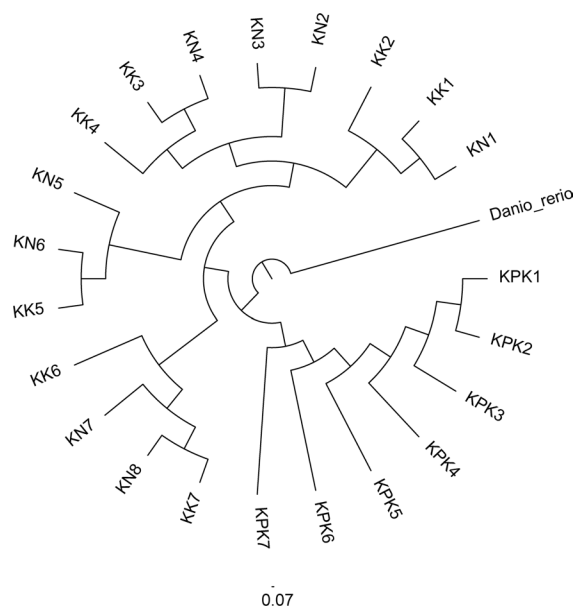


Fig. 3 Maximum likelihood phylogenetic tree of *G. stocki* individuals based on 1000 loci. The tree reveals two primary genetic clusters: one comprising individuals from the Swat River (KPK) and a second containing individuals from the Poonch (KK) and Neelum (KN) rivers. *D. rerio* was used as an out-group

Table 2 Genetic differentiation (F_{ST}) and significance (P -values) between three population pairs

Group	KK	KPK	KN
KK		$P < 0.001$	$P = 0.040$
KPK	0.134**		$P < 0.001$
KN	0.003*	0.175**	

* $P < .05$, ** $P < .01$

KK and KN populations exhibited heterozygote deficits ($F_{IS} = 0.22$ and 0.11 , respectively), whereas the KPK population showed a weaker signal consistent with heterozygote excess ($F_{IS} = -0.13$). Individual genomic inbreeding coefficients estimated from excess homozygosity (PLINK F) revealed pronounced variation among populations and among individuals. Median genomic inbreeding was highest in KK (median $F = 0.33$), intermediate in KN (0.15), and lowest in KPK (0.03), with substantial within-population heterogeneity. Overall, patterns of genomic inbreeding were broadly concordant with population-level F_{IS} estimates, indicating

elevated genomic homozygosity in the KK and KN populations relative to KPK.

Loci under selection (PBS)

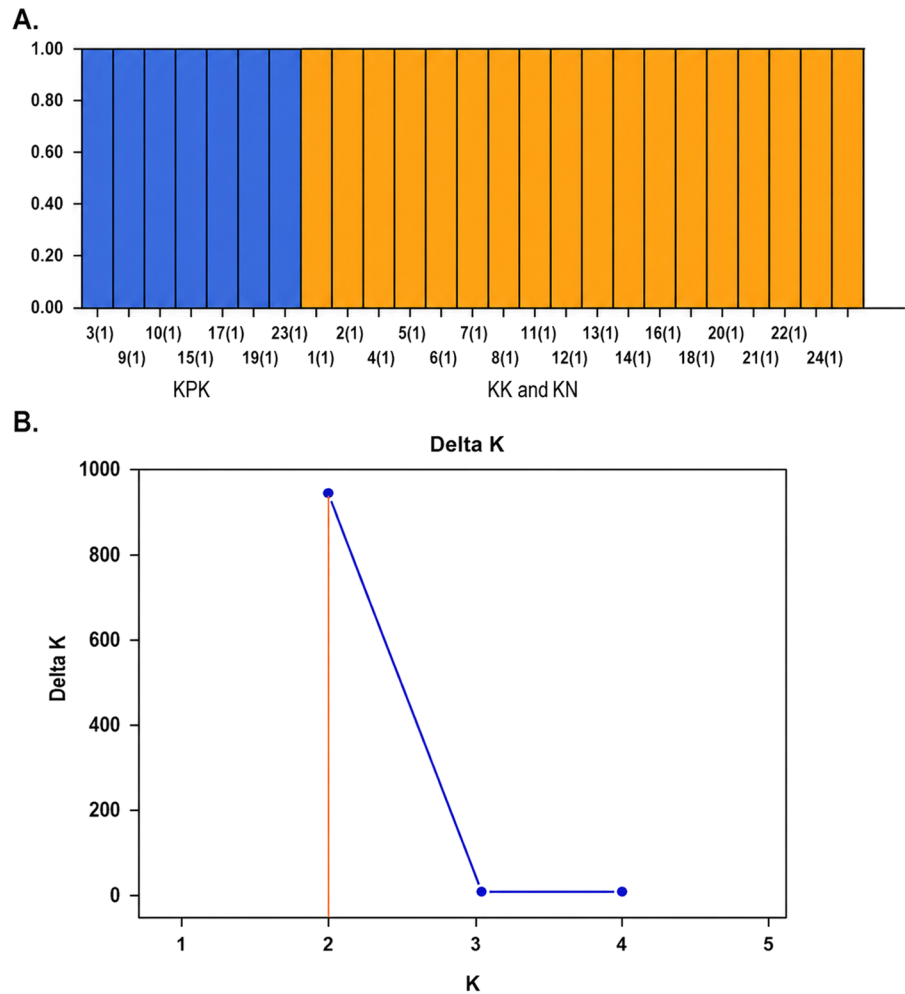
Genome-wide population branch statistic (PBS) identified a limited number of loci exhibiting elevated population-specific divergence across the three populations (Fig. 6). Elevated PBS values were restricted to discrete genomic regions, with the majority of loci showing low background differentiation. In each population, only a small subset of loci exceeded the 99th percentile threshold. Among the PBS outliers, some loci mapped near annotated genes, while others were located in regions lacking known gene annotations. In the KN population, one locus (Locus X) displaying elevated PBS was located near *col14a1b*, a gene involved in extracellular matrix organization and skeletal development. In the KK population, an elevated PBS signal at Locus VI was detected near a 5S rRNA gene cluster. In addition, two PBS outliers (Locus IV and Locus VII), identified in different populations, mapped to the same genomic region, corresponding to the uncharacterized gene ENSORLG00000028906. Overall, the strongest PBS outliers were predominantly located in noncoding or poorly annotated regions of the genome.

