Zoological Studies **65:**53 (2026)

Mitochondrial Lineage Diversity and Genetic Structure of the Mahseers *Tor tambra–Tor tambroides* Complex in Thailand and Peninsular Malaysia

Krittaporn Lempan[1], Hong-Chiun Lim[2], Wei-Jen Chen[3], and Monwadee Wonglapsuwan[1,4,*]

[1]*Division of Biological Science, Faculty of Science, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand. E-mail: krittaporn.near1259@gmail.com (Lempan)*

[2]*Kedah Agro Ausin Berhad, Wisma Kedah Agro, Lot 2152 & 2153, Jalan Tunku Abdul Halim, Taman Tunku Habsah, 05000 Alor Setar, Kedah, Malaysia. E-mail: hongchiun_01@yahoo.com (Lim)*

[3]*Institute of Oceanography, National Taiwan University, No.1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan. E-mail: wjchen.actinops@gmail.com (Chen)*

[4]*Center for Genomics and Bioinformatics Research, Faculty of Science, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand. *Correspondence: E-mail: monwadee.wo@psu.ac.th (Wonglapsuwan)*

ORCID

Wei-Jen Chen: https://orcid.org/0000-0003-4751-7632

Monwadee Wonglapsuwan: https://orcid.org/0000-0002-9272-684X

(Received 22 April 2026 / Accepted 18 August 2026 / Published -- 2026)

Communicated by Benny K.K. Chan

The mahseers (*Tor* spp.) are among the most ecologically, culturally, and economically significant freshwater fishes in Southeast Asia. Despite their importance, the taxonomic distinction between *T.* cf. *tambra* and *T.* cf. *tambroides* in Thailand and Peninsular Malaysia remains problematic because of overlapping morphological characteristics and the absence of genetic reference sequences anchored to name-bearing type specimens. In this study, three mitochondrial markers, COX1, 16S rRNA, and the D-loop, were used to investigate phylogenetic relationships and genetic diversity among *Tor* populations from Thailand and Peninsular Malaysia. COX1 phylogenetic reconstruction and species-delimitation analyses recovered four principal mitochondrial groups. G1, provisionally designated as *T.* cf. *tambra*, contained the majority of the examined samples from Thailand and Peninsular Malaysia. G2 comprised samples from the Salween River and was treated as an unresolved *Tor* lineage. G3 was subdivided into G3-1, an unresolved mitochondrial lineage represented by two Malaysian samples and one GenBank sequence reported from Java, and G3-2, provisionally designated as *T.* cf. *tambroides* based on Sumatran GenBank sequences. G4 comprised GenBank sequences provisionally designated as *T.* cf. *douronensis*. Population genetic analyses based on D-loop sequences revealed strong geographic structuring among populations from northern, central, and southern Thailand and Peninsular Malaysia, with comparatively weaker

differentiation among populations within the same region. The occurrence of divergent mitochondrial lineages in aquaculture stocks, particularly at Hulu Langat and Endau, suggests possible broodstock translocation or mixing. Overall, our results reveal substantial mitochondrial diversity within the *Tor tambra–T. tambroides* complex and provide a genetic framework for future taxonomic, conservation, and aquaculture studies.

Keywords: *Tor tambra*, *Tor tambroides*, Mitochondrial DNA, Phylogenetics, Species delimitation

Citation: Lempan K, Lim HC, Chen WJ, Wonglapsuwan M. 2026. Mitochondrial lineage diversity and genetic structure of the *Tor tambra–Tor tambroides* complex in Thailand and Peninsular Malaysia. Zool Stud **65:**53.

BACKGROUND

The Cyprinidae family is among the most diverse groups of freshwater fishes, comprising 3076 species in 296 genera distributed worldwide (Fricke et al. 2025a). Within this family, the mahseers (*Tor* spp.) are particularly notable for their ecological, cultural, and economic importance across Southeast Asia (Nguyen et al. 2006; Pinder et al. 2019; Jaafar et al. 2021). A total of 17 species of mahseer are currently recognized (Fricke et al. 2025b). Typically, *Tor* species inhabit oxygen-rich rivers and lakes with fast-flowing currents and migrate upstream to rocky spawning habitats near waterfalls (Jaafar et al. 2021).

