Borneo International Journal of Biotechnology
(BIJB) Vol. 5 (December 2025), 12-31.

e-ISSN 2716-697X

Geographic Isolation of Riparian Ecosystems Determine Population Diversity of *Tor* sp in Sabah, Malaysia.

Helena Biun[1], Ahemad Sade[2], Kenneth Francis Rodrigues[1]

[1]Biotechnology Research Institute, Universiti Malaysia Sabah, 88400, Kota Kinabalu, Sabah, Malaysia
[2]Jabatan Perikanan Sabah, 4th Floor, Wisma Pertanian Sabah, 88624 Kota Kinabalu, Sabah, Malaysia
*Corresponding author's email: kennethr@ums.edu.my
Received: 23 September 2025 | Accepted: 20 December 2025 | Published: 31 December 2025

ABSTRACT

Tor sp is a cyprinid riverine fish known as Pelian in the local dialect in Sabah, that is currently facing challenges due to climate change and habitat degradation. Applying conservation strategies and a sustainable management program for this species requires a primary assessment of the genetic diversity and overall structure of the geographically isolated populations. Eighteen wild Pelian populations from across ten river systems were analyzed based on eleven microsatellite markers. The gene diversity and allelic richness based on microsatellite loci ranged from 0.40-0.68 and 2.63-6.10, respectively. Pairwise FST values for microsatellites were significant ($P<0.05$) between the majority of populations. Microsatellites analyses of molecular variance, AMOVA analyses detected variation within populations that ranged from 60.85% to 50.74% within the entire watershed, with significantly high FST values. Mantel tests supported weak patterns of differentiation based on isolation by distance. The overall population comprised two distinct clusters that exhibited further sub structuring based on the watershed. The topographical features of the landscape in Borneo comprise a combination of isolation by distance, river fragmentation and historical isolation by the Crocker Trusmadi mountain range all support the hypothesis that population substructure was driven by isolation in distinct ecological niches. These results will be used in the application of conservation strategies and management program for Pelian in Sabah.

Keywords: Borneo, microsatellites, Pelian, population genetics, *Tor* sp

e-ISSN 2716-697X

INTRODUCTION

Tor tambra is a species of ray-finned fish of the family Cyprinidae in the genus *Tor*. In Sabah, East Malaysia, all *Tor* species are known by the ethnic group, the Dusuns, as 'Pelian'. *Tor* is referred in the local dialects as 'Empurau' and 'Semah' in Sarawak, and 'Kelah' in Peninsular Malaysia. The geographic range extends across Cambodia, Yunnan, Jawa, Kalimantan, Sumatera, Laos, Malaysia, Thailand, and Vietnam (Pinder et al., 2019). Pelian is endemic to Malaysian Borneo river basins and is presumably the most widespread Tor species recorded in Malaysian Borneo (Esa et al., 2008). Analysis of the mitochondrial DNA (mtDNA) *coxI* gene imply that specimens collected from Malaysian Borneo may represent a new *Tor* species (Walton et al., 2017). Pelian is an important fisheries resource to the local communities associated with riparian settlement that line the Kinabatangan, Kanarom, Labuk Sugut along the East coast of Sabah, and Pagalan Padas, Wariu Kadamaian, Tuaran, Moyogm Kimanis and Bongawan that all flow into the South China sea (Ahmad et al., 2006). Pelian inhabits the shallow, rocky areas of the river with strong currents. Their preferred diet comprises plant matter, insects, flowers and fruit. The breeding behaviour involves multiple spawners, with spawning occurring in streams up to 500 meters above sea level. Upstream migration from lower reaches of rivers by schools of Pelian starts when the breeding season commences, which coincides with the monsoon rains. Sub-adult Pelian can be found in the upper reaches of the water column at both higher and lower reaches of rivers, whereas larger adult Pelian are mostly found at the lower reaches of rivers, occupying the deepest portions of pools. Sarawak has reported the successful breeding of both *T. tambroides* and *T. douronensis* from brood stock reared in captive culture (Ingram et al., 2007). The riparian communities in Sabah are actively engaged in a community-based conservation scheme termed as "Tagal", that consists of villagers who manage and protect Pelian along selected stretches of the rivers. Fishing is prohibited in these areas and harvesting is only permitted at specific days in a year in order to promote sustainable fisheries. Within the island of Borneo, Pelian is genetically distinct as compared to the *Tor sp* endemic to the rivers of Sarawak (Esa et al., 2008). Microsatellite markers have been popular for population genetics studies of marine and freshwater fish species due to their codominance, polymorphism, ability to amplify specific loci, hyper variability, reproducibility across varieties, ease of application in laboratory setting and amenability to statistical analysis and interpretation using generally accepted statistical methods (Abdul-Muneer, 2014). Population characteristics of *T. tambra* such as fragmentation, isolation and unbalanced sex ratios observed in the wild, make the populations susceptible to the phenomenon of genetic drift, bottleneck and inbreeding processes. This study was conducted with the aims of demonstrating the utility of population genetic studies using microsatellite markers in understanding population diversity and structures of Pelian for appropriate conservation and management strategies in Sabah.

e-ISSN 2716-697X

MATERIALS AND METHODS

Sample Collection

The samples of *Tor tambra* were obtained from 18 village communities who were involved in managing Tagal systems in the respective villages as listed in **Table 1**. The location of the sampling sites is indicated in **Figure 1**. The identity of the species was established based on morphological features consistent with *Tor* sp (e.g., median lobe length, coloration). While recent mtDNA studies suggest potential cryptic diversity in Borneo, our initial screening and morphological validation treat the sampled populations as a single coherent genetic unit ('Pelian') for the purpose of population structure analysis. A single caudal fin clip was excised from each of the specimens (Figure 2), stored in a 50 ml tube containing 95% Ethanol and shipped to the laboratory at the Department of Fisheries, Likas, Sabah. Total DNA was extracted as per the manufacturers protocol using a Wizard Genomic DNA Extraction Kit (Promega, Madison WI, USA). The DNA was eluted in 50 microliters of elution buffer and the final concentration was adjusted to 50 nanograms per microliter using elution buffer. DNA was quantified using a Spectrophotometer and stored at -20^{o} C prior to PCR amplification.

Table 1: The river systems and abbreviations for the localities at which the samples were collected.