Discussion

Phylogenetic placement of *G. stocki*

The mitochondrial phylogenetic analysis provides additional insight into the evolutionary placement of *G. stocki* within Sisoridae. All three *G. stocki* individuals formed a well-supported monophyletic clade, consistent with their taxonomic identity and supporting the reliability of the assembled mitochondrial genomes. The close relationship between *G. stocki* and species such as *G. trilineatus* and *G. annandalei* is consistent with previously reported relations within the genus. Notably, *G. dolichonema* was recovered within the broader *Glyptothorax* clade, consistent with recent mitochondrial genome-based studies suggesting potential polyphyly of the genus (Schedel et al., 2022). Although the present study focuses primarily on genome-wide nuclear variation, the mitochondrial

Fig. 4 Population structure of *G. stocki* inferred from STRUCTURE analysis. (A) Bar plot at the optimal $K=2$ shows individuals assigned to two genetic clusters with no admixture: one specific to the KPK population and another comprising the KK and KN populations. (B) The Evanno plot (ΔK) confirms $K=2$ as the most likely number of genetic clusters



phylogeny complements these findings by placing *G. stocki* within a broader evolutionary framework.

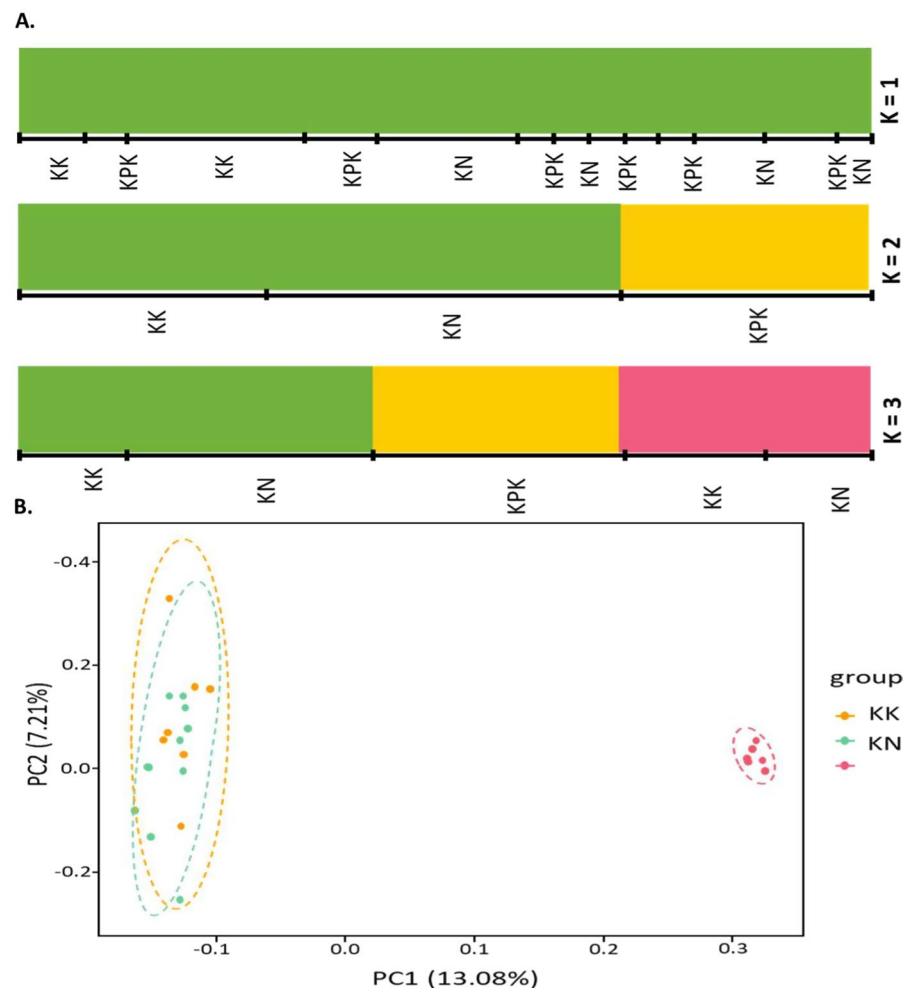
Genomic evidence for two evolutionarily significant units in *G. stocki*

This study presents the first genome-wide assessment of population structure and genetic diversity in the endangered endemic catfish *G. stocki* across its known range in Pakistan. Our analysis, based on high-quality SNPs, provides consistent and robust evidence for the existence of two genetically distinct populations: one comprising individuals from the Swat River in Khyber Pakhtunkhwa (KPK), and a second encompassing the closely related populations from the Poonch (KK) and Neelum (KN) Rivers in

Azad Jammu & Kashmir. This clear bifurcation, supported by STRUCTURE, ADMIXTURE, PCA, and phylogenetic analysis, is further validated by significant pairwise F_{ST} values. The pronounced differentiation between KPK and the Kashmir populations ($F_{ST}=0.134-0.175$) contrasts sharply with the negligible divergence between KK and KN ($F_{ST}=0.003$), underscoring a fundamental partition within the species with critical implications for its conservation.

Comparable patterns of phylogeographic subdivision have been reported in other South Asian freshwater fishes. For instance, population genetic studies on the large riverine catfish *Wallago attu* (Bloch & Schneider, 1801) in Pakistan revealed strong differentiation among Indus River subpopulations and low nucleotide diversity, largely

Fig. 5 Population genetic structure of *G. stocki*. **(A)** ADMIXTURE bar plots for $K = 1$ to $K = 3$. The optimal clustering at $K = 2$ supports two distinct genetic groups: KPK and a combined KK/KN population. **(B)** Principal component analysis (PCA) plot showing clear separation of the KPK population from a cluster containing the KK and KN populations



driven by hydrological isolation and anthropogenic modification of habitats (Sherzada et al., 2024). Similarly, *Tor putitora* (Hamilton, 1822) exhibits distinct genetic lineages across Himalayan drainages, corresponding to historical barriers and modern dam constructions (Yadav et al., 2020). The observed population structure may be influenced by both natural riverine features and anthropogenic barriers. In particular, the Mangla Dam represents a major artificial barrier that may restrict gene flow between these river systems.

