Journal of Environmental Biology
p-ISSN: 0254-8704 • e-ISSN: 2394-0379 • CODEN: JEBIDP

JEB®

Original Research Paper
Volume 45 Issue 6 November 2024 pp. 705-713
DOI: http://doi.org/10.22438/jeb/45/6/MRN-5337

J. Environ. Biol., 2024
© Triveni Enterprises, Lucknow (India)
Journal website : www.jeb.co.in ★ E-mail : editor@jeb.co.in

Morphometric trait normalization, cluster analysis and molecular phylogenetic analysis in delimiting sympatric *Neolissochilus hexastichus* and *N. hexagonolepis* in Umngi River, Meghalaya (India)

M.A. Laskar[1]*, D. Sarma[2], R. Nath[3], R. Kharbani[3], N. Shahi[4], R.S. Haldar[4] and P.A. Ganie[4]

[1]*Biotechnology Department, St. Anthony's College, Shillong-793 001, India*
[2]*Division of Aquaculture, ICAR-Central Institute of Fisheries Education, Mumbai-400 061, India*
[3]*Fishery Science Department, St. Anthony's College, Shillong-793 001, India*
[4]*ICAR-Directorate of Coldwater Fisheries Research, Bhimtal-263 136, India*

Received: 12 April 2024 **Revised:** 03 June 2024 **Accepted:** 21 August 2024

***Corresponding Author Email:** drlaskar@anthonys.ac.in ***ORCiD:** https://orcid.org/0000-0001-6294-0450

Online Copy

Abstract

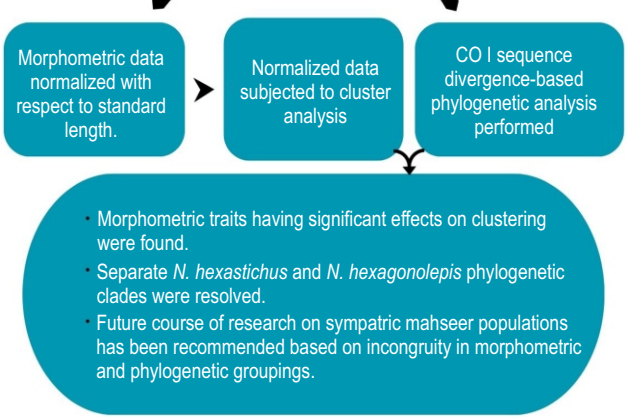
Aim: To assess the ability of an integrative approach comprising multivariate morphometric data analysis and mitochondrial cytochrome oxidase subunit I (CO I) sequence divergence analysis in delimiting sympatric *Neolissochilus hexastichus* (McClleland, 1839) and *N. hexagonolepis* (McClleland, 1839) mahseer populations in River Umngi, Meghalaya.

Methodology: Authoritative descriptions of mahseers were referred during the sampling of 40 individuals from River Umngi, Meghalaya. Multivariate morphometric data were normalized with respect to standard length. The normalization procedure maintained the particular shape factor for each sample. Normalised data were subjected to hierarchical clustering on principal components Analysis (HCPC) for grouping samples according to morphometric features. Genetic affinities based on CO I sequence divergence were assessed by Bayesian Inference (BI) and Automatic Barcode Gap Discovery (ABGD) protocols.

Results: The normalization protocols removed the variations in traits due to body size effects and facilitated the identification of allometric traits. HCPC analysis revealed morphometric features that had a significant impact on clustering. Nine *N. hexagonolepis* and 13 *N. hexastichus* individuals were identified by integrative morphological and molecular phylogenetic approach. The remaining 18 individuals had incongruence in their memberships in the morphometric and phylogenetic clusters.

Interpretation: The integrative approach, apart from delineating *N. hexagonolepis* and *N. hexastichus* congeners, highlighted the need to include nuclear genetic markers in future studies for confirming the presence of natural hybrids in sympatric populations of *Neolissochilus* mahseers.

Nelissochilus samples



Key words: Cluster, Morphometric, *Neolissochilus*, Normalization, Phylogenetics

How to cite: Laskar, M.A., D. Sarma, R. Nath, R. Kharbani, N. Shahi, R.S. Haldar and P.A. Ganie: Morphometric trait normalization, cluster analysis and molecular phylogenetic analysis in delimiting sympatric *Neolissochilus hexastichus* and *N. hexagonolepis* in Umngi River, Meghalaya (India). *J. Environ. Biol.*, **45,** 705-713 (2024).

Introduction

Cyprinid fishes belonging to the genus *Neolissochilus* are widely distributed in the freshwaters of South and South East Asia (Kechu *et al.,* 2023; Singh *et al.,* 2021). The term "Mahseer" used by anglers in the North-eastern Himalayan region recently refers to fish belonging to the genus *Neolissochilus* (Sarma *et al*., 2018). These fish are primarily found in the Brahmaputra River basin of North-east India, as well as in the Himalayan regions of Nepal and Bhutan, thriving at altitudes between 500 and 1500 m above sea level. Commonly known as chocolate mahseer or copper mahseer, they are renowned for their strong fighting abilities, reminiscent of a tiger, and their large scales and mouth. This makes the chocolate mahseer an iconic species in North-east India (Rainboth, 1985). Over twenty valid species of *Neolissochilus* exist across South-east Asia, recognized for their potential as game and sport fish, and are often sold at high market prices in an unregulated manner (Laskar *et al*., 2009). They are grouped under large-scaled carps known as mahseers. *Neolissochilus hexagonolepis* and *N. hexastichus* are popular food and sportfish (Sarma *et al.,* 2018). The natural population of these two species are declining in North east India, particularly in Meghalaya due to anthropogenic factors such as overexploitation, and damming of rivers and their siltation due to deforestation (Dash *et al.,* 2024; Chaudhury and Barbhuiya, 2020). Considerable success has been achieved from the conservation and breeding programmes for *N. hexagonolepis* that have been implemented in Meghalaya (Dash *et al.,* 2024) and such programmes will be necessary for *N. hexastichus* in the near future. These two species are present in the same drainage systems in Meghalaya that means in different streams of same river basin (Sarma *et al.,* 2018) and are genetically close (Singh *et al.,* 2021). Furthermore, according to Lalramliana *et al.* (2018) species of the genus *Neolissochilus* have overlapping morphological characters. Thus, in that context correct identification and delineation of two species is necessary for developing successful conservation and breeding programmes.