Tor tambra (Valenciennes 1842), locally known as pla pluang chomphu in Thailand (Surachat et al. 2022) or kelah in Malaysia (Esa et al. 2008), was originally described from Java, Indonesia. This species is characterized by a reddish to dark olive body, a blunt snout, and a relatively short lower median lobe (Jaafar et al. 2021). In contrast, *T. tambroides* (Bleeker 1854), originally described from Padang, West Sumatra, Indonesia, and locally known as pla wien numngen in Thailand, displays a more elongated lower lobe, a pointed snout, and a silver-to-bronze coloration with reddish hues and dark fins (Jaafar et al. 2021). However, previous research has shown that environmental variation can influence their morphology, and individuals of *T. tambra* have been observed with both short and long lower lobes (Walton et al. 2017). Consequently, species delimitation based solely on morphological characteristics remains problematic. In spite of this, both species are widely regarded as “true mahseer”, with distributions extending Laos, Myanmar, Vietnam, Thailand, Malaysia, and Indonesia (Jaafar et al. 2021).

In Thailand, *T. tambra* is considered a high-value aquaculture species, particularly in the southern provinces, Yala and Narathiwat, where it commands premium prices in the market. In

contrast, *T. tambroides* is primarily cultivated in the central and northern regions (Duangjai and Punroob, 2018). Similarly, Malaysia has actively promoted the cultivation of both species through government initiatives (Nguyen et al. 2006). At present, these fishes are sourced both from aquaculture and the harvest of wild juveniles, which are subsequently provided either for restocking programs or for commercial markets (Ramezani-Fard and Kamarudin 2012). However, effective stock management and conservation measures rely on accurate species identification, yet misidentification remains a persistent challenge. Dependence on local names and subtle morphological traits often results in confusion between *T. tambra* and *T. tambroides*. Species identification within the genus *Tor* has long been problematic due to overlapping morphological characteristics and inconsistent taxonomic interpretations (Roberts 1999; Kottelat 2013; Walton et al. 2017). This issue is particularly important because mahseer (*Tor* spp.) are among the most economically and culturally valued freshwater fishes in South and Southeast Asia (Ingram et al. 2007; Walton et al. 2017). In Malaysia, previous genetic work has also highlighted persistent confusion in the identification of local *Tor* species and demonstrated their high value for aquaculture, broodstock management, food, ornamental use, and recreational fisheries (Chew et al. 2021). Consequently, uncertainty in species identification raises questions about whether *Tor* populations from different regions truly belong to the same species or represent distinct taxonomic identities. Such uncertainties may influence not only trade value but also conservation planning and breeding strategies.

Genetic analyses—particularly DNA barcoding, which enables precise species identification and delimitation when morphological methods fall short—offer an effective approach for resolving this issue (Karim et al. 2016). DNA barcoding typically targets short, conserved mitochondrial regions (400–800 bp) that are easily amplified via Polymerase Chain Reaction (PCR) across diverse taxa, with the cytochrome *c* oxidase subunit I (COX1) gene being the most widely used marker in fish taxonomy (Ward et al. 2009; Bhattacharya et al. 2016; Cawthorn et al. 2012). Additionally, *COX1* has proven especially effective in distinguishing various “cryptic” species or species complexes, including *Cyprinus carpio* (Torgunakova et al. 2012), freshwater halfbeak (Lim et al. 2016), Pennah croakers (Lim et al. 2021), *Barbus xanthopterus* (Al-Khafaji et al. 2025), and *Tor* spp. that inhabit southeastern Asia (Ward et al. 2005; Ardura et al. 2010; Bingpeng et al. 2018; Haque et al. 2023). In addition, the mitochondrial 16S rRNA *gene*, which is characterized by conserved regions interspersed with variable site, has been broadly employed in phylogenetic reconstruction of teleost fishes, including *Tor* species (Nguyen et al. 2008; Quraishia et al. 2015; Alyamani 2024). Furthermore, the mitochondrial control region (D-loop), a fast-evolving non-coding segment, serves as a valuable marker for detecting population-level genetic variation (Brown et al. 1986; Sbisà et al. 1997; Biun et al. 2021; Fang et al. 2022).