Drivers of population divergence

The observed genetic structure is a likely consequence of both historical biogeography and contemporary

anthropogenic fragmentation. The primary separation reflects the distinct hydrographic basins these populations inhabit. The KK and KN populations are tributaries of the Jhelum River, which feeds into the Indus River system via the massive Mangla Dam reservoir. In contrast, the KPK population belongs to the Swat River system, a tributary of the Kabul River, representing a separate sub-basin within the greater Indus catchment (Mirza, 1994). This natural hydrological separation has likely existed for a significant evolutionary period, potentially exacerbated by historical geological events such as river captures and Pleistocene climatic shifts that have shaped vicariance in other Himalayan aquatic taxa (He & Chen, 2006; Duan et al., 2009).

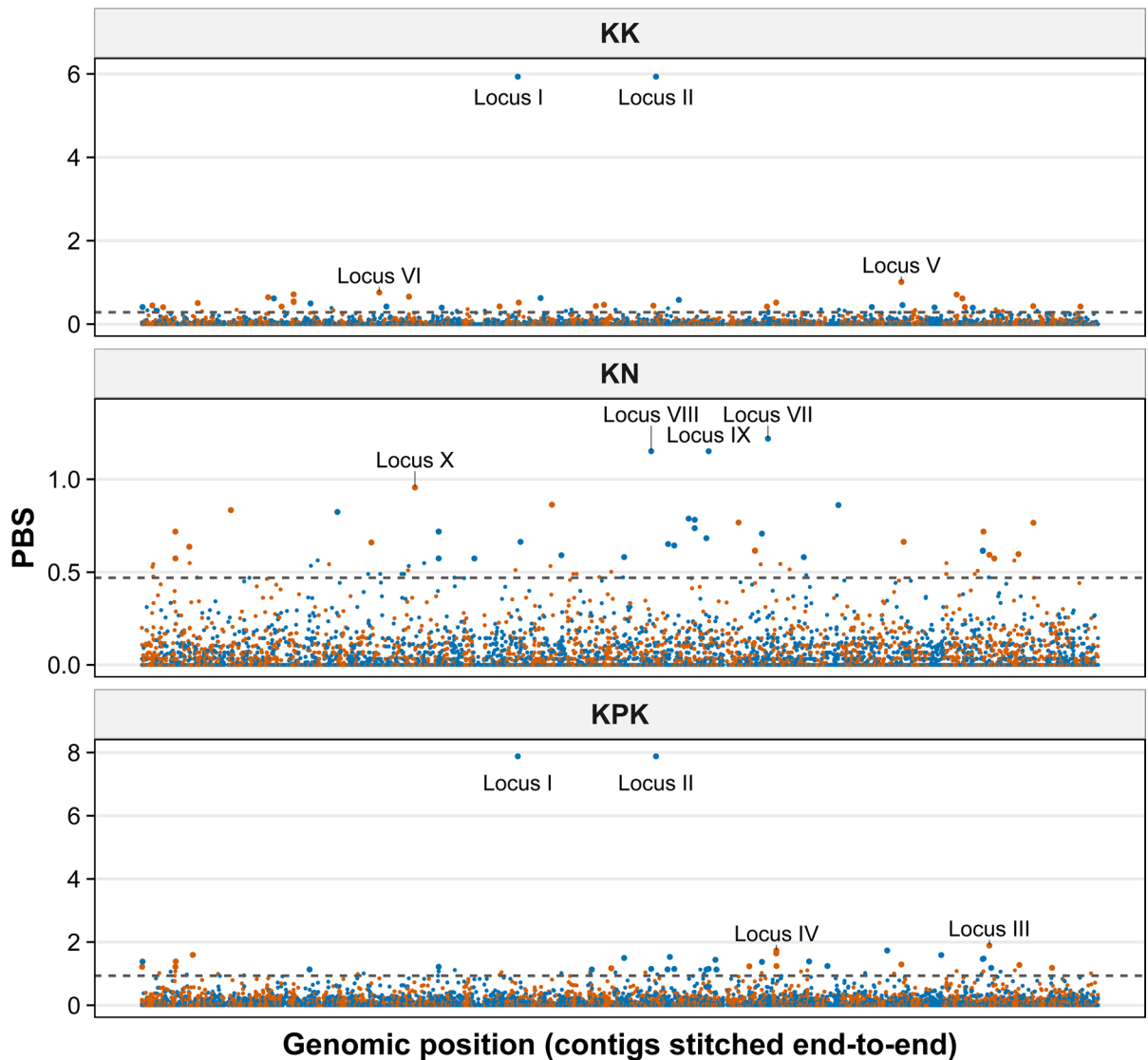


Fig. 6 Genome-wide population branch statistic (PBS) scans across three populations. Manhattan-style plots show PBS values calculated for each SNP across the genome for the KK, KN, and KPK populations. SNPs are plotted according to cumulative genomic position, with contigs stitched end-to-end

and alternating colors indicating adjacent contigs. Horizontal dashed lines denote the 99th percentile of the empirical PBS distribution within each population. Labeled points indicate the highest PBS scoring loci in each population based on predefined locus identifiers

However, this natural division has been severely amplified by modern infrastructure. The Mangla Dam constitutes a formidable barrier to gene flow, disrupting longitudinal connectivity and isolating upstream populations. Such anthropogenic barriers promote genetic divergence through genetic drift in fragmented subpopulations, a well-documented phenomenon in riverine ecosystems worldwide (Nilsson et al., 2005; Zarfl et al., 2015). Our findings thus align

with studies on other South Asian fishes, such as the golden mahseer (*T. putitora*), where mitochondrial DNA reveals phylogeographic breaks correlated with both natural drainage basins and dams (Yadav et al., 2020). The genomic signature in *G. stocki* therefore bears the dual imprint of deep-time evolutionary history and recent human impact, highlighting that effective conservation must address both legacies.