Species delimitation exercises yield accurate results when morphometric traits are normalized and subjected to multivariate analysis (Chan and Grismer, 2021). Recent studies on mahseer species delimitation (Sarmah *et al.,* 2024; Johnson *et al.,* 2023) used ratios of morphometric traits to total length, standard length and head length to remove body size effects. But ratios do not remove the allometric effects of size (Chen and Grismer, 2021). Furthermore, these workers performed univariate comparisons of traits. Hence, this study was conducted to test the effectives of multivariate analysis of normalized traits in delimiting sympatric populations *N. hexagonolepis* and *N. hexasticus.*

Hierarchical clustering on principal components was implemented for a comprehensive assessment of similarities and differences in the morphometric features of *Neolissochilus* samples in this study. Hierarchical clustering, employed to uncover similarities, and the hidden patterns among observations (Granato *et al.*, 2018) was considered appropriate for evaluating the inherent relationships in the multivariate morphometric data. The workflow included facilities for the statistical validation of cluster solutions, which enabled the results to be reported with confidence. However, since the cluster analysis lacked a phylogenetic basis, phylogenetic analysis based on mitochondrial cytochrome oxidase subunit I (COI) sequence divergence was also conducted. The congruence of the morphometric and phylogenetic analyses was assessed to infer *Neolissochilus* species delimitation. This study aimed to elucidate the population structure of *Neolissochilus* species in the Umngi River of Meghalaya, thereby generating scientific data to resolve species ambiguities in this crucial group of threatened species.

Materials and Methods

Forty mahseer individuals were sampled from two sites, Domsopian (25°18'30"N, 92°33'0''E) and Kenmynsaw Syntein (25°18'30"N, 91°33'30"E) of river Umngi of the Barak-Meghna-Surma drainage (Fig. 1) that flows through South west Khasi Hills district of Meghalaya state, India. Authoritative descriptions of mahseers (Lalramliana *et al.,* 2019, Laskar *et al.,* 2013; Talwar and Jhingran, 1991) were used during sampling, which was done during August – September, 2021. Mahseers of sub-adult stages (standard length, 51.14 mm to 144.17 mm) were sampled with the help of local communities using angling and cast net method. Measurements of standard length and 19 additional traits (Table 1) were recorded to the nearest millimeter with a digital slide calliper. These traits were among those mentioned in the authoritative descriptions of mahseers. Meristic data were also recorded for all 40 individuals. All samples were allotted isolate codes and preserved in 70% ethanol after collecting fin clips from pectoral fins and deposited in the fish museum of Fishery Department of St. Anthony's College, Shillong, Meghalaya, India. Statistical analyses of the present study were done with R statistical computing environment version 4.2.2 (R Core Team, 2023).

Normalization of morphometric traits: The procedure of Lleonart *et al.* (2000) as implemented in the R software package allomr (Schar, 2022) was used to remove the effects of body size on morphometric traits and to identify allometric traits. Normalization of 19 morphometric traits (Table 1) was done with respect to standard length (SL) measurements of 40 mahseer individuals. The dataframe for normalization comprised forty rows, one for each mahseer sample and twenty columns, one for values of each morphometric trait of these samples.

Normalization of morphometric traits with respect to standard length (SL) was done as:

$$Y_i^* = Y_i [X_0 / X_i]^b \quad (1)$$

where, Y_i^* is a normalized trait of individuali; X_i, is the standard length (SL) of the individual, Y_i is a trait of the individual subjected to normalization and X_0 is the mean standard length of forty mahseer individuals; b is the allometric exponent whose