River system	Locality	Abbreviation
Kanarom	Sorinsim	KSO
	Tangkol	KTA
	Sunsui	KSU
	Marak Parak	KMP
Sugut	Luanti	SLU
	Poring	SPO
Labuk	Kinarasan	SKI
	Paus	SPA
Kinabatangan	Pinipi	SPI
	Kironggu	SKIR
Pagalan	Menserulong	IME
	Rugading	IRU
Kadamaian	Tokulung	WCTO
	Lingkubang	WCLI
Tuaran	Gontung	WCGO
Moyog	Babagon	WCBA
Kimanis	Doingin	WCDO
Bongawan	Telantang	WCTE

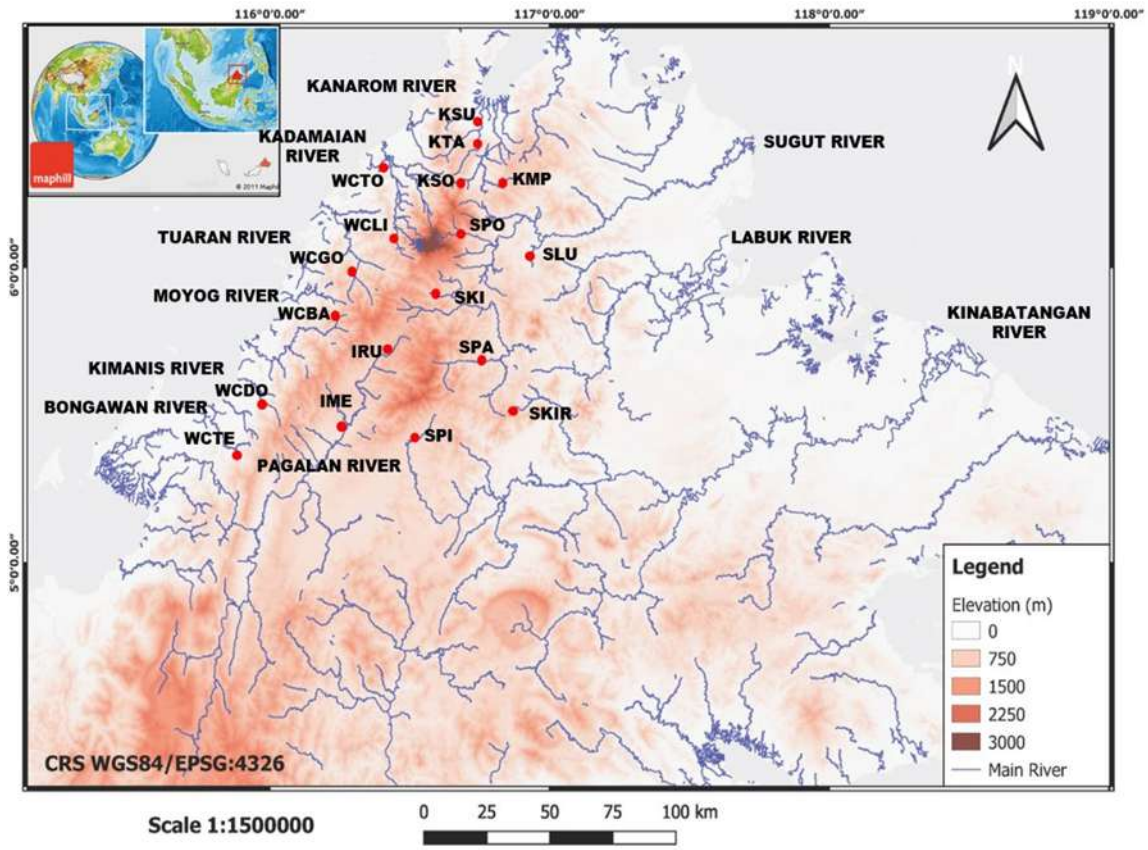

Figure 1: The locations at which sample were collected. The abbreviations are listed in Table 1.

Microsatellite analysis

For microsatellite analysis, polymerase chain reaction amplification was set up in a 50 µL reaction mixture containing 1× TopTaq Master Mix buffer (Qiagen, Germany), 0.2 µM of each forward and reverse primers, and 50 - 100 ng of template DNA. Twelve primer pairs were used to amplify anonymous microsatellite loci: 11 of these (loci Tt1.A06, Tt2.DO1, Tt2.H08, Tt1.C10, Tt2.F04, Tt2.F07, Tt1.E11, Tt2.B02, Tt1.B01, Tt1.F02, and Tt2.B07) were described from *Tor tambroides* (Nguyen et al., 2007) while the locus MFW7 was derived from *Cyprinus carpio* (Crooijmans et al., 1997) **(Table 2)**. The forward primers were labelled with the TAMRA/HEX/FAM/ROX dye at the 5' end. Thermal cycling protocol for all loci was as follows: initial denaturation at 94 °C for 3 min followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 48 °C for 30 s, and extension at 72 °C for 30 s, with a final extension step at 72 °C for 10 min.

e-ISSN 2716-697X

Table 2: The 12 primers that were used in this study indicating the locus, the repeat motifs, the expected size of the amplicons and the primer sequence.

Locus	Repeat Motif	Size (bp)	Primer Sequence (5'-3')
A06	(ATTT)3(GTTT)4(ATT)8	217-231	F: CCGAAATGCATTCTTGTCTT R: GGACTGACACTGGGGATCAT
DO1	(TG)15	210-248	F: CCATTACGCCTTTGGAGTGT R: TGGGAGATGTTGTTTCTCCA
H08	(TC)14	178-188	F: GGCTGTGAATGTGTTTGTGG R: GCCAGGATGATGAGCATGTA
C10	(TG)13	181-217	F: GCTGAAGCAGGTGAATCTGA R: TGATGCCTGTCAAACCTGTG
F04	(AC)11	120-150	F: ATGCCAGCTACAGGTCCAAT R: CGTGTGTATGATGCCACCTC
F07	(AG)9	141-180	F: GAGACGACTCTAGTCGCTGACA R: GTGTGGCCAGTGTAGCTGAA
E11	(TG)23	146-230	F: GAGTCCCTACAGACGTATTTCCA R: TTCAGCTCAGAGGGGACACT
B02	(TG)15...(TG)3	137-203	F: CTGGGAACGTCAGTTTACGG R: GTCCCCACAAGGATAGCAGA
B01	(AC)8AT(AC)3...(TC)3	245-265	F: GAGGGGCATTTTGTTCTTGA R: GCTTCCCCTCATAAGCCTTC
F02	(TG)2TA(TG)4TC(TG) 2	246-258	F: CATGGACCAAATTACAAGGATTT R: AACCTGTGAGGGATGTCCAG
B07	(TG)3(TC)6TT(TC)3T T(TC)6	225-255	F: TGGAAATTGAGACAAAGCTTCA R: TATGTGGTTTCAGGCAGCAG
MFW	CA	284-384	F: TACTTTGCTCAGGACGGATGC R: ATCACCTGCACATGGCCACTG

e-ISSN 2716-697X

All the DNA amplifications were performed in a Mastercycler® gradient (Eppendorf AG, Germany). Upon successful amplification of microsatellite, PCR products were dispatched for fragment analysis service using Genetic Analyser ABI3730XL system (Applied Biosystems). Aliquots of amplified products with different fluorescent dyes were multiplexed. One microlitre of this mixture was mixed with 9 µL of highly deionised formamide spiked with GeneScan 500 LIZ Size Standard (Applied Biosystems) and denatured at 95 ºC for 5 min. The amplicons were resolved on an ABI 3730XL Genetic Analyser.