Building on the DNA sequences from *COX1* and 16S rRNA genes, the present study aims to investigate the genetic diversity and phylogenetic relationships of *Tor* species centered on *T. tambra* and *T. tambroides* from both wild and farmed populations in Thailand and Peninsular Malaysia. By incorporating the D-loop sequences, we further assess whether the high-value *Tor* from southern Thailand is genetically distinct from populations in other regions. The resulting insights will strengthen species identification, delimitation of species distribution, support aquaculture management, and inform conservation strategies. Ultimately, this work contributes to the sustainable utilization of these culturally and economically valuable fish while ensuring accurate species authentication for future cultivation and trade.

MATERIALS AND METHODS

Sample collection

A total of 104 *Tor* spp. samples were analyzed in this study. The samples were collected from both aquaculture farms and wild populations in Malaysia and Thailand to represent the main distribution sources of this species in markets of the region. Farmed specimens (Fig. S1) were obtained from five aquaculture locations in Malaysia: Sungai Tembat (Terengganu), Sungai Loh (Terengganu), Endau (Pahang), Gua Musang (Kelantan), and Hulu Langat (Selangor). In Thailand, farmed specimens were collected from two aquaculture farms located in Yala and Phetchaburi provinces. In addition, wild-caught specimens were obtained from three natural locations in Thailand: Narathiwat, Kanchanaburi, and the Salween River. Among the 104 total samples analyzed, 15 specimens were derived from wild populations, while the remaining samples originated from aquaculture sources. The sampling sites are illustrated in figure 1, and the details of sample collection are listed in table 1. The morphology of all specimens was examined (Table S1) based on the key features (Roberts 1999; Kottelat 2013; Walton et al. 2017; Jaafar et al. 2021) to identify up to Genus level. The numbers of predorsal scales (Pred. S), scales above the lateral line (TRa), scales below the lateral line (TRb), and gill rakers were recorded for each specimen following the methods described by Maurice Kottelat (2013) and Tyson R. Roberts (1999), which are commonly applied in taxonomic studies of *To*r species. All fish were handled according to the Ethics guidelines by the Animal Research for Scientific Work guidelines (PSU Animal Standard, Prince of Songkla University, Hat Yai, Songkhla, Thailand) (Ref. AQ082/2025). The specimens were photographed using a digital camera. They were then euthanized with an overdose of anesthetics, 2-Phenoxyetanol (Sigma-Aldrich, Germany). The epaxial muscle tissues were excised

and preserved in 95% ethanol for DNA extraction. Subsequently, the specimens were fixed in formalin and stored in 70% ethanol for long-term preservation at the Faculty of Sciences, Prince of Songkla University, Thailand.