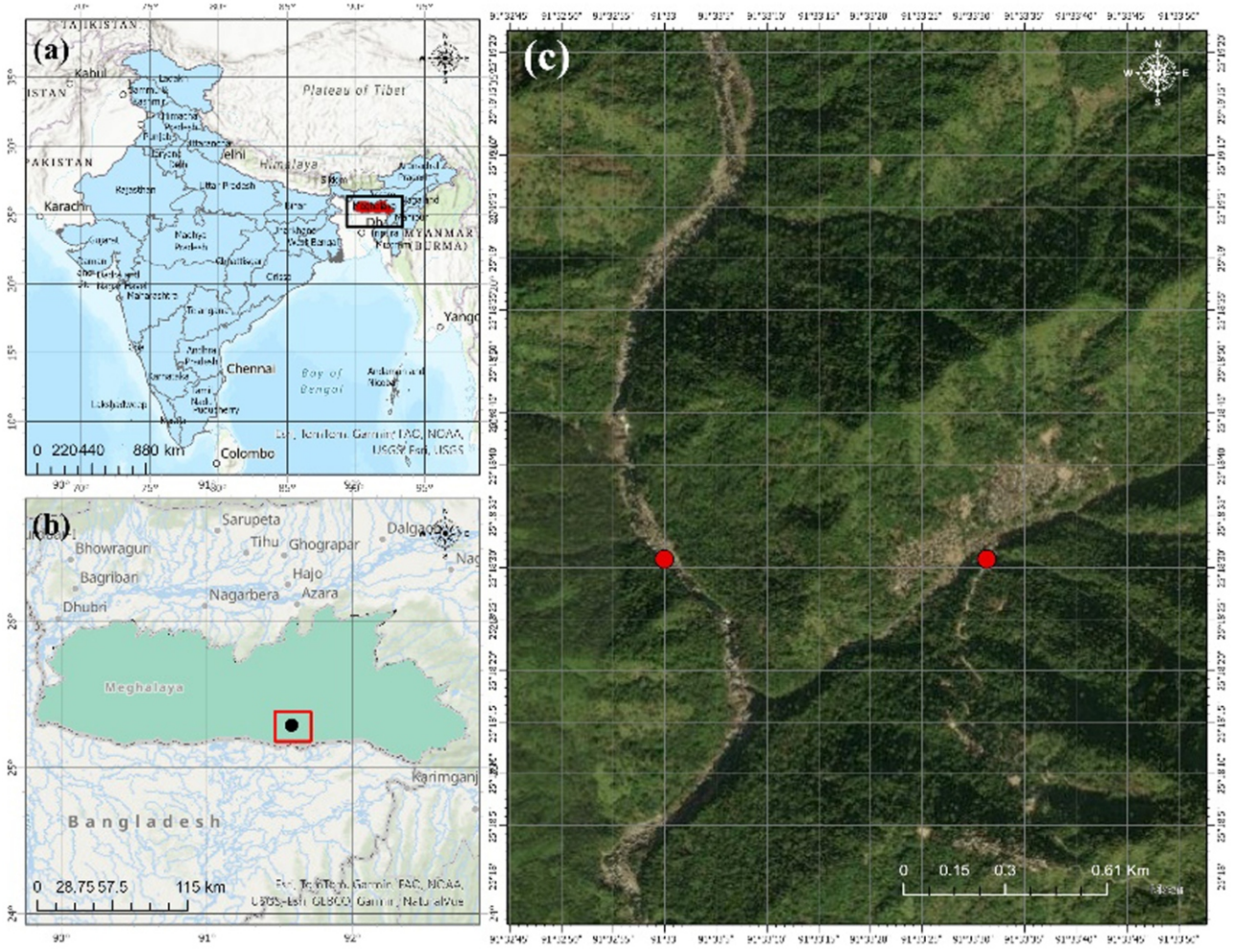


Fig. 1: Map of sampling sites Domsopian (25°18'30"N, 92°33'0"E) and Kenmynsaw Syntein (25°18'30"N, 91°33'30"E) of *Neolissochilus hexagonolepis* and *N. hexastichus* in river Umngi of Meghalaya.

value was obtained as slope of the log regression equation:

$$\ln (Y) = \ln(a) + b[\ln(X)] \quad (2)$$

where, Y is a vector comprising measurements of a morphometric trait of forty individuals; X is a vector of the corresponding SL values. a and b are the intercept and slope of the regression line. The slope (b) of log regression line (Eq. 2) of each trait was compared via t-test with a slope of 1 to determine if traits differed from isometric growth. t-statistics were obtained from the equation furnished below at 38 degrees of freedom (N-2) and $\alpha = 0.05$:

$$t = (b-1)/\text{standard error (S.E) of slope} \quad (3)$$

HCPC analysis of normalized morphometric traits: The dataset comprising 19 normalized morphometric traits of 40 mahseer individuals were analysed by the Hierarchical Clustering on Principal Components (HCPC) procedure. The R packages FactoMineR (Le *et al.,* 2008) and *factoextra* were used to perform PCA and obtain the results. PCA solutions of first four principal components were utilised for cluster analysis by HCPC function of FactoMineR. Ward's criterion was used to agglomerate clusters. The appropriate number of clusters were validated by implementing 50 iterations of K-means clustering algorithm. The number of clusters obtained by HCPC result was also separately validated using the *NbClust* package of R, while the completeness and homogeneity of cluster was verified by Silhouette Width.

Molecular phylogenetic analysis: For molecular phylogenetic analysis, total genomic DNA was extracted from pectoral fin clips of 40 *Neolissochilus* individuals using DNeasy (Qiagen) DNA extraction kit. PCR amplification of a 655 bp segment of the 5' end of the mitochondrial cytochrome c oxidase subunit I (COI) gene was conducted using the standard protocol with the primers FishF1 and FishR1 as described by Ward *et al*. (2005). The amplicons were column-purified, and Sanger sequencing was

performed by an external service provider.

Chromatograms of initial sequences were scrutinized for quality calls of nucleotide bases using Seqtrace (Stucky, 2012); low-quality bases were edited and primer sequences were trimmed. Finished sequences were checked for correct open reading frame using ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/) and submitted to GenBank (https://www.ncbi.nlm.nih.gov/nucleotide/) as partial coding sequences (cds). The GenBank accession numbers and isolate codes of these coding sequences are mentioned in Fig. 3.

Taxon sampling was done by retrieving from GenBank the COI sequences of 40 *Neolissochilus* isolates of the present study and seven submitted by previous research groups; three of these were of *N. hexastichus,* three of *N. hexagonolepis* and one of *Puntius* (*Naziritor*) chelynoides, which was selected as the outgroup, based on its genetic proximity to mahseers of genus *Neolissochilus* (Shahi *et al.,* 2023). These 47 COI sequences were aligned according to the Clustal W algorithm using MEGA 11 (Tamura *et al.,* 2021). There were 705 nucleotide positions or sites in the multiple sequence alignment (MSA) that was used in the Nexus format for molecular phylogenetic analysis. Isolate codes and accession numbers obtained from GenBank were used for describing the results of phylogenetic and cluster analysis. Bayesian inference (BI) of phylogeny as implemented in MrBayes 3.2 (Ronquist *et al.,* 2012b) was used to build the most probable phylogenetic tree based on the General Time Reversal (GTR) nucleotide substitution model with gamma distributed variation across sites and a proportion of invariant sites.

Four independent Bayesian analyses were performed simultaneously. In each analysis, four Markov Chain Montecarlo (MCMC) chains ran for 650000 generations with default priors. The trees were sampled every 5000 generations and the first 25 % of trees were discarded as 'burn-in'. Standard deviation of split frequencies < 0.01 and potential scale reduction factor equal to 1.0 were considered as convergence of runs. The post-burn-in trees were used to construct a 50% majority-rule consensus tree and to calculate Bayesian posterior probability (BPP) of nodes. A barcode gap analysis and species delimitation procedure were performed by ABGD web-interface (Puillandre *et al.,* 2012). Kimura two parameter (K2P) model of nucleotide substitution (Kimura, 1980) was used, in the ABGD procedure, for estimating of sequence divergence.