For microsatellite analyses, allele scoring, and genotyping of microsatellite traces was performed using Geneious Prime (Biomatters, 2014). Binning of microsatellite allele lengths was done using Tandem (Matschiner & Salzburger, 2009). For each locus in each population sample MICRO-CHECKER 2.2.3 (Van Oosterhout et al., 2004) was used to detect the presence of null alleles, large allele drop out and stuttering. ML-NullFreq (Kalinowski & Taper, 2006), a maximum likelihood estimator of the frequency of null alleles in a sample with or without missing data was used to estimate the frequency of null alleles for each locus in each population. To account for the bias introduced by null alleles, we utilized the ENA (Excluding Null Alleles) correction method implemented in the software FreeNA (Chapuis and Estoup, 2007) to estimate global and pairwise Fst values. These corrected values were compared against uncorrected estimates to ensure the robustness of the population structure inference. GENEPOP 4.2 (Raymond & Rousset, 1995) was employed at each locus for each population to test for statistically significant departures of genotype frequencies from expectations under Hardy–Weinberg equilibrium (HWE) using Fisher's exact test with P-values estimated via Markov Chain with 10 000 dememorizations, 200 batches, 5000 iterations and Bonferroni correction. Linkage between microsatellite loci was investigated using GENEPOP 4.2 The null hypothesis was tested using the Bonferroni correction (Rice et al., 2008). Freedom from selection pressure was verified using Ewens–Watterson test of neutrality (Slatkin, 1994) that involves testing the null hypothesis that observed homozygosity value for each locus lies within the 95% confidence interval (CI) of the expected homozygosity estimates provided by 1,000 simulated samples. Phylogeographical structuring among the populations using the microsatellites marker was inferred using a distance-based method. A matrix of pairwise genetic distance (Takezaki & Nei, 1996) was generated in Microsatellite Analyser (MSA) v4.05 (Dieringer & Schlötterer, 2003) with 10,000 iterations. A majority-rule consensus tree from 10,000 sets of unweighted paired group method (Dieringer & Schlötterer, 2003) with arithmetic mean (UPGMA) dendrogram was constructed using PHYLIP v3.965 (Felsenstein, 1989). To infer the directionality and magnitude of contemporary gene flow between river basins, we performed a Bayesian assignment test using the software [BAYESASS v3.0 / Migrate-n]. This analysis distinguishes between historical connectivity and recent migration rates (m). Additionally, evidence for recent genetic bottlenecks was rigorously tested using BOTTLENECK v1.2.02, applying the Two-Phase Model (TPM) with 95% single-step mutations and 5% multi-step mutations, as this model is most appropriate for microsatellite data.

e-ISSN 2716-697X

Population Structure Inference

Population genetic clustering was further assessed using STRUCTURE v2.3.4 under the admixture model with correlated allele frequencies. Ten independent runs were performed for K = 1–10, each with 100,000 burn-in iterations followed by 500,000 MCMC iterations. The optimal number of clusters (K) was determined using the ΔK method implemented in STRUCTURE HARVESTER.

Isolation by Distance

Isolation by distance was evaluated using a Mantel test implemented in GenAlEx v6.5, correlating pairwise genetic distances (F_{st}) with geographic distances (river-to-river) based on 10,000 permutations to assess significance.

RESULTS AND DISCUSSION

The microsatellite loci selected for this study were successfully employed across all the individuals across populations, with the following exceptions. Locus B02 did not show discernible amplicons in 34 individuals across ten populations (6-SLU, 8-SKI, 4-SKIR, 2-WCTE, 2-WCBA, 2-WCTO, 3-KSO, 7- KSU, 1-IRU, 2-KMP). Locus A06 did not show amplification patterns in 1 individual from SKI. Microchecker detected null alleles for D01 (19-28%), H08 (32-35%), C10 (16-32%), E11 (16-25%) and B02 (23-33%) at certain populations **(Table 3).**

Table 3: The total number of samples, number of alleles, observed heterozygosity, expected heterozygosity, and Hardy Weinberg Equilibrium observed at each of the 12 loci used in this study. Two of the loci (DO1) and (MFW7) were excluded from the statistical analysis due to the ambiguity in their amplification profile.

Locus	N	Na	Ho	He	HWE
A06	174	8	0.661	0.667	0.032
DO1*	175	22	0.726	0.889	0.005
H08*	173	10	0.225	0.499	0.000
C10*	175	23	0.646	0.920	0.001
F04	174	5	0.316	0.368	0.992
F07	174	4	0.155	0.164	0.831
E11*	174	46	0.506	0.945	0.000

B02*	141	22	0.468	0.777	0.000
MFW7*[a]	175	6	0.434	0.592	0.003
B01	175	13	0.566	0.804	0.549
F02[b]	175	2	0.011	0.011	NA
B07*	175	11	0.703	0.836	0.000

*The locus departed significantly from Hardy–Weinberg equilibrium after Bonferroni correction. N: sample size, Na: number of alleles, Ho: observed heterozygosity, He: unbiased expected heterozygosity, HWE (P): Probability at $P<0.05$,

NA: tests were not applicable, HS: Highly significant. [a]Excluded due to possible misgenotyping,

[b]Excluded due to being monomorphic

Higher null alleles frequencies were recorded when stutters were present. The locus E11 was reported to be not reliable and was rejected from the analysis to avoid errors in interpretation of the results. Post sequential Bonferroni correction, with ($P<0.05$), 16 of 216 comparisons were determined to deviate significantly from Hardy–Weinberg equilibrium exact tests. After computing the fraction of significant test on each vector and overall combined of P value using Fisher's method, it was shown that all HWE deviation could be traced to 5 loci (H08, B02, D01, C10, E11) with departures representing heterozygotes deficiencies.