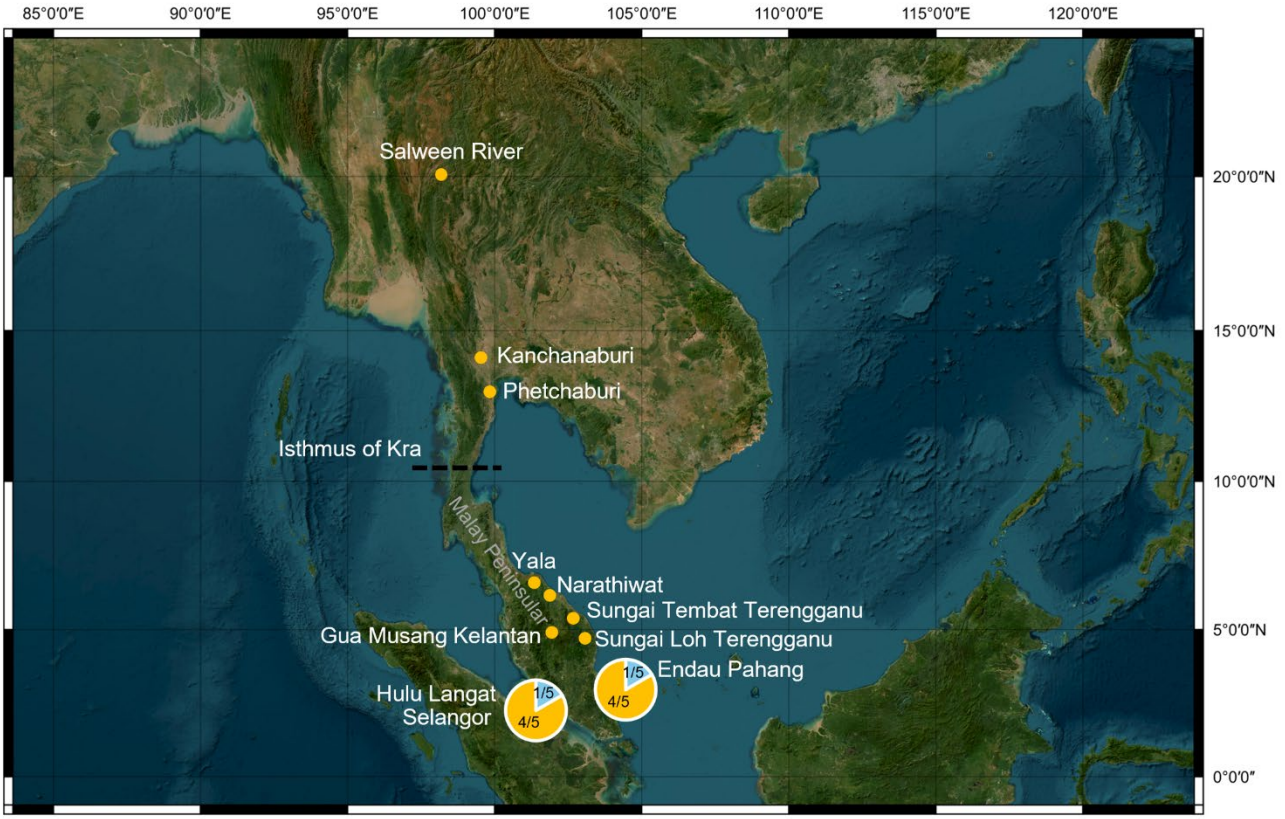

Fig. 1. Map of mainland Southeast Asia showing the sampling localities of *Tor* specimens examined in this study. Yellow circles indicate localities where all *COX1*-sequenced specimens belonged to the principal *T.* cf. *tambra* mitochondrial assemblage. Pie charts at Hulu Langat, Selangor and Endau, Pahang, show the mitochondrial-lineage composition among the five specimens with available *COX1* sequences at each locality: the yellow sector represents four specimens belonging to G1, whereas the white sector represents one specimen belonging to the taxonomically unresolved G3-1 lineage. Samples from Kanchanaburi, the Salween River, and Narathiwat were collected from the wild, whereas samples from the remaining localities were obtained from aquaculture farms. The dashed line indicates the approximate position of the Isthmus of Kra. The map was prepared using QGIS Geographic Information System (QGIS.org 2024), an Open Source Geospatial Foundation project.

Table 1. Sample information, including sample IDs, sampling locations in Thailand and Peninsular Malaysia, numbers of successfully amplified mitochondrial markers, and *COX1* phylogenetic group assignments based on figure 2. The number of samples is indicated by *n*. An asterisk denotes wild-caught individuals. A dash indicates that no *COX1* group could be assigned because a *COX1* sequence was unavailable. Groups G3-2 and G4 comprised only GenBank reference sequences and are therefore not represented among the newly collected samples in this table

No.	Sample ID	Location	n			Total	Group
			COX1	16S rRNA	D-loop		
1	Tor_TEM001	Sungai Tembat, Terengganu, Malaysia	5	5	5	5	G1
2	Tor_TEM002						
3	Tor_TEM003						
4	Tor_TEM004						
5	Tor_TEM005						
6	Tor_LOH001	Sungai Loh, Terengganu, Malaysia	5	5	10	10	G1
7	Tor_LOH002						
8	Tor_LOH003						
9	Tor_LOH004						
10	Tor_LOH005						
11	Tor_LOH006		–	–			–
12	Tor_LOH007						
13	Tor_LOH008						
14	Tor_LOH009						
15	Tor_LOH010						
16	Tor_KLT001	Gua Musang, Kelantan, Malaysia	5	5	10	10	G1
17	Tor_KLT002						
18	Tor_KLT003						
19	Tor_KLT004						
20	Tor_KLT005						
21	Tor_KLT006		–	–			–
22	Tor_KLT007						
23	Tor_KLT008						
24	Tor_KLT009						
25	Tor_KLT010						
26	Tor_HUL001	Hulu Langat, Selangor, Malaysia	5	5	9	9	G1
27	Tor_HUL002						
28	Tor_HUL003						
29	Tor_HUL004						
30	Tor_HUL005						G3-1
31	Tor_HUL006		–	–			G1
32	Tor_HUL007						
33	Tor_HUL008						
34	Tor_HUL009						
35	Tor_PHG001	Endau, Pahang, Malaysia	5	5	10	10	G1
36	Tor_PHG002						
37	Tor_PHG003						G3-1
38	Tor_PHG004						G1
39	Tor_PHG005						
40	Tor_PHG006		–	–			–
41	Tor_PHG007						