Results and Discussion

Measurements of standard length (SL), which was used as an indicator of body size of mahseers in the present study, ranged from 51.14 mm to 144.17 mm with a mean of 73.01 mm. Table 1 lists the regression results of log-log transformed data, the t-test comparing b (the slope of regression equation 2) versus a slope of 1. The residual standard error (R.S.E) values of regression test were low, whereas the r^2 values were high (Table 1). These values validated a perfect fit of the regression

Table 1: Important measures of data normalization, allometry and clustering

Trait (response variable)	Regression Analysis (goodness of fit test)		t-Test for allometry evaluation			η^2
	R.S.E	r^2	B	t-statistic	p-value	
Head Length	0.035	0.978	1.012	0.506	0.308	0.53
Body Depth	0.049	0.961	1.054	1.586	0.061	0.13
Eye Diameter	0.069	0.804	0.617	-7.855	<0.001	ns
Snout Length	0.097	0.891	1.193	2.852	0.003	0.66
Post Orbital Length	0.048	0.946	0.885	-3.375	<0.001	ns
Inter Orbital Length	0.127	0.682	0.804	-2.196	0.017	0.75
Upper Jaw Length	0.092	0.859	0.977	-0.356	0.362	ns
Lower Jaw Length	0.138	0.732	1.009	0.094	0.463	0.41
Anal Length	0.017	0.995	1.037	2.985	0.002	ns
Head Depth	0.053	0.946	0.958	-1.138	0.131	ns
Dorsal Fin Length	0.053	0.931	0.829	-0.465	0.322	0.24
Dorsal Fin Base	0.065	0.925	0.982	-0.395	0.348	ns
Pectoral Fin Length	0.048	0.959	1.004	0.126	0.451	0.29
Pectoral Fin Base	0.141	0.734	1.011	0.121	0.452	0.29
Anal Fin Length	0.078	0.897	0.962	-0.725	0.236	0.13
Anal Fin Base	0.074	0.919	1.084	1.618	0.057	ns
Ventral Fin Length	0.045	0.962	0.967	-1.052	0.149	0.35
Ventral Fin Base	0.104	0.879	1.213	2.918	0.003	ns
Caudal Fin Length	0.071	0.891	0.875	-2.516	0.008	0.5

Standard Length (SL) was the predictor variable for regression analysis; R.S.E.: Residual Standard Error; r^2 = Coefficient of Determination.b: slope of regression lines (equation 2). Traits with p-values <0.05 (in bold) are allometric. η^2 values: significant effects of traits on clustering

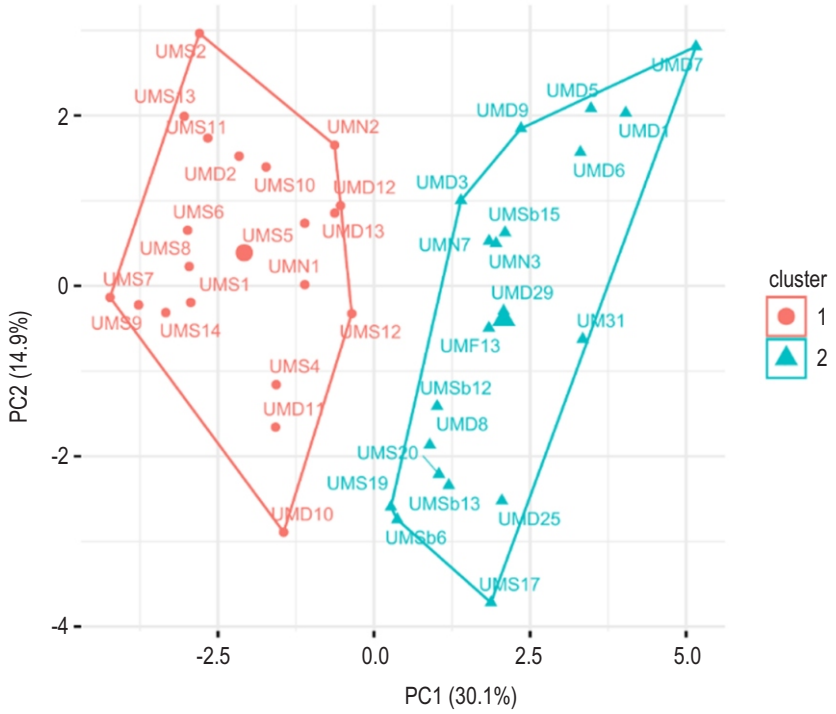


Fig. 2: PCA-HCPC plot depicting clustering of 40 *Neolissochilus* samples based on multidimensional variances in morphometric traits. Variances retained by the first two principal components, PC1 and PC2, are mentioned.

model, emphasizing the confidence in the slope b as the allometric exponent in the normalization equation (Eq. 1), and also as a test variable in the t-test (Eq. 3) for allometry evaluation.

The results of t-test statistics and associated p-values (Table 1), revealed that the seven traits, eye diameter (ED), post orbital length (POL), inter orbital length (IOL) and caudal fin length (CFL) were negatively allometric, while snout length (SNL), anal length (AL) and ventral fin base (VFB) were positively allometric and the remaining 12 traits were isometric. Allometry describes the relationship between size and morphometric traits (Goodwin and Arbour, 2023). This phenomenon is governed by ontogenetic and evolutionary factors. Recent reports on mahseers (Kechu *et al.,* 2024; Rana and Nautiyal, 2023) focussed on the allometric effects of size on body weight. Sarmah *et al.* (2024) reported allometric relationships of morphometric traits of mahseers of Jia Bhoroli river, Assam, which is line with the findings present work. However, in-depth studies on influences of habitat, food and feeding habits on allometric growth of mahseers are required in the future. The information generated from such research will be useful in resolving taxonomic ambiguities and in designing effective conservation and breeding programmes.