F02 was almost exclusively homozygous (99%) and therefore regarded as monomorphic. Thus, ten loci were retained in subsequent analysis (A06, DO1, H08, C10, F04, F07, B02, MFW7, B01 and B07) after confirming neutrality and absence of linkage. The markers were polymorphic, yielding 124 alleles combined and with an average of 15 alleles per locus. The lowest number of alleles was observed in locus F07 with 4 and the highest in locus C10 with 23. Locus DO1 was the most polymorphic with respect to heterozygosity (Ho = 0.726) while F04 was the least polymorphic (Ho = 0.155) (Table 4).

Table 4: Bottleneck analysis

Subpopulation	Wilcoxon Test P (one tail for Hex)
River Kanarom	
Marak-Parak (KMP)	0.752
Sunsui (KSU)	0.809
Tangkol (KTA)	0.344
Sorinsim (KSO)	0.344
River Sugut	
Luanti (SLU)	0.722
Poring (SPO)	0.455
River Labuk	
Kinarasan (SKI)	0.125
Paus (SPA)	0.410
River Kinabatangan	
Kironggu (SKIR)	0.180
Pinipi (SPI)	0.545
River Pagalan	
Rugading (IRU)	0.320
Menserulong (IME)	0.248
River Kadamaian	
Tokulung (WCTO)	0.156
Lingkubang (WCLI)	0.500
River Tuaran	
Gontung (WCGO)	0.125
River Moyog	
Babagon (WCBA)	0.014
River Kimanis	
Doingin (WCDO)	0.273
River Bongawan	
Telantang (WCTE)	0.064

P, probability, Hex, heterozygosity excess. ($P<0.05$).

e-ISSN 2716-697X

The subpopulation in IME possessed the highest genetic diversity with respect to the expected heterozygosity and allelic richness (n = 9, uHE = 0.549, AR = 2.316). Allelic richness (Ar) is based on the smallest sample size for a site of 2 diploid individuals. Samples from KSU had the lowest genetic diversity (n = 8, uHE = 0.325, AR = 1.677). An insignificant excess and deficient heterozygote were observed in 12 (FIS: –0.018 to –0.875) and 8 localities (FIS: 0.002 to 0.362), respectively. After sequential Bonferroni correction, with ($P<0.05$), two localities departed significantly from HWE; WCGO and WCTE. Results from the BOTTLENECK analysis (Table 4) indicated that across nine populations, there was no evidence of genetic variation. Specifically, the lack of significant heterozygosity excess under the TPM model confirms the absence of recent severe bottlenecks in the majority of populations.

The graphical representation (Figure 3) shows the geographic distance-based substructure of the populations as two different groups that branch into three discrete clades. Rivers that originate in the Crocker Range and drain into the South China Sea are represented by the Western clade. The five rivers that flow into the Sulu Sea are host to populations from the Eastern clade.

e-ISSN 2716-697X

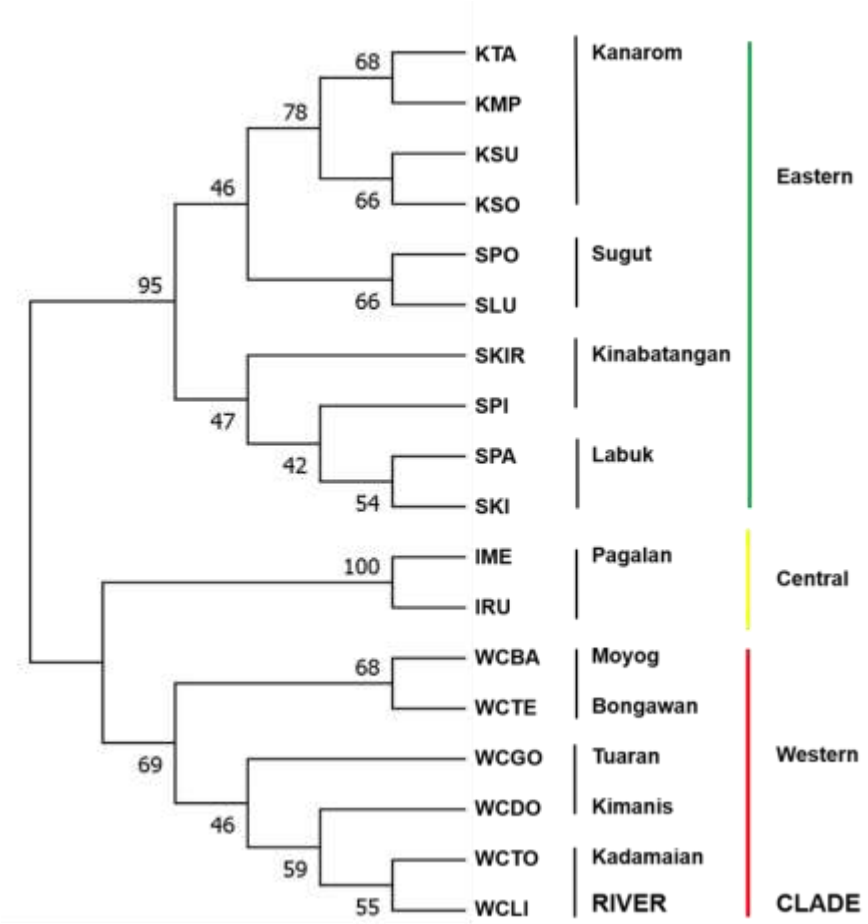

Figure 3: Majority-rule consensus tree of Pelian populations showing geographic structuring with some of the branches with lower statistical support (42–100%) in which two distinct major groups correspond to two independent evolutionary clades: (1) the western clade representing populations from the rivers Kadamaian, Tuaran, Moyog, Kimanis, and Bongawan; (2) the eastern and central clade representing populations from the rivers Kanarom, Sugut, Labuk and Kinabatangan, and Pagalan. The bootstrap values are shown in percent in the nodes

Directional Gene Flow:

Bayesian migration analysis revealed asymmetric gene flow patterns. Migration rates were generally low ($m < 0.05$), consistent with high Fst values. However, significant directional flow was observed from upstream to downstream populations in the Eastern clade, suggesting unidirectional dispersal likely driven by water currents and limited upstream migration due to physical barriers. The amount of variance explained by each model (Models 1, 2, and 3; populations grouped by current river distributions; and Model 4; populations grouped by current river distributions relative to the Crocker Trusmadi Range separation) was estimated using AMOVAs for the microsatellites data set

e-ISSN 2716-697X

In comparison to groups of five (70.27%, F_{st} = 0.230, p = 0.000 0.000), six (70.77%, F_{st} = 0.293, p = 0.000 0.000), and two (66.34%, F_{st} = 0.337, p = 0.000 0.000), rivers grouped as ten accounted for the biggest variability (72.21%, F_{st} = 0.278, p = 0.000 0.000). According to the AMOVA data, nearly all of the variation (70%) occurred within the populations, suggesting extensive gene flow between groups.