42	Tor_PHG008						
43	Tor_PHG009						
44	Tor_PHG010						
45	Tor_Sal1	Salween River*, Mae Hong Son, Thailand	4	4	4	4	G2
46	Tor_Sal2						
47	Tor_Sal3						
48	Tor_Sal4						
49	Tor_Phet01	Tha Yang, Phetchaburi, Thailand	5	5	10	10	G1-bis
50	Tor_Phet02						
51	Tor_Phet03						
52	Tor_Phet04						
53	Tor_Phet05						
54	Tor_Phet06		–	–			–
55	Tor_Phet07						
56	Tor_Phet08						
57	Tor_Phet09						
58	Tor_Phet10						
59	Tor_Kan001	Sangkhla Buri, Kanchanaburi*, Thailand	3	3	3	3	G1-bis
60	Tor_Kan002						
61	Tor_Kan003						
62	Tor_N1	Sai Buri River, Narathiwat*, Thailand	5	5	8	8	G1
63	Tor_N2						
64	Tor_N3						
65	Tor_N4						
66	Tor_N5						
67	Tor_N6		–	–			–
68	Tor_N7						
69	Tor_N10						
70	Tor_YL001	Muang, Yala, Thailand	5	5	5	35	G1
71	Tor_YL002						
72	Tor_YL003						
73	Tor_YL004						
74	Tor_YL005						
75	Tor_TBR2	Than To, Yala, Thailand	2	2	4		G1
76	Tor_TBR3						
77	Tor_TBR5		–	–			–
78	Tor_TBR7		6	6	–		G1
79	Tor_TBR8				1		
80	Tor_TBR10				–		
81	Tor_TBY2				5		
82	Tor_TBY4						
83	Tor_TBY5						
84	Tor_TBY6		–	–			–
85	Tor_TBY10		2	2			G1
86	Tor_TBY20						
87	Tor_YLB1		–	–	18		–
88	Tor_YLB2						
89	Tor_YLF1						
90	Tor_YLF2						

91	Tor_YLF4						
92	Tor_YLF7						
93	Tor_YLS1						
94	Tor_YLS3						
95	Tor_YLS5						
96	Tor_YLS6						
97	Tor_YLS7						
98	Tor_YLS9						
99	Tor_YLS10						
100	Tor_YLS11						
101	Tor_YLS12						
102	Tor_YLS13						
103	Tor_YLS14						
104	Tor_YLS16						
		Total	57	57	102	104	

DNA Extraction

DNA was extracted using the Gentra Puregene Tissue Kit (Qiagen Sciences, Maryland, USA) according to the manufacturer's protocol. Approximately 15 mg of muscle tissue was excised from the preserved sample and thoroughly washed with deionized water. It was then dried on Kim Wipes (Kimberly-Clark Co, USA). The cleaned tissue was further minced and placed in a 1.5 mL tube containing 600 µL of cell lysis solution (Qiagen Sciences) and 60 µg/ml Proteinase K. The mixture was incubated at 55°C for 3 hours to facilitate tissue lysis. 12 µg/ml RNase A was added, and the sample was incubated at 37°C for 30 minutes to degrade RNA. After incubation, the sample was allowed to cool at room temperature for 5 minutes. 200 µL of protein precipitation solution (Qiagen Sciences) was added, and the tube was cooled on ice for 15 minutes to precipitate proteins. The sample was then centrifuged at 7500 x g at 4°C for 15 minutes, and the supernatant was transferred to a new tube. 600 µL of isopropanol was added to precipitate DNA, and the sample was incubated at –20°C for 2 hours. After incubation, the sample was centrifuged at 7500 x g at 4°C for 15 minutes. The supernatant was discarded, and the DNA pellet was washed with 600 µL of 70% ethanol. The tube was centrifuged at 7500 x g at 4°C for 5 minutes, and the ethanol was removed. The DNA pellet was air-dried for 20 minutes, resuspended in 25 µL of deionized water, and stored at –20°C for future use.