The normalization procedure adopted in this study was originally developed by Thorpe (1975); it was found by Reist (1985) to be superior to several other such techniques. Lleonart *et al.* (2000) presented Thorpe's normalization function in a computationally simple form and enunciated the incorporation of specific shape factor in the function. Application of Thorpe's normalization function to our morphometric data transformed these into values that mahseer individuals of 73.01 mm standard length would have. Normalization of morphometric data was deemed essential prior to multivariate analysis by Lleonart *et al.* (2000) and Thorpe (1976) so that variations due to body size difference are removed. A common practice for removing body size effects in fishes is the use of ratios. Measured traits are divided by standard length and expressed as a percentage. However, ratios cannot remove the effect of size from allometric characters (Lleonart *et al.,* 2000) and hence, may not be effective in analysis of multivariate morphometric traits.

The variance retained by the first four principal components (PC_s) of PCA solution were utilized for downstream processing by HCPC protocol following the suggestion of Peres-Nato *et al.* (2005) that PCA solutions of axes that have eigenvalues greater than 1 should be utilized. The cumulative variance retained by these four axes were 62.18%, while the eigenvalues and variance percentage retained on them were PC_1 (5.72& 30.1%), PC_2 (2.83&14.9%), PC_3 (1.9&10.25%), PC_4 (1.32&6.9%), respectively. Eigenvalues on PC5 and higher dimensions were ≤1. The Ward method (Ward,1963) of hierarchical clustering agglomerated the multidimensional variance. The K-means partitioning algorithm performed iteration

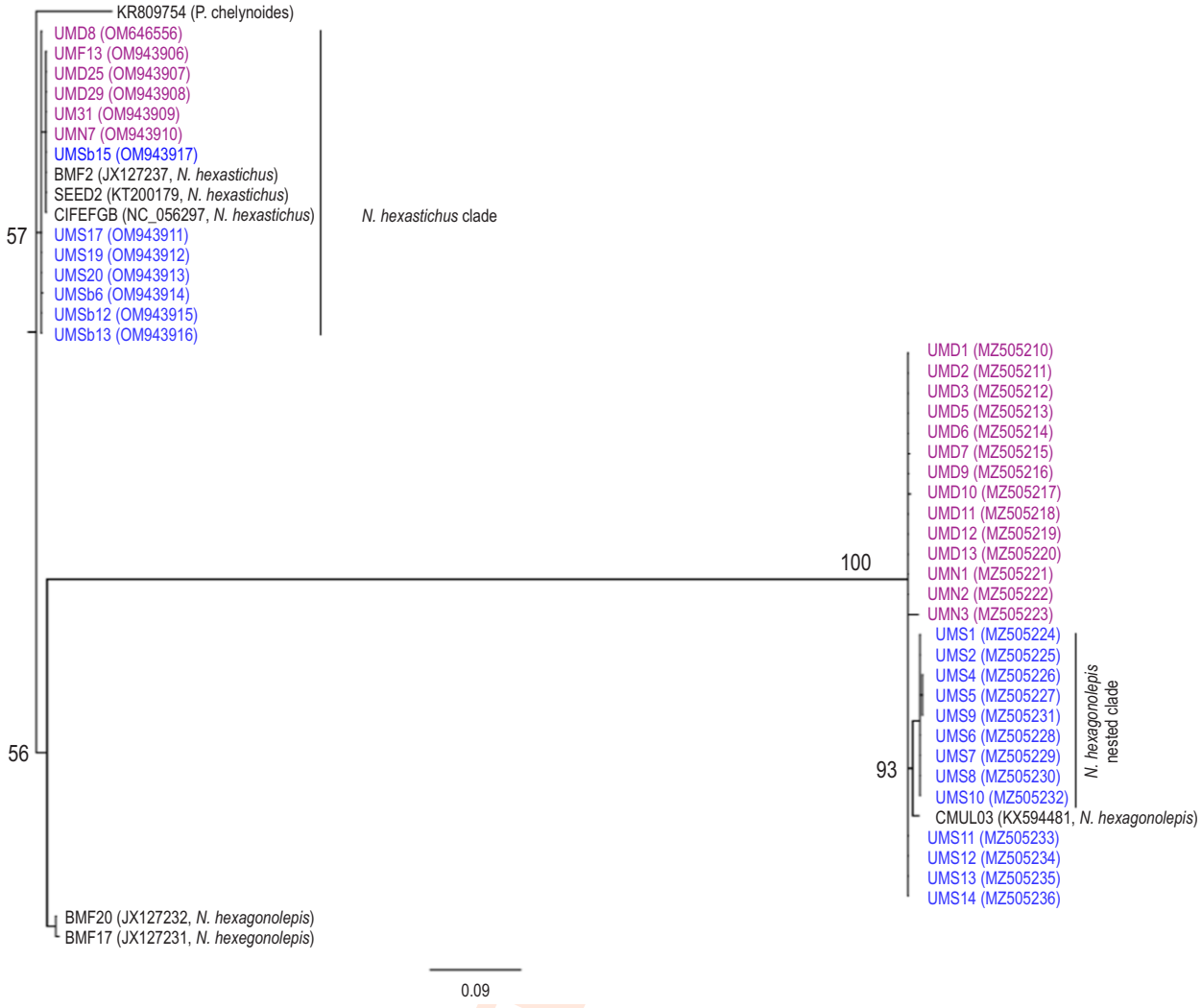


Fig. 3: Bayesian Inference phylogeny of 47 mahseer individuals (vide Section 2.3) based on CO I sequence divergence. Isolate codes in blue followed by GenBank ids in parenthesis represent *Neolissochilus* spp. sampled from Domsopian site and those in magenta represent *Neolissochilus* spp. sampled from Kenmynsaw Syntein site. *N. hexastichus, N. hexagonolepis* and *P. cheloynoides* (outgroup) reported by previous workers are in black. Numbers on nodes are Bayesian Posterior Proability (BPP) support values in percentages.

of the hierarchical clustering solution to confirm or suggest a better clustering solution.