Population branch statistic (PBS) and environment-associated genomic differentiation

Genome-wide analysis of population branch statistic (PBS) values revealed that population-specific divergence in *G. stocki* is highly localized, with only a small number of discrete genomic regions showing elevated PBS values against a background of generally low differentiation. This pattern may be characteristic of taxa in which adaptation occurs through selective pressures acting on specific loci rather than across the entire genome (Shpak et al., 2025). The observed distribution of PBS peaks is distinct among populations but spatially restricted, which may point to independent adaptive responses to local environmental conditions or selective regimes. Similar patterns have been reported in other systems, such as in migratory birds and marine species, where local adaptation arises through limited genomic islands of divergence (Tigano & Friesen, 2016; Hämälä & Savolainen, 2019).

Among the identified PBS outliers, in the KN population, an elevated PBS signal near *coll4a1b*, a gene associated with extracellular matrix organization and skeletal development, suggests possible selection on traits related to morphology or structural adaptation. Functional parallels have been observed in adaptive divergence studies of vertebrates, where genes involved in collagen synthesis and skeletal development frequently contribute to environmental adaptation (Cuevas et al., 2021). Similarly, the elevated PBS signal detected near a 5S rRNA gene cluster in the KK population may indicate selection on ribosomal function or gene regulation, consistent with reports that rRNA cluster variation can evolve rapidly under environmental stress or metabolic demand (López-Maury et al., 2008). Intriguingly, two PBS outliers mapped to the same genomic region corresponding to the uncharacterized gene ENSORLG00000028906, suggesting that this locus may represent a previously unrecognized target of selection. The repeated signal across populations implies a conserved selective mechanism, potentially involving regulatory elements or structural genomic features. The predominance of PBS outliers in noncoding or poorly annotated regions suggests that adaptive divergence often involves changes in regulatory sequences that

fine-tune gene expression, rather than mutations in protein-coding regions (Booker et al., 2021).

Population branch statistics and related branch rescaled methods have been validated as effective for identifying loci associated with local adaptation across populations using SNP data, often outperforming classical F_{ST} -based scans in detecting population-specific sweeps from genome-wide SNP sets (Nandanpawar et al., 2024). This further supports the interpretation of discrete PBS peaks in *G. stocki* as candidate signals of population-specific differentiation rather than genome-wide divergence/adaptation. Future functional assays and transcriptomic analysis will be critical to validate the adaptive relevance of these loci and elucidate their roles in population-specific environmental responses.

Contrasting genetic health signals demand tailored management

Beyond revealing population structure, our analysis uncovered critical differences in genetic health indicators, necessitating distinct conservation strategies for each unit. The KPK population displays a concerning genetic profile characterized by markedly low nucleotide diversity ($\pi=0.0256$) compared to the Kashmir populations (KK $\pi=0.0697$; KN $\pi=0.0646$). Low nucleotide diversity directly reduces adaptive potential, increasing vulnerability to environmental change and stochastic events (Frankham, 2005). This pattern is characteristic of small, isolated populations that have experienced a reduction in effective population size, possibly due to a historical bottleneck or sustained habitat constraints. Paradoxically, this same KPK population exhibits a heterozygote excess ($F_{IS}=-0.13$). While this could theoretically indicate balancing selection, a more plausible explanation for an isolated population with low overall diversity is a recent bottleneck. In such a scenario, a rapid reduction in population size reduces allelic diversity (lowering π) faster than heterozygosity, leading to a transient excess of heterozygotes until a new equilibrium is reached. This pattern has been empirically observed in freshwater fish populations exhibiting low genetic variation as a consequence of recent demographic decline (e.g., *Labeo rohita* (Hamilton, 1822) showing heterozygosity excess and low allelic richness consistent with a bottleneck) (Nandanpawar

et al., 2024). Importantly, estimates of individual genomic inbreeding based on excess homozygosity (PLINK F) indicate that genomic inbreeding in KPK is generally low to moderate (median $F=0.03$), but with substantial variation among individuals. This combination signals a population that is genetically depauperate and potentially at high risk, despite a temporary surplus of heterozygotes that may mask the severity of its situation.

In contrast, the Kashmir populations (KK and KN) present a different challenge. They retain comparatively high nucleotide diversity, suggesting a historically larger effective population size and a healthier genetic base. However, both show positive F_{IS} values (KK $F_{IS}=0.22$; KN $F_{IS}=0.11$), indicating a heterozygote deficit. Estimates of individual genomic inbreeding corroborate this pattern, with elevated median genomic inbreeding in KK (median $F=0.33$) and intermediate levels in KN (median $F=0.15$). This suggests that within the Jhelum system, even though KK and KN are genetically similar, there may be finer-scale barriers or reduced dispersal leading to mating among relatives. Positive F_{IS} values and heterozygote deficiencies are commonly interpreted as signals of inbreeding or mating among related individuals within small, fragmented populations, which has been documented in other freshwater fish where restricted gene flow and small local effective sizes contribute to elevated inbreeding (Castric et al., 2002). In summary, the high overall species-level nucleotide diversity ($\pi=0.1196$) indicates that *G. stocki* retains significant ancestral variation, but this is now partitioned into units facing distinct threats: the KPK population suffers from a lack of genetic variation, while the Kashmir populations risk inbreeding depression despite their broader genetic base.

Similar demographic signatures have been observed in Pakistani populations of *W. attu*, where microsatellite analysis identified recent bottlenecks and reduced allelic richness, consistent with population contraction and isolation (Zafar, 2016). These findings suggest that small, isolated riverine populations across Pakistan frequently experience loss of variation and transient heterozygote excess, analogous to the genetic condition of the KPK *G. stocki* population.

Conservation implications

The consistent genetic evidence mandates the formal recognition of at least two management units (MUs) for *G. stocki*: a KPK MU (Swat River) and a Kashmir MU (KK and KN populations combined). Treating these as separate units is essential to prevent outbreeding depression from mixing genetically distinct groups and to allow for targeted interventions. For the KPK MU, the priority is to halt further genetic erosion and restore genetic variation. Strategies must include stringent habitat protection to support population growth. Given its low diversity, more active interventions like genetic rescue through carefully managed translocation of individuals from the genetically healthy Kashmir MU could be considered in the future, though this would require rigorous risk assessment, including evaluation of the evolutionary history and duration of population separation, particularly whether divergence predates recent anthropogenic barriers such as the construction of the Mangla Dam.