Primer design and PCR in silico

The D-loop primers were designed using the complete mitochondrial genome sequence of *T. tambra* obtained from GenBank (GenBank: MW471078.1 and OK442459) as a reference (Table

S2). To encompass the entire region of D-loop, the forward primer was designed to target the control region of the mitochondrial sequence near tRNA-Pro, and the reverse primer was designed to target tRNA-Phe. In silico PCR was performed by SnapGene version 8.2 (Dotmatics Ltd., Boston, MA, USA; Dotmatics 2025).

mtDNA Markers Amplification

The fragments of *COX1* and 16S rRNA gene, and D-loop were amplified by PCR. Each 25 µL reaction contained 5× GoTaq® Flexi Buffer (1× final concentration), 0.2 mM dNTPs, 2.5 mM $MgCl_2$, 3 U GoTaq® Flexi DNA Polymerase (Promega, Madison, WI, USA), 0.2 µM of each primer (Macrogen, Inc., Korea), and 20–200 ng genomic DNA template. Primer details are provided in Table S2. PCR conditions consisted of an initial denaturation at 94°C for 4 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing at the temperatures and times as specified in Table S2, and 72°C for 30 s, with a final extension at 72°C for 5 min and a hold at 4°C. Amplicons were verified on a 1.5% agarose gel and sequenced by Macrogen (Seoul, South Korea). The forward and reverse sequences were examined manually in the chromatograms to check the sequence quality, and low-quality nucleotides were trimmed and assembled using MEGA11 (Tamura et al. 2021).

Dataset

The datasets used in this study consisted of both sequences generated and homologous *Tor* sequences retrieved from the NCBI GenBank (Tables S3-S4). The *COX1* dataset comprises 77 sequences among which 57 are newly generated in this study (Table 1). These sequences represent the following nominal *Tor* species: *T. tambra*, *T. tambroides*, *T. dongnaiensis*, *T. douronensis*, *T. mekongensis*, plus two outgroup samples from *Labeo gonius*. The 16S rRNA dataset contains 90 sequences, including 57 sequences newly generated in this study and 33 sequences from GenBank. The sequences are from the following nominal species of *Tor*: *T. tambra*, *T. tambroides*, *T. douronensis*, plus two outgroup samples from *Labeo gonius*. The D-loop dataset comprised 102 sequences of "*T. tambra"*, "*T. tambroides",* all of which were generated in this study. All sequences newly generated in this study were deposited in the NCBI GenBank database, and their accession numbers are provided in table S5.

Phylogenetic tree reconstruction, species delimitation analyses, and haplotype network generation

Two mitochondrial markers, *COX1* and 16S rRNA, were used to reconstruct phylogenetic relationships among *Tor* species because these loci are widely applied in fish phylogenetics and provide reliable resolution for species-level relationships. The mitochondrial control region (D-loop), which evolves more rapidly, was used to construct haplotype networks to examine genetic diversity and population-level structure among the sampled individuals. Sequences generated in this study were combined with reference sequences of *Tor* species retrieved from GenBank to provide broader taxonomic representation in the phylogenetic analyses.

The compiled DNA sequences of the *COX1* gene,16S rRNA, and D-loop were managed, edited, and trimmed using MEGA 11 package (Tamura et al. 2021). The sequence alignments were performed using automatic sequence alignment program MUSCLE (Edgar 2004) as implemented in MEGA 11. The maximum likelihood (ML) method was employed to construct phylogenetic trees of *COX1* and 16S rRNA genes, employing the GTR+GAMMA model of nucleotide substitution. Node support was assessed using 1,000 rapid bootstrap replicates in raxmlGUI 2.0 using RAxML (Edler et al. 2021). Bayesian phylogenetic inference was performed using MrBayes v3.2.7 (Ronquist et al. 2019) with an MCMC chain of 10,000,000 generations and a sampling frequency of 1,000 in order to get 10,000 sampling trees. We discarded the first 20% of the sampling trees as burn-in. The IQ-TREE (Trifinopoulos et al. 2016) web server was used to determine the best-fit model for the Bayesian phylogenetic trees. In the Bayesian Inference (BI) analysis, the GTR + I + gamma model was applied to the *COX1* dataset, and the K2P + I model was applied to the 16S rRNA dataset. The species delimitation tools used in this study were Assemble Species by Automatic Partitioning (ASAP) with pairwise genetic distances calculated (Puillandre et al. 2021), including the Bayesian Poisson Tree Process (bPTP) based on the Bayesian method of phylogenetic tree to separate species in bPTP web server (https://species.h-its.org) (Zhang et al. 2013) and automatic barcode gap discovery (ABGD) with a method to locate the barcode-gap distance (Puillandre et al. 2012) employing Kimura's two-parameter (K2P) model (Kimura 1980) to identify OTUs via the open-source iTaxoTools program (https://itaxotools.org/) (Vences et al. 2021).