HCPC categorized 40 mahseer individuals into two clusters (Fig. 2). The NbClust function validated the appropriate number of clusters as two according to the majority rule. Each cluster were assigned twenty individuals as per HCPC protocol. No outliers were detected. Eleven traits, HL, SnL, LJL,IOL, BD, PFB, PFL, DFL, VFL, CFL and AFL had significant effects in separating individuals into two clusters, as revealed by η^2 scores (Table 1). Of these SnL, LJL, HL, BD were among the diagnostic features in the morpho-taxonomic descriptions of mahseers by Lalramliana *et al.* (2019), Laskar *et al.* (2013) and Talwar and Jhingran (1991). Furthermore, this aspect of our finding is in line with the observation of Rainboth (1985) who proposed that the head and body shape are useful indicators of species differences. Thus, the η^2 statistic of HCPC results revealed an important information in present dataset relevant to species delimitation. The silhouette widths of Cluster 1 individuals ranged from 0.42 to 0.116; the median value was 0.312. The silhouette widths of Cluster 2 individuals ranged from 0.39 to 0.118; the median value was 0.261. These values reveal intra-cluster variations in morphometric traits. Analysis of divergence of mitochondrial COI sequences by the Bayesian Inference (BI) approach recovered *Neolissochilus* congeners in two clades and the outgroup *Puntius (Naziritor) chelynoides,* the dark mahseer, in a separate branch from the basal node (Fig. 3). The *N. hexastichus* congeners BMF2 (Laskar *et al.,* 2013), SEED2 and CIFEFGB-SERB-NH-01 (Singh *et al.,* 2019) were resolved in one clade having 57% Bayesian Posterior Probability (BPP) support, with the isolates UMD8,

Table 2: Mean (± SD) values of normalised morphometric traits

Trait	*N. hexatichus* n = 13	*N. hexagonolepis* n = 9	*Intermediates* n = 18
Head Length	18.113±0.44	19.109±0.31	18.949 ±0.59
Body Depth	18.564±0.66	19.198±0.95	19.477±0.91
Eye Diameter	5.149±0.24	5.348±0.55	5.198±0.32
Snout Length	5.269±0.21	6.48846728	5.904±0.52
Post Orbital Length	10.647±0.41	11.093±0.43	11.102±0.57
Inter Orbital Length	8.442±0.51	6.74±0.53	7.885±0.97
Upper Jaw Length	5.574±0.24	5.575±0.33	5.797±0.86
Lower Jaw Length	5.026±0.26	4.23±0.7	4.816±0.58
Anal Length	52.127±0.62	51.729±0.74	51.924±1.12
Head Depth	13.385±0.45	13.555±0.83	13.809±0.74
Dorsal Fin Length	18.785±0.57	17.947±0.57	18.967±1.2
Dorsal Fin Base	10.588±0.6	10.281±0.56	10.929±0.73
Pectoral Fin Length	15.233±0.46	14.480±0.72	15.226±0.72
Pectoral Fin Base	2.925±0.35	3.595±0.38	3.223±0.41
Anal Fin Length	13.09±1.18	12.714±0.43	13.138±0.87
Anal Fin Base	5.273±0.42	5.673±0.36	5.542±0.35
Ventral Fin Length	13.558±0.23	12.992±0.49	13.462±0.72
Ventral Fin Base	3.396±0.37	3.637±0.36	3.62±0.35
Caudal Fin Length	23.985±1.07	21.552±0.77	23.397±1.78

Mean SL was 73.01 mm (vide sections 2.1 and 3.1)

UMF13, UMD25, UMD29, UM31, UMN7, UMSb15, UMS17, UMS19, UMS20, UMSb6, UMSb12, and UMSb13 of the present study. The congeners *N. hexagonolepis* BMF20 and BMF17 from Laskar *et al.* (2013), as well as CMUI.03 from Sharma *et al*. (2019), were grouped in a clade with twenty-seven isolates from the present study, which was supported by a 56% Bayesian posterior probability (BPP) (Fig. 3). The topology of this *N. hexagonolepis* clade suggested that BMF20 and BMF17 were genetically distant from the rest of the members, which were grouped as a subclade with 100% BPP support. CMUI.03 formed a nested clade, having 93% BPP support, with the isolates UMS1, UMS2, UMS4, UMS5, UMS9, UMS6, UMS7, UMS8, UMS10 in the present study (Fig. 3). A comparative analysis of the morphological features of *Neolissochilus* congeners within the nested clade of the present study, as well as those in the *N. hexastichus* clade, is provided below. The mouth was truncated in nested clade vs smoothly rounded in *N. hexastichus* clade. Gill raker counts on the lower arm of the gill arch was eight in nested clade vs nine in *N. hexastichus* clade. These descriptions are in line with those of *N. hexagonolepis* vs *N. hexastichus* (Laskar *et al*., 2013). In contrast, all 40 mahseer specimens examined exhibited a green coloration on the back, with reddish fins that faded toward the tips, and a coppery tinge on the scales above the lateral line. These observations diverge from the descriptions provided by Laskar et al. (2013), who reported that *N. hexagonolepis* displayed a green back, while *N. hexastichus* had a brownish back. According to their descriptions, the scales above the lateral line of *N. hexagonolepis* had a coppery tinge at the ends, whereas those of *N. hexastichus* were reddish-brown. Additionally, *N. hexagonolepis* was described as having deep slate-colored fins that pale toward the margins, while *N. hexastichus* fins were characterized by red tips. Furthermore, Talwar and Jhingran (1991) described interrupted labial groove in *N. hexagonolepis* but continuous in *N. hexastichus.* In contrast, all forty *Neolissochilus* congeners in the present study had interrupted labial groove. The lateral line scales ranged from 27 – 30, while the dorsal (D), ventral (V), pectoral (P) and anal (A) fin ray (hard + soft) counts were D iv+9, V i+ 7-8, P i+15-16, A iii+5. These observations were irrespective of cluster or clade memberships of the samples.