There was a significant but weak correlation between genetic and geographical distances (R2 = 0.206), with the exception of populations within a single riparian system, pairwise F_{st} suggested substantial genetic variance across the populations (F_{st} = 0.575-1.000, p 0.05) (Figure 4).

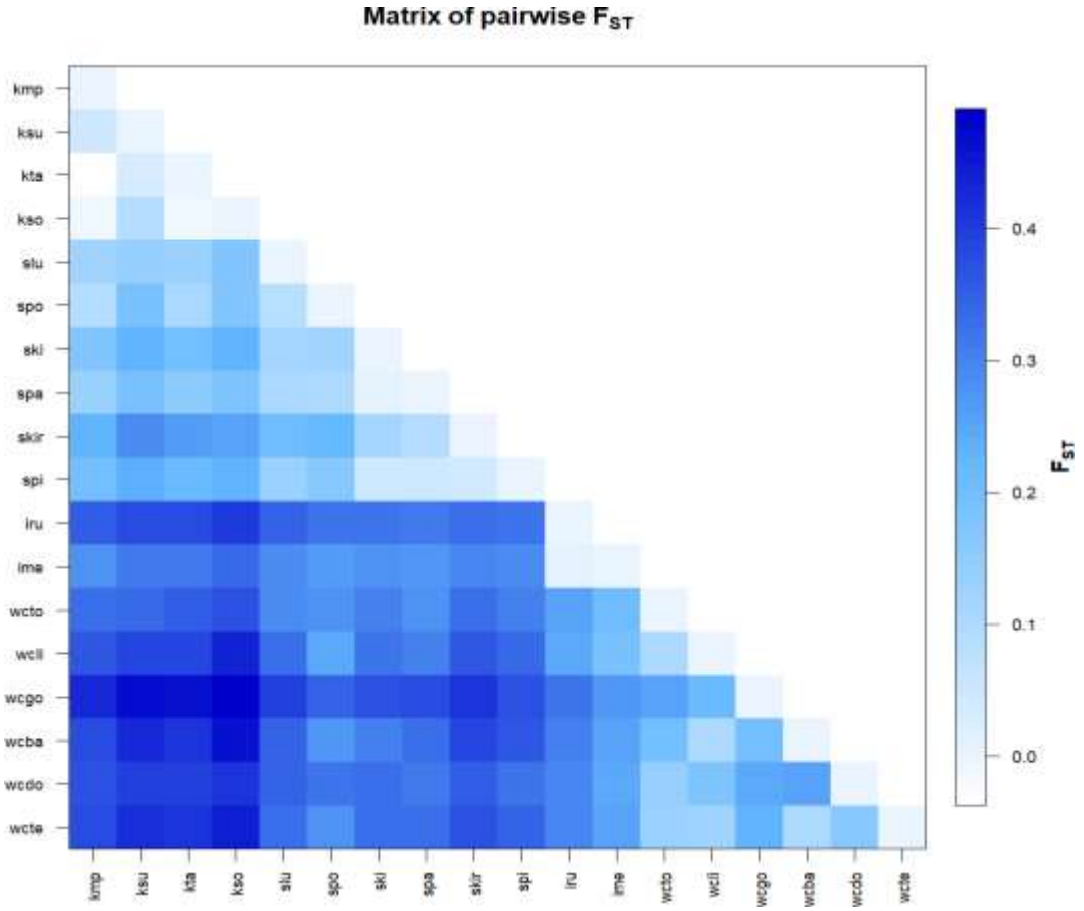

Figure 4. Heat map of pairwise F_{st} estimates for each locality. Pairwise F_{st} estimates for each locality, using microsatellites; dark blue squares represent high F_{st} values, and light blue squares represent low F_{st} values between localities.

Using microsatellite markers, the main genetic clusters, K with respect to ΔK, were at K = 2 (ΔK = 426.51) where grouping of the 175 individuals was consistent with the two major clades outlined using the UPGMA clustering method. The first cluster consisted of individuals from Clade 1 (Q = 97.1 – 99.6%) while the second cluster was made up of clade 2 (Q = 81.9 – 99.6%). Relatively weak differentiation was observed at K = 4 (ΔK = 43.35). The smallest mean value of Ln[P(X|K)] was at K = 1 {Ln[P(X|K)] = –5,570} before plateauing at K = 5 to K = 10 with mean Ln[P(X|K)] between -4138.89 and -3895.94. At K = 4, the proportion of ancestry in cluster one and cluster two was ranging from 40% to 98%, indicating some degree of genetic exchange between clusters whereas cluster three and four consistently above 78% indicating limited genetic exchange between the genetic clusters. The first cluster comprised individuals from KMP, KSU, KTA, KSO, SLU and SPO (Q = 34.0 – 98.1%), the second K comprised SLU, SPO, SKI, SPA, SKIR, and SPI (Q = 40.6 – 97.9%), the third K comprised IRU and IME (Q = 92.6 – 98.0), and the fourth K comprised WCTO, WCLI, WCGO, WCBA, WCDO, and WCTE (Q = 77.8 – 98.7%).

During the initial process of screening the microsatellite loci for deviations from HWE, the robustness of the HWE and Fis results were evaluated by comparing results with and without the inclusion of loci that did not amplify. Fis and HWE analysis were run with the complete 12 loci but removing a different locus at a time for each new analysis using the same parameters to identify the level of information provided by each locus on the level of HWE and Fis determination. When E11 was included in the determination of HWE across 18 populations, 8 populations deviated from HWE. The departures in these populations reflected a consistent deficiency of heterozygotes. Factors that generally result in heterozygote deficiencies include null alleles, the Wahlund effect and inbreeding. All the loci were affected equally thus eliminating the possibility of inbreeding being the cause of heterozygote deficiency. When this locus was discarded, those populations were in HWE except for 2 populations (WCGO and WCTE). E11 is suspected to be the source of HWE deviation to the populations because it is an unreliable locus to genotype which might lead to errors. Nguyen (2006) reported that this locus was unreliable in his study and our findings in this study concurred with his observations, leading to the decision to discard this locus.