To assess population structuring and visualize genealogical relationships among haplotypes, a haplotype network based on D-loop sequences was constructed for *Tor* spp., comprising *T. tambra* and its closely related lineages from Thailand and Peninsular Malaysia (based on the reconstructed *COX1* and 16S rRNA phylogenetic results), using the median-joining algorithm in PopART (University of Otago, http://popart.otago.ac.nz). Haplotypes were first identified using DnaSP6 (Rozas et al. 2017).

Estimates of genetic diversity and population differentiation were performed using the D-loop dataset. Nucleotide diversity (π) and haplotype diversity (h) were calculated for each sampled locality using all 102 D-loop sequences in Arlequin v.3.5 (Excoffier and Lischer 2010). Pairwise ΦST comparisons were then performed among populations within the broader G1/G2 mitochondrial assemblage after excluding Tor_PHG003 and Tor_HUL005, which belonged to the taxonomically unresolved G3-1 lineage. Pairwise ΦST values were calculated based on p-distance, with statistical significance assessed using 100,000 permutations.

RESULTS

Morphology examination, Phylogenetic inference, species delimitation analyses and genetic divergence

Morphology examination is only able to identify specimen up to genus level while meristic variation among the examined *Tor* specimens was limited and largely overlapping across sampling localities (Fig. S1, Table S1). Predorsal scales ranged from 8 to 10, TRa was usually 4 with occasional counts of 5, TRb was constant at 3 in all specimens. Gill-raker counts mostly ranged from 14 to 24 among the examined specimens. Although two specimens showed low counts of 14 and 15, these values were close to several other low-end counts of 16–17 and were therefore interpreted as part of the observed meristic variation. In contrast, one specimen, Tor_N2 from Narathiwat, showed a higher count of 28, which was separated from the remaining specimens. Because this high value was observed in only one individual, it was treated cautiously and was not used as a diagnostic character for species-level identification. Overall, these characters showed substantial overlap among populations and did not provide clear diagnostic separation among the sampled groups.

The length of the aligned sequences is 555 bp for the *COX1* dataset and 530 bp for the 16S rRNA dataset. Excluding the outgroup sequences, the *COX1* and the 16S rRNA datasets contained 35 and 31 parsimony-informative sites, respectively. The number of variable sites was 39 (7%) for the *COX1* gene and 44 (8.3%) for the 16S rRNA gene. Figure 2 illustrates the *COX1* gene tree of *Tor* spp. reconstructed using the ML method, together with the results of the species delimitation analyses using three different approaches. Both the phylogenetic analysis and ABGD analysis resolved the sampled individuals into four distinct clades or operational taxonomic units (OTUs), whereas ASAP and bPTP grouped them into three and five OTUs, respectively. Bayesian inference (BI) supports the ML topology, with nodal support values based on posterior probabilities (PP)

consistent with the ML bootstrap support values (BS). The values of MLBS greater than 70% and BI PP greater than 0.94 are considered to represent strong confidence in the inferred phylogenetic tree.