The congeners of the present study in nested clade were placed in Cluster 1 of the HCPC result (Fig. 2), while those of the *N. hexastichus* clade were placed in cluster 2 of the HCPC result. It is pertinent to reiterate here that Clusters were based on morphometric similarities and clades were based on genetic affinities. Thus, going by the concordance of the morphometric groupings with genetic groupings as well as by the gill rakers counts, we report the isolates UMD8, UMF13, UMD25, UMD29, UM31, UMN7, UMSb15, UMS17, UMS19, UMS20, UMSb6, UMSb12, and UMSb13 as *N. hexastichus* individuals and the isolates UMS1, UMS2, UMS4, UMS5, UMS9, UMS6, UMS7, UMS8, UMS10 as *N. hexagonolepis* individuals. Normalized morphometric traits of these individuals are presented in Table 2. Eighteen isolates of the present study, UMD1, UMD2, UMD3, UMD5, UMD6, UMD7, UMD9, UMD10, UMD11, UMD12, UMD13, UMN1, UMN2, UMN3, UMS11, UMS12, UMS13 and UMS14 were on multifurcating branches within the *N. hexagonolepis* sub-clade. Their Gill raker count on the lower arm of the gill arch ranged from eight to eleven. They could not be distinguished based on their colour in life. Seven of these isolates (UMD9, UMD5, UMD7, UMD3, UMD6, UMD1, and UMN3) were

Online Copy

placed in Cluster 2 of HCPC result, and the remaining eleven were placed in Cluster 1 of HCPC result (Fig. 2). These individuals are probably the natural hybrids of *N. hexastichus* and *N. hexagonolepis.* We report them as intermediates (Table 2). Genetic proximity among sympatric populations is conducive to natural hybridization in freshwater fishes (Pinheiro *et al.,* 2019). The topology of BI tree (Fig. 3) and ABGD results reveal genetic proximity among the *N. hexastichus* and *N. hexagonolepis* congeners of present study.

ABGD computed the barcode gap distance as 0.667 and returned 3 partitions or groups at maximum intra-species COI sequence divergence of 0.059. One partition comprised *P. chelynoides,* while another comprised the taxa of *N. hexagoloepis* sub-clade in the BI tree (Fig. 3), and the third partition had the taxa of *N. hexastichus* clade of the BI tree. *N. hexagoloepis* congeners, BMF17 and BMF20, Laskar *et al.* (2013) were also in third partition indicating their genetic affinities to congeners of *N. hexastichus* clade as well as to those of *N. hexagonolepis* sub-clade of BI tree (Fig. 3). The low BPP values of the node supporting *N. hexagonolepis* and *N. hexastichus* clades, 56% and 57% respectively, also indicate close genetic kinship. Singh *et al.* (2019) and Laskar *et al.* (2013) have reported genetic proximity of *N. hexastichus* and *N. hexagonolepis.* Thus, the integrative approach, of this study, apart from reiterating genetic nearness between the two species, also identified the existence of individuals having intermediate morphological characteristics, probably due to natural hybridizations between them.

Ogale and Kulkarni (1987) reported the hybridization between *Tor khudree* and *T. tor* in ponds without the use of pituitary hormone or any other inducers. Hence, it is likely that hybridization also occurs among the sampled populations of *Neolissochilus* in river Umngi of Meghalaya. In the light of our findings, we recommend using combinations of nuclear and organelle DNA markers for confirming the existence of natural hybridization. Exact identification of true-breeding natural populations of *N. hexagonolepis* and *N. hexastichus* is required for successful aquaculture and captive breeding programmes.

Acknowledgments

The authors are grateful to Dr. P.K. Pandey, Director, ICAR-DCFR, Bhimtal for funding and continued support and Dr. (Br) A.L. Dkhar, Principal, St. Anthony's College, Shillong for providing infrastructure for the present study

Authors' contribution: M.A. Laskar, D. Sarma and R. Nath: Conceptualized and executed the research project and prepared the manuscript; **R. Kharbani:** Assisted in sample collection, morphometric and meristic data collection; **N. Sahi and R.S. Haldar:** Provided technical advice and managed fund availability; **P.A. Ganie:** Prepared the GIS map of sample collection sites.

Funding: Funded by ICAR-Directorate of Cold Water Fisheries Research, Bhimtal, Uttarakhand. Grant Reference No.: F.No. 3-12(1)/2017-Estt./Vol.I Sanction No. 140/20-21 Date:29.07.2020

Research content: The manuscript contains original research content that has not been published elsewhere.

Ethical approval: Guidelines of the National Biodiversity Board were followed during fish sampling.

Conflict of interest: The authors declare that they have no conflict of interest.

Data availability: Data of this study are available on request from the corresponding author.

Consent to publish: All authors agree to publish the paper in *Journal of Environmental Biology.*

References

Chan, K.O. and L.L. Grismer: A standardized and statistically defensible framework for quantitative morphological analyses in taxonomic studies. *Zootaxa,* **5023**, 293 – 300 (2021).

Dash, P., R.S Tandel and D. Sarma: Breeding and conservation of chocolate mahseer in captivity. In: Aquaculture and Conservation of Inland Coldwater Fishes (Eds.: D. Sarma, S. Chandra, S.K. Malik). Springer Nature Singapore Pte Ltd, Singapore, pp. 69 – 81 (2024).

Goodwin, M.R. and J.H. Arbour: Darter fishes exhibit variable intraspecific head shape allometry and modularity. *Anat. Rec.,* **306**, 446 – 456 (2023).

Granato, D., J.S. Santos, G.B. Escher, B.L. Ferreira and R.M. Maggio: Use of principal components analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: a critical perspective. *Trends Food Sci.,* **72**, 83-90 (2018).

Johnson, J.A., P. Nanda, B. Pant, S.S. Mane and V. Kolipakam: Redescription of an endemic mahseer, *Tor mahanadicus* (David, 1953) from Mahanadi river basin, India based on integrated morphological and molecular techniques. *Plos ONE,* **18**, e0291436 (2023).

Kechu, M., S. Ezung and P.P. Pankaj: Bionomics, length-weight relationship and condition factor of Chocolate mahseer, *Neolissohilus hexagonolepis* from Dikhu river, Nagaland. *J. Environ. Biol.,* **44**, 706 – 712 (2023).