For the Kashmir MU, the goal is to mitigate inbreeding by restoring connectivity and gene flow within the Jhelum River system. Conservation efforts should advocate for environmental flow regimes and the construction of fish passes at barriers like the Mangla Dam to facilitate natural movement. Additionally, habitat rehabilitation in the Poonch and Neelum tributaries is crucial to support larger, more interconnected populations. Similarly, conservation frameworks have been proposed for other regional fish species. The golden mahseer (*Tor putitora*) in Himalayan rivers has been managed using drainage-based conservation units to preserve genetic integrity while maintaining adaptive potential across basins (Yadav et al., 2020). Likewise, reviews of Pakistan's freshwater biodiversity emphasize that effective management must align with river-specific evolutionary histories and address fragmentation due to dams and water abstraction (Ali et al., 2025). Adopting such an integrated strategy for *G. stocki* will ensure that each management unit retains its distinct genetic legacy and adaptive capacity.

Conclusion

Overall, the present study is the first and most detailed investigation of the population structure and genetic variation of the endemic catfish *G. stocki* in Pakistan. Incorporating high-resolution SNP data, phylogeny, and population genetic methods, two genetically distinct lineages were revealed in the tributaries of the Kashmir region and the KPK basin. Diversity in nucleotides and inbreeding coefficients showed significant variation: Kashmir populations exhibited high values of inbreeding coefficients, while KPK populations revealed a heavy excess of heterozygotes along with low nucleotide diversity. The implications of these findings are the need to formulate at least two management units and strategies that promote gene flow, reduce the risks of inbreeding, and maintain the local adaptive potential.

Acknowledgements We would like to express our sincere gratitude to Dr. Muhammad Jahangir, University of Okara, for his valuable support during this research. We are also thankful to the Director (Naeem Dar) and Deputy Director (Sakhi-uz-Zaman) of Wildlife and Fisheries, Azad Jammu & Kashmir, for their cooperation and assistance in facilitating and supporting the collection of samples, without which this study would not have been possible.

Author contributions Mohsan Raza was involved in conceptualization, writing-original draft, writing—reviewing and editing, visualization, methodology, data curation, formal analysis, and validation. Arif Muhammad Khan was responsible for writing-reviewing and editing, methodology, and resources. Chenhong Li took part in conceptualization, supervision, writing-reviewing and editing, methodology, data curation, and validation.

Funding The present work was funded by the Science and Technology Commission of Shanghai Municipality under the Project numbers STCSM 19410740500; 19050501900.

Data availability The raw sequencing data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under the BioProject accession number PRJNA1404215. The three assembled mitochondrial genome sequences of *G. stocki* are available on GenBank under the accession numbers PZ295967-PZ295969.

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval Permission for data and sample collection of fish species for scientific purposes was obtained from the Department of Wildlife and Fisheries, Azad Jammu & Kash-

mir (Permission No. 1230–32). All experimental procedures, including specimen handling, followed the guidelines and protocols approved by the Ethics Committee for the Use of Animal Subjects, Approval No. SHOU-DW-2021–090 of August 09, 2021, Shanghai Ocean University, China.

References

- Ali, G., S. Abbas, S. Nagai, N. Mohd Arshad & S. Bhassu, 2025. Threats of climate change to freshwater ecosystems in Pakistan: eDNA monitoring will be the next-generation tool used in biodiversity, conservation, and management. *Biology* 14: 1191. <https://doi.org/10.3390/biology14091191>.
- Allendorf, F. W., G. H. Luikart & S. N. Aitken, 2012. Conservation and the Genetics of Populations, John Wiley & Sons.
- Bleeker, P., 1858. *Ichthyologiae archipelagi indici prodromus*. Lange et Company.
- Bloch, M. E., J. F. Hennig, & J. G. Schneider, 1801. ME Blochii. *Systema ichthyologiae iconibus CX illustratum/post obitum auctoris opus inchoatum absoluit, correxuit, interpolavit Jo. Gottlob Schneider, Saxo. Sumtibus auctoris impressum et Bibliopolio Sanderiano commissum*, (No Title).
- Blyth, E., 1860. Report on some fishes received chiefly from the Sitang River and its tributary streams. *The Journal of the Asiatic Society of Bengal* 29: 138–174.
- Booker, T. R., S. Yeaman & M. C. Whitlock, 2021. Global adaptation complicates the interpretation of genome scans for local adaptation. *Evolution Letters* 5: 4–15. <https://doi.org/10.1002/evl3.208>.
- Butchart, S. H. M., M. Walpole, B. Collen, A. Van Strien, J. P. W. Scharlemann, R. E. A. Almond, J. E. M. Baillie, B. Bomhard, C. Brown, J. Bruno, K. E. Carpenter, G. M. Carr, J. Chanson, A. M. Chenery, J. Csirke, N. C. Davidson, F. Dentener, M. Foster, A. Galli, J. N. Galloway, P. Genovesi, R. D. Gregory, M. Hockings, V. Kapos, J.-F. Lamarque, F. Leverington, J. Loh, M. A. McGeoch, L. McRae, A. Minasyan, M. H. Morcillo, T. E. E. Oldfield, D. Pauly, S. Quader, C. Revenga, J. R. Sauer, B. Skolnik, D. Spear, D. Stanwell-Smith, S. N. Stuart, A. Symes, M. Tierney, T. D. Tyrrell, J.-C. Vié & R. Watson, 2010. Global biodiversity: indicators of recent declines. *Science* 328: 1164–1168. <https://doi.org/10.1126/science.1187512>.
- Castric, V., L. Bernatchez, K. Belkhir & F. Bonhomme, 2002. Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus Fontinalis* Mitchell (Pisces, Salmonidae): a test of alternative hypotheses. *Heredity* 89: 27–35. <https://doi.org/10.1038/sj.hdy.6800089>.
- Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores & W. A. Cresko, 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22: 3124–3140. <https://doi.org/10.1111/mec.12354>.
- Chen, C., H. Chen, Y. Zhang, H. R. Thomas, M. H. Frank, Y. He & R. Xia, 2020. TBtools: an integrative toolkit