The signature of the Wahlund effect was not detected in the multilocus analysis across loci and subsamples. The Wahlund effect registers concordantly across loci instead of just one or a few whereas null alleles are locus specific. Wahlund effect is a phenomenon when two genetically distinct groups are inadvertently grouped into a single sampling unit, either because they co-exist but rarely interbreed, or because the spatial scale chosen for sampling a site is larger than the true scale of a population, there will be more homozygotes than expected under HWE. Often the mutations that lead to the formation of null alleles will only occur in one or a few populations, so a heterozygote deficit might not be apparent across all populations in which this pattern is detected. Global significant deviation from Hardy–Weinberg proportions showed excessive variation across loci of F-statistics which represents a signature of null alleles that also recovered in this study. Thus , the factor that is most likely to produce HWE deviations in WCTE and WCGO population islikely to be presence of null alleles, and this has been reiterated by

the divergent conclusions can be arrived at when comparing biological significance and statistical significance (Waples & Allendorf, 2015).

Inbreeding is quite common in *Tor,* however, in this study, inbreeding was not evident in *T. tambra* populations which was represented by insignificant global Fis value across populations. There is very little evidence to suggest that there was a significant bottleneck event in the past that impacted the composition of the population and genetic diversity in Sabah, despite the decline in abundance observed in prior decades. The latest discovery appears to confirm that *T. tambra* is resistant to challenges presented by human interference in the riparian ecosystems examined in this study.

The population IME exhibited the highest microsatellite diversity (uHe=0.549, Ar=2.316) and KSU the lowest diversity uHE = 0.325, AR = 1.677). Across nine populations, there was no evidence of genetic variation, regardless of sample size. Nuclear loci exhibit biparental inheritance and allow different alleles to be distributed and combined at higher rates (Allendorf, 2017; Freeland et al., 2000). This suggests that colonization began in Sandakan before populations dispersed to eastern clades.

The major clades are in congruence with the results of the study of *Channa striata* (Robert et al., 2019),in which it has been reported that the populations are subdivided into two sub structured populations that are referred to as western and eastern clusters, respectively. The major clade grouping and the separation of Sabah into two parts by the Crocker and Trusmadi Range (CTR) coincide, even though the major river systems and the subclades do not. The historical and current linkages between rivers have a significant impact on the phylogeographical distribution of freshwater fish species, which is used to assess the genetic links between populations that have been detected. The main geological barrier separating Sabah's western and eastern riverine and floodplain environments is represented by the CTR. The Crocker Range (CR), which spans nearly 200 km and has mountains up to 4000 m in height, is the tallest range in Sabah's hilly western region. It has an average elevation of 2000 m and is more than 40 km broad. The second-highest mountain peak in Malaysia is located inside the 80 km Trusmadi Range (TR), which is next to the CR. On Sabah's east, middle, and north coasts, rivers that drain into the South China Sea originate in CTR and its mountains. Lower mountain ranges and plains with intermittent hills separating significant rivers like Kinabatangan River, Labuk River, and Sugut River, which flow into the Sulu Sea, extend into the western beaches, southern plains, and the interior or central part of Sabah. The eastern and central rivers range in length from 100 to 560 km, with the Kinabatangan River being the longest and widest floodplain. In general, the length of the western and northern rivers is less than 100 km. With the elevation of the Borneo Mountain ranges, the division of the eastern, central, northwestern, and southwestern clades took place in the late Miocene and Pliocene (5–1 mya). Through dispersal and subsequent population diversification brought on by geographic isolation, climate variations throughout the Pleistocene may have affected modern species distribution patterns. "The major clade grouping coincides with the Crocker and

Trusmadi Range (CTR). It is crucial to distinguish the temporal signals of different markers: while mitochondrial DNA (e.g., Cytb or D-loop) typically reflects ancient divergence events driven by geological barriers like the CTR, our microsatellite data reveals finer-scale, contemporary population structuring. The discordance between broad historical separation (mtDNA) and the highly fragmented current population structure (microsatellites) highlights that recent river fragmentation acts as a potent barrier to gene flow, superimposed on the historical vicariance templatesAs an alternative, the Pleistocene's low sea level permitted dispersal at river mouths and eventual isolation during sea level increases. Isolation by distance (IBD) (Briñoccoli et al., 2021), isolation by barriers (IBB) (Tsuji et al., 2022), and isolation by resistance (IBR) (De Queiroz et al., 2017) have been reported to influence genetic variation in riverine freshwater fishes.

The historical isolation by CTR and the current physical isolation of the rivers following the SHM, respectively, are demonstrated by the notable high genetic differentiation between clades and river systems, which are the causes of genetic variances in Tor populations. Additionally, the Mantel test revealed that, according to the IBD model, geographic distance had an impact on the genetic makeup of Tor populations. Due to the rugged terrain in Sabah, the majority of the rivers are divided from one another by mountain ridges; as a result, the populations along adjoining rivers are not connected genetically. However, some populations from geographically isolated but contemporaneously separated river basins, such as WCBA (Moyog River) and WCGO (Tuaran River), WCDO (Kimanis River) and WCTE (Bongawan River), SKIR (Kinabatangan River) and SKI, SPA (Labuk River), were found to be highly homogeneous. This finding lends support to the idea that there are gene flows between rivers. Using microsatellites, it was also found that the Moyog and Tuaran populations have a significant degree of affinity. Dispersal via a geomorphological phenomenon, such as river capture, stream piracy, flood occurrences, the historical connection between river systems, and manmade activity, including restocking programs and translocation, could account for the lack of genetic patterning.

If a population or group of populations has been reproductively isolated for long enough to contain unusual evolutionary combinations that are unlikely to revolve around an ecological time scale, then it fits the requirements for an evolutionarily significant unit (ESU) (Hallerman & Hilsdorf, 2014; Hutama et al., 2017). The analysis of data obtained from microsatellite markers implicated that the genetic differentiation between ESU is due to barriers by CTR. This study suggests that three clades be designated as ESU since they are all historically isolated, monophyletic, and have a high degree of genetic differentiation.

The Eastern clade comprises the Rivers Kanarom, Sugut, Kinabatangan, and Labuk. The Central clade encompasses the Pagalan River, the Western clade includes the of Rivers Bongawan, Kimanis, Moyog, Tuaran and Kadamaian. Short-term management depends on the identification of Management Units (MU), which can be used to direct current management initiatives including stocking programmes, quota allocation, alternative harvesting scenario modelling, and monitoring programme design. Populations from different river systems may have become isolated from each other due to habitat fragmentation.