Kimura, M.: A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.,* **16**, 111-120 (1980).

Lalramliana, S. Lalronunga, S. Kumar and M. Singh: DNA barcoding revealed a new species of *Neolissochilu*s Rainboth, 1985 from the Kaladan river of Mizoram, North East India. Mitochondrial DNA PartA, DNA Mapp. *Seq. Anal.,* **30**, 52 – 59 (2019).

Laskar, B. A., D. N. Das and B. C. Tyagi: Growth performance of the *Neolissochilus hexagonolepis* (McClld) in pond system in Arunachal Pradesh. *Ind. J. Fish.*, **56**, 55-59 (2009).

Le, S., J. Josse and F. Husson: FactoMiner: An R package for multivariate analysis. *J. Stat. Softw.,* **25**,1-8 (2008).

Lleonart, J., J. Salat and G.J. Torres: Removing allometric effects of body size in morphological analysis. *J. Theor. Biol.,* **205**, 85 – 93 (2000).

Ogale, S.N. and C.V Kulkarni: Breeding of pond raised hybrids of mahseer fish *Tor khudree,* Sykes and Tor tor, Ham. *J. Bombay Nat. Hist. Soc.,* **84**, 332 – 335 (1987).

Peres-Neto, P.R., D.A. Jackson and K.M Somers: How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Comput. Stat. Data Anal.,* **49**, 974-997 (2005).

Pinheiro, A.P.B, R.M.C.M Melo, D.F. Teixeira, L.J.O. Birindelli, D.C. Carvalho and E. Rizzo: Integrative approach detects natural hybridization of sympatric lambaris species and emergence of infertile hybrids. *Scientific Rep.,* **9**, 4333 (2019).

Puillandre, N., A. Lambert, S. Brouillet and G. Achaz: ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.,* **21**, 1864-1877 (2012).

R Core Team:_R:A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2023). <https://www.R-Project.org/>.

Rainboth, W.J.: *Neolissochilus*, A new genus of South Asian cyprinid fishes. *Beufortia,* **35**, 25 – 35 (1985).

Rana, P. and P. Nautiyal: Body growth and condition of endangered *Tor putitora* (Hamilton, 1822) (Actinopterygii: Cypriniformes: Cyprinidae) in the crucially important breeding grounds of the Ganga stock. *J. Threat. Taxa,* **15**, 24255 – 24260 (2023).

Reist, J.D.: An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. *Can. J. Zool.,* **63**, 1429 – 1439 (1985).

Ronquist, F., M. Teslenko, P.V.D Mark, D. Ayres and A. Darling: MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.,* **61**, 539 – 542 (2012b).

Sarma, D., M.S. Akhtar, P. Sharma and A.K. Singh: Resources, breeding, eco-tourism, conservation, policies and issues of Indian mahseer: A review. *J. Coldwater Fisheries,* **1**, 4-12 (2018).

Sarmah, R., U.K. Das, D. Sarma, N. Sahi, P. Ahmed, P.K. Pandey and R. Dutta: Morphometric and molecular assessment of *Tor putitora* and *Neolissochilus hexagonolepis* population in Jia Bhoroli river ecotone zone, Assam. *J. Animal Res.,* doi: 10.18805/IJAR.B-5156 (2024).

Shahi, N., D. Sarma, B. Singh, S.K. Mallik, D. Baruah, R. Posti, R. Haldar and I. Linthoingambi: Characterizing the dark mahseer, *Naziritor chelynoides* (McClelland, 1839): A morphological, osteological, and molecular approach. *Genet. Aquat. Org.,* **7**, Ga601 (2023).

Sharma, L., S. Ali, A. Barat, R. Kumar, V. Pande, M.A. Laskar, P.K. Sahoo and S. Sumer: Molecular Identification and genetic diversity analysis of Chocolate masheer (*Neolissochilus hexagonolepis*) populations of North East India, using mitochondrial DNA markers. Mitochondrial DNA PartA DNA Mapp. *Seq. Anal.,* **30**, 397 – 406 (2019).

Schar, S.: allomr: Removing body size in morphological analysis. R package version 0.3.0, https://CRAN.R-project.org/package=allomr (2023). DOI: 10.32614/CRAN.package.allomr

Singh, S., A. Mishra, A. Pavan-Kumar, S.B. Singh, K.V. Radhakrishnan, N.S. Nagpure and A. Chaudhuri: Characterization of the complete mitochondrial genome of *Neolissochilus hexasticus* (McClleland, 1839). *Mitochondrial DNA BResour.,* **9**, 2461-2463 (2021).

Stucky, B.J.: SeqTrace: A graphical tool for rapidly processing DNA sequencing chromatograms. *J. Biomol. Tech.,* **23**, 90-93 (2012).

Talwar, P.K. and J.G. Jhingran: Inland fishes of India and adjacent countries. **Vol. 1,** Oxford IBH Publication, N. Delhi, pp. 229 – 230 (1991).

Tamura, K., G. Stecher and S. Kumar: MEGA: Molecular Evolutionary Genetic Analysis version 11. *Mol. Biol. Evol.,* **38**, 3022-3027 (2021).

Thorpe, R.S.: Quantitative handling of characters useful in snake systematics with particular reference to intraspecific variation in the Ringed snake *Natrix natrix* (L.). *Biol. J. Linn. Soc. Lond.,* **7**, 27-43 (1975).

Thorpe, R.S.: Biometric analysis of geographic variation and racial affinities. *Biol. Rev. Camb. Phil. Soc.,* **51**, 407 – 452 (1976).

Ward, J.H.: Hierarchical grouping to optimize and objective function. *J. Am. Stat. Assoc.,* **58**, 236 – 244 (1963).

Ward, R.D., T.S. Zemlak, B.H. Innes, P. RLast and P.D.N Hebert: DNA barcoding of Australia's fish species. *Philos. Trans. R. Soc. Lond. B Biol. Sci.,* **360**, 1847 – 1857 (2005).

On Line Copy