- developed for interactive analyses of big biological data. Elsevier. *Molecular Plant* 13: 1194–1202.
- Cincotta, R. P., J. Wisniewski & R. Engelman, 2000. Human population in the biodiversity hotspots. *Nature* 404: 990–992. <https://doi.org/10.1038/35010105>.
- Collen, B., F. Whittton, E. E. Dyer, J. E. M. Baillie, N. Cumberlidge, W. R. T. Darwall, C. Pollock, N. I. Richman, A. Soulsby & M. Böhm, 2014. Global patterns of freshwater species diversity, threat and endemism. *Global Ecology and Biogeography* 23: 40–51. <https://doi.org/10.1111/geb.12096>.
- Cuevas, A., M. Ravinet, G. Sætre & F. Eroukmanoff, 2021. Intraspecific genomic variation and local adaptation in a young hybrid species. *Molecular Ecology* 30: 791–809. <https://doi.org/10.1111/mec.15760>.
- Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, & 1000 Genomes Project Analysis Group, 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>.
- DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. Del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernysky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler & M. J. Daly, 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* 43: 491–498. <https://doi.org/10.1038/ng.806>.
- DeWoody, J. A., A. M. Harder, S. Mathur & J. R. Willoughby, 2021. The long-standing significance of genetic diversity in conservation. *Molecular Ecology* 30: 4147–4154. <https://doi.org/10.1111/mec.16051>.
- Dierckxsens, N., P. Mardulyn, & G. Smits, 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research* : gkw955. <https://doi.org/10.1093/nar/gkw955>.
- Duan, Z., K. Zhao, Z. Peng, J. Li, R. Diogo, X. Zhao & S. He, 2009. Comparative phylogeography of the Yellow River schizothoracine fishes (Cyprinidae): Vicariance, expansion, and recent coalescence in response to the quaternary environmental upheaval in the Tibetan Plateau. *Molecular Phylogenetics and Evolution* 53: 1025–1031. <https://doi.org/10.1016/j.ympev.2009.03.034>.
- Dudgeon, D., A. H. Arthington, M. O. Gessner, Z. Kawabata, D. J. Knowler, C. Lévêque, R. J. Naiman, A. Prieur-Richard, D. Soto, M. L. J. Stiassny & C. A. Sullivan, 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews* 81: 163–182. <https://doi.org/10.1017/S1464793105006950>.
- Evanno, G., S. Regnaut & J. Goudet, 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 14: 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.
- Frankham, R., 2003. Genetics and conservation biology. *Comptes Rendus. Biologies* 326: 22–29. [https://doi.org/10.1016/S1631-0691\(03\)00023-4](https://doi.org/10.1016/S1631-0691(03)00023-4).
- Frankham, R., 2005. Genetics and extinction. *Biological Conservation* 126: 131–140. <https://doi.org/10.1016/j.biocon.2005.05.002>.
- Funk, W. C., J. K. McKay, P. A. Hohenlohe & F. W. Allendorf, 2012. Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution* 27: 489–496. <https://doi.org/10.1016/j.tree.2012.05.012>.
- Gehrke, P. C., D. M. Gilligan & M. Barwick, 2002. Changes in fish communities of the Shoalhaven River 20 years after construction of Tallowa Dam, Australia. *River Research and Applications* 18: 265–286. <https://doi.org/10.1002/rra.669>.
- Hamilton, F., 1822. An account of the fishes found in the river Ganges and its branches. Archibald Constable.
- Hammer, M. P., C. M. Bice, A. Hall, A. Frears, A. Watt, N. S. Whiterod, L. B. Beheregaray, J. O. Harris, & B. P. Zampatti, 2013. Freshwater fish conservation in the face of critical water shortages in the southern Murray–Darling Basin, Australia. *Marine and Freshwater Research* 64: 807. <https://doi.org/10.1071/MF12258>.
- Harrisson, K. A., A. Pavlova, M. Telonis-Scott & P. Sunnucks, 2014. Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications* 7: 1008–1025. <https://doi.org/10.1111/eva.12149>.
- He, D. & Y. Chen, 2006. Biogeography and molecular phylogeny of the genus *Schizothorax* (Teleostei: Cyprinidae) in China inferred from cytochrome *b* sequences. *Journal of Biogeography* 33: 1448–1460. <https://doi.org/10.1111/j.1365-2699.2006.01510.x>.
- Hora, S. L., 1923. Notes on fishes in the Indian Museum-on the composite genus *Glyptosternon* McClelland. *Records of the Zoological Survey of India*: 01–44.
- Hussain, S. & S. P. S. Dutta, 2016. Fish fauna of tehsil mendhar, district-poonch, Jammu region, J&K state, with a new record of botiabirdi, ompokpabda and *Glyptothorax punjabiensis* from Poonch district. *Proceedings of the Zoological Society of India* 15: 1–11.
- Hämälä, T. & O. Savolainen, 2019. Genomic patterns of local adaptation under gene flow in *Arabidopsis lyrata*. *Molecular Biology and Evolution* 36: 2557–2571. <https://doi.org/10.1093/molbev/msz149>.
- IUCN, 2020. *Glyptothorax stocki*: Rafique, M., Narejo, N., Qadir, A., Hasan, Z., Ahmad, I., Ramzan, M., Daniels, A. & Dahanukar, N. The IUCN Red List of Threatened Species 2021: e.T128723276A128723295.
- Ishaq, M., S. Khan, J. Khan, N. Akhtar & K. Saeed, 2014. Study on ichthyofaunal biodiversity of River Swat. *World Journal of Fish and Marine Sciences* 6: 313–318.
- Iyyappan, S., S. Ghara, I. A. Bhat, I. A. Khan, & M. A. Rather, 2026. Uncovering evolutionary and phylogenetic relationships in *Glyptothorax* species through comparative mitochondrial genomics. Elsevier. *Journal of Genetic Engineering and Biotechnology* 24: 100629.
- Javed, M. N., S. Kalsoom & K. Pervaiz, 2013. Catfishes of the genus *Glyptothorax* Blyth (Pisces: Sisoridae) from Pakistan. *Biological Society of Pakistan* 59: 69–83.
- Jiang, W., H. H. Ng, J. Yang, & X. Chen, 2012. A taxonomic review of the catfish identified as *Glyptothorax zanaensis* (Teleostei: Siluriformes: Sisoridae), with the descriptions of two new species. 165.
- Kardos, M., H. R. Taylor, H. Ellegren, G. Luikart & F. W. Allendorf, 2016. Genomics advances the study of