Furthermore, based on our gene flow analysis, translocation strategies should prioritize 'upstream' conservation. Since downstream populations show higher admixture, they may serve as sources for genetic rescue, but care must be taken not to disrupt the unique local adaptations of isolated upstream populations (ESUs). Future barrier mitigation (e.g., fish ladders) should target areas where our Bayesian analysis identified blocked migration pathways. Consequently, gene flow might not be able to reverse genetic depletion if these populations were significantly affected by specific conditions, including overfishing. To maintain the persistence of the local populations that make up a fished stock, fisheries managers should manage local populations independently so that a sufficient proportion of individuals from each local population escape catch and reproduce.

Populations of MU show less evolutionary divergence than reciprocal monophyly and have significantly different allele frequencies at numerous loci (Moritz, 1999); however, allele frequency differentiation cannot be taken as direct proof of demographic independence (Palsbøll et al., 2007) proposed that the amount of genetic divergence at which populations become demographically independent should be used to identify MUs from population genetic data. MU status would then be given when the observed estimate of genetic divergence is significantly higher than a predetermined threshold value (Karjalainen et al., 2022) . Microsatellite results for Tor populations in Sabah appear to be significantly structured into geographically distinct subpopulations across major drainages, supporting data from mtDNA, where inter-stock dispersal and migration patterns were constrained by habitat fragmentation. (Biun et al., 2021). Gene flow is at least somewhat limited between populations, yet it still occurs frequently enough to guarantee genetic structure is generally homogenised to 5 genetic clusters.

The Department of Fisheries could use watersheds to delineate provisional MU. A similar study on Brook Trout, that reported on the delineation of structure based on riparian watersheds revealed that each watershed contained a population that was genetically distinct and demographically independent (Habera & Moore, 2005). Additionally, significant positive correlations between genetic differentiation and geographic distance among Tor populations might generate unique adaptations in each population due to variation of habitat conditions and contemporary environmental conditions. Studies has reported that local adaptation can contribute to the evolution of reproductive barriers and differences in life (Rajkov et al., 2018). *Tor* appear to be river specific in their migrations, indicating that the subpopulations acquired behavioural and or physiological adaptations to each river (Haryani, 2022; Hestirianoto et al., 2021).

Although studies of local adaptations have not been done to support this, minor genetic differences at neutral genomic loci, are indicators of fixation of alleles that can be interpreted as adaptations to a unique ecological niche (Larsen et al., 2007). Isolated populations undergo evolution in situ over time and can acquire distinct traits as well as reproductive barriers that curtail any attempt at artificial breeding with recruits from other populations.. The ten rivers investigated in for study were categorised as separate management unis of Tor all over Sabah drainage. The large spatial distribution and the diversity of ecological niches are representative

e-ISSN 2716-697X

of unique opportunities for evolution and adaptation and the outcome of this study recommend that the populations in each management unit be divided into multiple zones to conserve the locally adapted populations from being stocked with individuals from other populations.

CONCLUSION

The objective of this study was to assess the population diversity of Pelian in the riparian ecosystems of Sabah from the perspective of conservation of the species and the subsequent establishment of Management Units and Conservation Units. This was achieved by the application of 12 microsatellite loci from published literature across ten river systems. There was a clear sub structuring of the populations based on geographic barriers that supports existing theories of speciation by isolation. The findings of this study will serve as a reference for monitoring of the populations and contribute to the development of a large scale conservation and breeding program in Sabah.

ACKNOWLEDGEMENTS

This research was funded by UMSGreat, grant number **GUG0166-02/2017** and Department of Fisheries Sabah, grant number **D28 0005 0024**.

REFERENCES

Abdul-Muneer, P. M. (2014). Application of Microsatellite Markers in Conservation Genetics and Fisheries Management: Recent Advances in Population Structure Analysis and Conservation Strategies. *Genetics Research International*, *2014*. https://doi.org/10.1155/2014/691759

AHMAD, A., TN-NEK, S. A. R., & AMBAK, MOHD. A. (2006). Preliminary study on fish diversity of ulu tungud, meliau range, sandakan, sabah. *Journal of Sustainability Science and Management*, *1*(2).

Allendorf, F. W. (2017). Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology*, *26*(2). https://doi.org/10.1111/mec.13948

Biun, H., Sade, A., Robert, R., & Rodrigues, K. F. (2021). Phylogeographic structure of freshwater tor sp. In river basins of sabah, Malaysia. *Fishes*, *6*(4). https://doi.org/10.3390/fishes6040044

Briñoccoli, Y. F., Jardim de Queiroz, L., Bogan, S., Paracampo, A., Posadas, P. E., Somoza, G. M., Montoya-Burgos, J. I., & Cardoso, Y. P. (2021). Processes that drive the population structuring of Jenynsia lineata (Cyprinidontiformes, Anablepidae) in the La Plata Basin. *Ecology and Evolution*, *11*(11). https://doi.org/10.1002/ece3.7427

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. Molecular biology and evolution, 24(3), 621-631.

Crooijmans, R. P. M. A., Bierbooms, V. A. F., Komen, J., Van Der Poel, J. J., & Groenen, M. A. M. (1997). Microsatellite markers in common carp (Cyprinus carpio L.). *Animal Genetics*, *28*(2). https://doi.org/10.1111/j.1365-2052.1997.00097.x

De Queiroz, L. J., Torrente-Vilara, G., Quilodran, C., da Costa Doria, C. R., & Montoya-Burgos, J. I. (2017). Multifactorial genetic divergence processes drive the onset of speciation in an Amazonian fish. *PLoS ONE*, *12*(12). https://doi.org/10.1371/journal.pone.0189349

Dieringer, D., & Schlötterer, C. (2003). MICROSATELLITE ANALYSER (MSA): A platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, *3*(1). https://doi.org/10.1046/j.1471-8286.2003.00351.x

Esa, Y. B., Siraj, S. S., Daud, S. K., Ryan, J. J. R., Rahim, K. A. A., & Tan, S. G. (2008). Molecular systematics of mahseers (Cyprinidae) in Malaysia inferred from sequencing of a mitochondrial Cytochrome C Oxidase I (COI) gene. *Pertanika Journal of Tropical Agricultural Science*, *31*(2).

Felsenstein, J. (1989). {PHYLIP} -- Phylogeny Inference Package ({V}ersion 3.2). *Cladistics*, *5*.

Freeland, J. R., Romualdi, C., & Okamura, B. (2000). Gene flow and genetic diversity: A comparison of freshwater bryozoan populations in Europe and North America. *Heredity*, *85*(5). https://doi.org/10.1046/j.1365-2540.2000.00780.x

Habera, J., & Moore, S. (2005). Managing southern Appalachian brook trout: A position statement. *Fisheries*, *30*(7).