- inbreeding depression in the wild. *Evolutionary Applications* 9: 1205–1218. <https://doi.org/10.1111/eva.12414>.
- Kinsella, R. J., A. Kahari, S. Haider, J. Zamora, G. Proctor, G. Spudich, J. Almeida-King, D. Staines, P. Derwent, A. Kerhornou, P. Kersey, & P. Flicek, 2011. Ensembl BioMart: a hub for data retrieval across taxonomic space. *Database* 2011: bar030–bar030. <https://doi.org/10.1093/database/bar030>.
- Knaus, B. J. & N. J. Grünwald, 2017. vcfr : a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources* 17: 44–53. <https://doi.org/10.1111/1755-0998.12549>.
- Lakra, W. S., U. K. Sarkar, R. S. Kumar, A. Pandey, V. K. Dubey & O. P. Gusain, 2010. Fish diversity, habitat ecology and their conservation and management issues of a tropical River in Ganga basin, India. *The Environmentalist* 30: 306–319. <https://doi.org/10.1007/s10669-010-9277-6>.
- Li, H. & R. Durbin, 2009. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25: 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- López-Maury, L., S. Marguerat & J. Bähler, 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics* 9: 583–593. <https://doi.org/10.1038/nrg2398>.
- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17: 10. <https://doi.org/10.14806/ej.17.1.200>.
- McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernysky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly & M. A. DePristo, 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20: 1297–1303. <https://doi.org/10.1101/gr.107524.110>.
- Meador, M. R. & R. M. Goldstein, 2003. Assessing water quality at large geographic scales: relations among land use, water physicochemistry, riparian condition, and fish community structure. *Environmental Management* 31: 504–517. <https://doi.org/10.1007/s00267-002-2805-5>.
- Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. Von Haeseler & R. Lanfear, 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 37: 1530–1534. <https://doi.org/10.1093/molbev/msaa015>.
- Mirza, M. R., 1994. Geographical distribution of freshwater fishes in Pakistan: a review. *Punjab University Journal of Zoology* 9: 93–108.
- Mirza, M. & H. Nijssen, 1978. *Glyptothorax stocki*, a new sisorid catfish from Pakistan and Azad Kashmir (Siluriformes, Sisoridae). *Bulletin Zoologisch Museum* 6: 79–85.
- Mirza, M. R., & K. M. Kashmiri, 1971. A note on the fishes of the genus *Glyptothorax Blyth* (Osteichthyes, Sisoridae) from West Pakistan with the description of a new subspecies.
- Moyle, L. C., Ecology and Evolutionary Biology Dept., Brown Univ, USA, B. R. Hudgens, Dept. of Biology, Duke Univ, USA, W. F. Morris, & Dept. of Biology, Duke Univ, USA, 2003. Conservation genetics in the recovery of endangered animal species: a review of US endangered species recovery plans (1977–1998). *Animal Biodiversity and Conservation* 26: 85–95. <https://doi.org/10.32800/abc.2003.26.2085>.
- Namkoong, G., 2000. Genetic conservation strategies: Who needs data? In Mátyás, C. (ed), *Forest Genetics and Sustainability* Springer, Netherlands, Dordrecht: 211–216.
- Nandanpawar, P., B. Sahoo, L. Sahoo, K. Murmu, D. Reang, A. P. Kumar, A. Chaudhari & P. Das, 2024. Unveiling population structure and selection signatures of riverine and genetically improved rohu, *Labeo rohita* using genome wide SNPs. *Molecular Biology Reports* 51: 926. <https://doi.org/10.1007/s11033-024-09866-y>.
- Nei, M., 1987. *Molecular Evolutionary Genetics*, Columbia University Press.
- Nei, M. & W. H. Li, 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences* 76: 5269–5273. <https://doi.org/10.1073/pnas.76.10.5269>.
- Ng, H. H., & R. K. Hadiaty, 2009. *Glyptothorax ketambe* a new catfish (Teleostei: Sisoridae) from northern Sumatra. *Zootaxa* 2085: 61–68. <https://doi.org/10.11646/zootaxa.2085.1.5>.
- Nilsson, C., C. A. Reidy, M. Dynesius & C. Revenga, 2005. Fragmentation and flow regulation of the World's Large River Systems. *Science* 308: 405–408. <https://doi.org/10.1126/science.1107887>.
- Paradis, E., 2010. Pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26: 419–420. <https://doi.org/10.1093/bioinformatics/btp696>.
- Paradis, E. & K. Schliep, 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Oxford University Press. *Bioinformatics* 35: 526–528.
- Pavlova, A., L. B. Beheregaray, R. Coleman, D. Gilligan, K. A. Harrison, B. A. Ingram, J. Kearns, A. M. Lamb, M. Lintermans, J. Lyon, T. T. T. Nguyen, M. Sasaki, Z. Tonkin, J. D. L. Yen & P. Sunnucks, 2017. Severe consequences of habitat fragmentation on genetic diversity of an endangered Australian freshwater fish: a call for assisted gene flow. *Evolutionary Applications* 10: 531–550. <https://doi.org/10.1111/eva.12484>.
- Pierson, J. C., D. J. Coates, J. G. B. Oostermeijer, S. R. Beissinger, J. G. Bragg, P. Sunnucks, N. H. Schumaker & A. G. Young, 2016. Genetic factors in threatened species recovery plans on three continents. *Frontiers in Ecology and the Environment* 14: 433–440. <https://doi.org/10.1002/fee.1323>.
- Poff, N. L. & J. D. Allan, 1995. Functional organization of stream fish assemblages in relation to hydrological variability. *Ecology* 76: 606–627. <https://doi.org/10.2307/1941217>.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. De Bakker, M. J. Daly & P. C. Sham, 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* 81: 559–575. <https://doi.org/10.1086/519795>.
- Rahman, M. I., R. Pervaiz & M. Umar, 2025. Proximate body composition of fish species of the River Panjkora, Dir Upper Khyber Pakhtunkhwa, Pakistan. *Annual Methodological Archive Research Review* 3: 239–250.