Hallerman, E., & Hilsdorf, A. W. S. (2014). Conservation genetics of tilapias: Seeking to define appropriate units for management. *Israeli Journal of Aquaculture - Bamidgeh*, *66*. https://doi.org/10.46989/001c.20769

Haryani, G. S. (2022). Migratory freshwater fish in Indonesia: Threats and conservation efforts. *IOP Conference Series: Earth and Environmental Science*, *1062*(1). https://doi.org/10.1088/1755-1315/1062/1/012001

Hestirianoto, T., Sulistiono, Wildan, D. M., Simanjuntak, C. P. H., Rahardjo, M. F., Syafruddin, D., Nugroho, T., & Ismanto, A. J. (2021). Application of radiotelemetry in tracking the movement of mahseer fish (Tor spp.) in Batang Toru River, North Sumatera Province Indonesia. *IOP Conference Series: Earth and Environmental Science*, *744*(1). https://doi.org/10.1088/1755-1315/744/1/012028

Hutama, A., Dahruddin, H., Busson, F., Sauri, S., Keith, P., Hadiaty, R. K., Hanner, R., Suryobroto, B., & Hubert, N. (2017). Identifying spatially concordant evolutionary significant units across multiple species through DNA barcodes: Application to the conservation genetics of the freshwater fishes of Java and Bali. *Global Ecology and Conservation*,*12*. https://doi.org/10.1016/j.gecco.2017.11.005

Ingram, B., Sungan, S., Tinggi, D., Sim, S. Y., & De Silva, S. S. (2007). Breeding performance of Malaysian mahseer, Tor tambroides and T. douronensis broodfish in captivity. *Aquaculture Research*, *38*(8). https://doi.org/10.1111/j.1365-2109.2007.01716.x

Kalinowski, S. T., & Taper, M. L. (2006). Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. In *Conservation Genetics* (Vol. 7, Issue 6). https://doi.org/10.1007/s10592-006-9134-9

Karjalainen, J., Sjövik, R., Väänänen, T., Sävilammi, T., Sundberg, L. R., Uusi-Heikkilä, S., & Marjomäki, T. J. (2022). Genetic-based evaluation of management units for sustainable vendace (Coregonus albula) fisheries in a large lake system. *Fisheries Research*, *246*.https://doi.org/10.1016/j.fishres.2021.106173

Larsen, P. F., Nielsen, E. E., Williams, T. D., Hemmer-Hansen, J., Chipman,J. K., Kruhøffer, M., Grønkjær, P., George, S. G., Dyrskjøt, L., & Loeschcke, V. (2007). Adaptive differences in gene expression in European flounder (Platichthys flesus). *Molecular Ecology*, *16*(22). https://doi.org/10.1111/j.1365-294X.2007.03530.x

Matschiner, M., & Salzburger, W. (2009). TANDEM: Integrating automated allele binning into genetics and genomics workflows. *Bioinformatics*, *25*(15). https://doi.org/10.1093/bioinformatics/btp303

Moritz, C. (1999). Conservation units and translocations: Strategies for conserving evolutionary processes. In *Hereditas* (Vol. 130, Issue 3). https://doi.org/10.1111/j.1601-5223.1999.00217.x

Nguyen, T., Baranski, M., Rourke, M., & McPartlan, H. (2007). Characterization of microsatellite DNA markers for a mahseer species, Tor tambroides (Cyprinidae) and cross-amplification in four congeners. *Molecular Ecology Notes*, *7*(1). https://doi.org/10.1111/j.1471-8286.2006.01546.x

Palsbøll, P. J., Bérubé, M., & Allendorf, F. W. (2007). Identification of management units using population genetic data. *Trends in Ecology and Evolution*, *22*(1). https://doi.org/10.1016/j.tree.2006.09.003

Pinder, A. C., Britton, J. R., Harrison, A. J., Nautiyal, P., Bower, S. D., Cooke, S. J., Lockett, S., Everard, M., Katwate, U., Ranjeet, K., Walton, S., Danylchuk, A. J., Dahanukar, N., & Raghavan, R. (2019). Mahseer (Tor spp.) fishes of the world: status, challenges and opportunities for conservation. In *Reviews in Fish Biology and Fisheries* (Vol. 29, Issue 2). https://doi.org/10.1007/s11160-019-09566-y

Rajkov, J., Weber, A. A. T., Salzburger, W., & Egger, B. (2018). Immigrant and extrinsic hybrid inviability contribute to reproductive isolation between lake and river cichlid ecotypes. *Evolution*, *72*(11). https://doi.org/10.1111/evo.13612

Raymond, M., & Rousset, F. (1995). GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity*, *86*(3). https://doi.org/10.1093/oxfordjournals.jhered.a111573

Rice, T. K., Schork, N. J., & Rao, D. C. (2008). Methods for Handling Multiple Testing. In *Advances in Genetics* (Vol. 60). https://doi.org/10.1016/S0065-2660(07)00412-9

Robert, R., Amit, N. H., Sukarno, N. M., Majapun, R. J., & Kumar, S. V. (2019). Population genetic structure of Asian snakehead fish (Channa striata) in North Borneo: Implications for conservation of local freshwater biodiversity. *Ecological Research*, *34*(1). https://doi.org/10.1111/1440-1703.1008

Slatkin, M. (1994). An exact test for neutrality based on the Ewens sampling distribution. *Genetical Research*, *64*(1). https://doi.org/10.1017/S0016672300032560

Takezaki, N., & Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, *144*(1). https://doi.org/10.1093/genetics/144.1.389

Tsuji, S., Shibata, N., Sawada, H., & Watanabe, K. (2022). Differences in the genetic structure between and within two landlocked Ayu groups with different migration patterns in Lake Biwa revealed by environmental DNA analysis. *Environmental DNA*. https://doi.org/10.1002/edn3.345

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3). https://doi.org/10.1111/j.1471-8286.2004.00684.x

Walton, S. E., Gan, H. M., Raghavan, R., Pinder, A. C., & Ahmad, A. (2017). Disentangling the Taxonomy of the Mahseers (Tor spp.) of Malaysia: An Integrated Approach Using Morphology, Genetics and Historical Records. In *Reviews in Fisheries Science and Aquaculture* (Vol. 25, Issue 3). https://doi.org/10.1080/23308249.2016.1251391

Waples, R. S., & Allendorf, F. (2015). Testing for hardy-weinberg proportions: Have we lost the plot? In *Journal of Heredity* (Vol. 106, Issue 1). https://doi.org/10.1093/jhered/esu062