- Rambaut, A., 2009. FigTree. Tree Figure Drawing Tool. Institute of Evolutionary Biology, University of Edinburgh. <http://tree.bio.ed.ac.uk/software/figtree/>.
- Regan, C. T., 1905. XIII – A Synopsis of the Species of the Silurid Genera *Parexostoma*, *Chimarrichthys* and *Exostoma*. *Annals and Magazine of Natural History* 15: 182–185. <https://doi.org/10.1080/03745480509442815>.
- Sala, O. E., F. Stuart Chapin, J. J. Iii, E. Armesto, J. Berlow, R. Bloomfield, E. Dirzo, L. F. Huber-Sanwald, R. B. Huenneke, A. Jackson, R. Kinzig, D. M. Leemans, H. A. Lodge, M. Mooney, N. L. Oesterheld, M. T. Poff, B. H. Sykes, M. Walker, Walker & D. H. Wall, 2000. Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774. <https://doi.org/10.1126/science.287.5459.1770>.
- Sayyadzadeh, G., F. Zarei, & H. R. Esmaili, 2022. Glyptothorax (Teleostei: Sisoridae) from the middle east: an integrated molecular and morphological insight into its taxonomic diversity. *multidisciplinary digital publishing institute. Diversity* 14: <https://doi.org/10.3390/d14100884>.
- Schedel, F. D. B., A. Chakona, B. L. Sidlauskas, M. O. Popoola, N. Usimesa Wingi, D. Neumann, E. J. W. M. N. Vreven & U. K. Schliwen, 2022. New phylogenetic insights into the African catfish families Mochokidae and Austroglanididae. *Journal of Fish Biology* 100: 1171–1186. <https://doi.org/10.1111/jfb.15014>.
- Sherzada, S., N. Hussain, A. Hussain, M. A. M. El-Tabakh & S. A. Khan, 2024. Diversity and genetic structure of freshwater shark *Wallago attu*: an emerging species of commercial interest. *Environmental Science and Pollution Research* 31: 15571–15579. <https://doi.org/10.1007/s11356-024-32117-3>.
- Shpak, M., K. N. Lawrence, & J. E. Pool, 2025. The precision and power of population branch statistics in identifying the genomic signatures of local adaptation. *Genome Biology and Evolution* 17: evaf080. <https://doi.org/10.1093/gbe/evaf080>.
- Taylor, R. H., N. Dussex, & Y. van Heezik, 2017. Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Global Ecology and Conservation* 10: 231–242. <https://doi.org/10.1016/j.gecco.2017.04.001>.
- Thomson, A. W., & L. M. Page, 2006. Genera of the Asian catfish families Sisoridae and Erethistidae (Teleostei: Siluriformes). *Zootaxa* 1345: <https://doi.org/10.11646/zootaxa.1345.1.1>.
- Tigano, A. & V. L. Friesen, 2016. Genomics of local adaptation with gene flow. *Molecular Ecology* 25: 2144–2164. <https://doi.org/10.1111/mec.13606>.
- Ullah, M., A. M. Yousafzai & A. Muhammad, 2025. First comprehensive study on ichthyofaunal diversity of river Khurram Khyber Pakhtunkhwa Pakistan. *Journal of Wildlife and Biodiversity* 9: 300–313.
- Van Der Auwera, G. A., M. O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, & M. A. DePristo, 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics* 43: <https://doi.org/10.1002/0471250953.bi1110s43>.
- Weir, B. S. & C. C. Cockerham, 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358. <https://doi.org/10.2307/2408641>.
- Wickham, H., 2016. ggplot2, Springer International Publishing, Cham:
- Yadav, P., A. Kumar, S. A. Hussain & S. K. Gupta, 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: e0234377. <https://doi.org/10.1371/journal.pone.0234377>.
- Yelmen, B., D. Marnetto, L. Molinaro, R. Flores, M. Mondal, & L. Pagani, 2021. Improving selection detection with population branch statistic on admixed populations. *Genome Biology and Evolution* 13: evab039. <https://doi.org/10.1093/gbe/evab039>.
- Yi, X., Y. Liang, E. Huerta-Sanchez, X. Jin, Z. X. P. Cuo, J. E. Pool, X. Xu, H. Jiang, N. Vinckenbosch, T. S. Korneliusen, H. Zheng, T. Liu, W. He, K. Li, R. Luo, X. Nie, H. Wu, M. Zhao, H. Cao, J. Zou, Y. Shan, S. Li, Q. Yang, P. Asan, G. Ni, J. Tian, X. Xu, T. Liu, R. Jiang, G. Wu, M. Zhou, J. Tang, T. Qin, S. Wang, G. Li, Feng, J. Huang, W. Luosang, F. Wang, Y. Chen, X. Wang, Z. Zheng, Z. Li, G. Bianba, X. Yang, S. Wang, G. Tang, Y. Gao, Z. Chen, L. Luo, Z. Gusang, Q. Cao, W. Zhang, X. Ouyang, H. Ren, H. Liang, Y. Zheng, J. Huang, L. Li, K. Bolund, Y. Kristiansen, Y. Li, X. Zhang, R. Zhang, S. Li, H. Li, R. Yang, J. Wang, Nielsen & J. Wang, 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329: 75–78. <https://doi.org/10.1126/science.1190371>.
- Yuan, H., C. Atta, L. Tornabene & C. Li, 2019. Assexon: assembling exon using gene capture data. *Evolutionary Bioinformatics* 15: 1176934319874792. <https://doi.org/10.1177/1176934319874792>.
- Yuan, H., J. Jiang, F. A. Jiménez, E. P. Hoberg, J. A. Cook, K. E. Galbreath & C. Li, 2016. Target gene enrichment in the cyclophyllidean cestodes, the most diverse group of tapeworms. *Molecular Ecology Resources* 16: 1095–1106. <https://doi.org/10.1111/1755-0998.12532>.
- Zafar, M., 2016. Microsatellite DNA marker based genetic polymorphism in *Wallago attu*. *International Journal of Life Sciences* 4: 176–184.
- Zarfl, C., A. E. Lumsdon, J. Berlekamp, L. Tydecks & K. Tockner, 2015. A global boom in hydropower dam construction. *Aquatic Sciences* 77: 161–170. <https://doi.org/10.1007/s00027-014-0377-0>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.