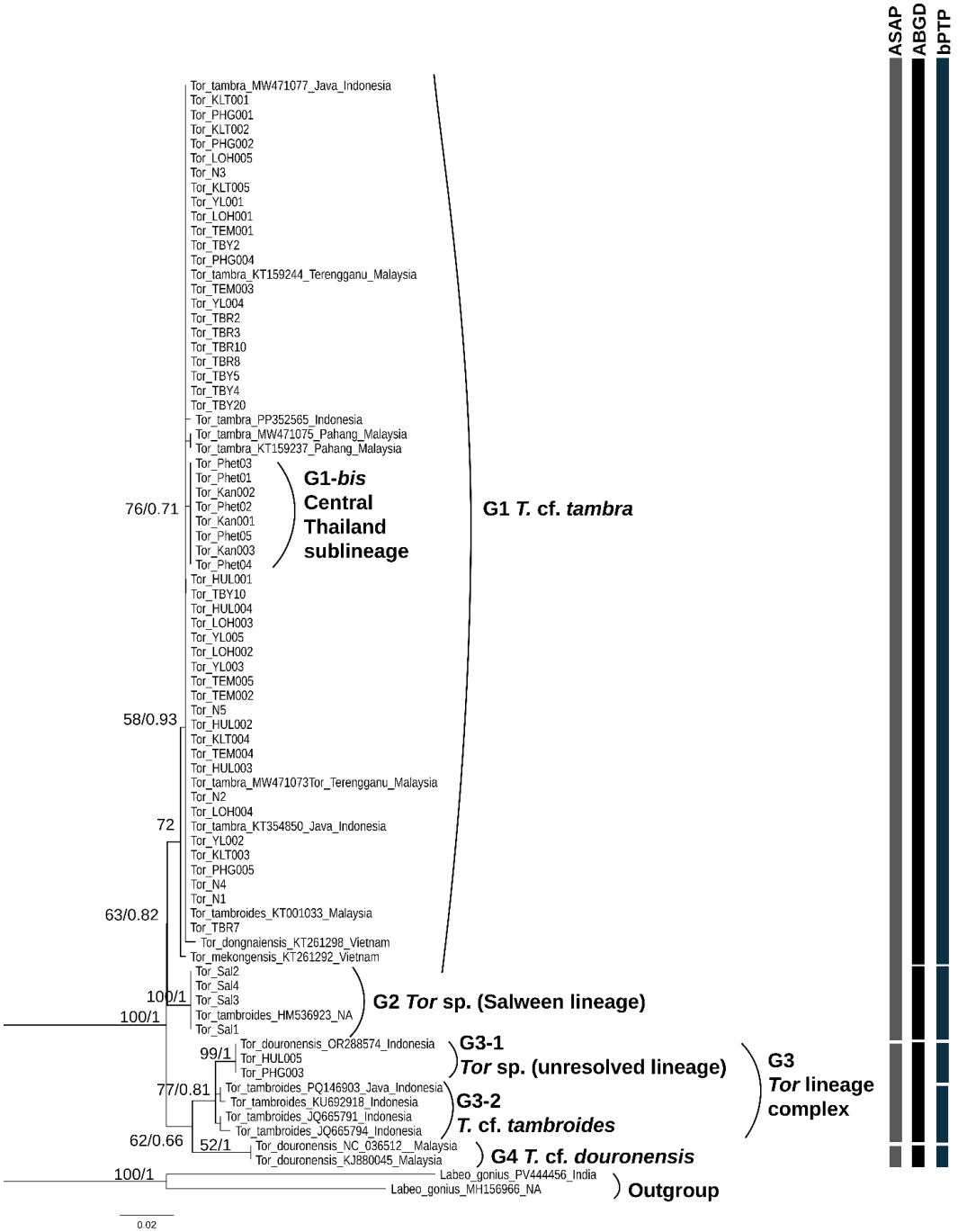

Fig. 2. Maximum-likelihood phylogenetic tree of *Tor* species based on *COX1* sequences, with *Labeo gonius* used as the outgroup. Results of species-delimitation analyses (ASAP, ABGD, and bPTP) are presented as vertical bars on the right. Numbers at nodes represent ML bootstrap values (BS) and Bayesian inference posterior probabilities (PP). Taxon information is provided in tables 1 and S3. Taxa labelled with accession numbers represent sequences retrieved from GenBank.

Further inspection revealed that Group (G) 1 comprised sequences of *Tor* samples collected from southern Thailand, central Thailand, and Peninsular Malaysia, together with three *T. tambra*

sequences from Java, Indonesia; one sequence identified as "*T. dongnaiensis*" from Vietnam, one identified as "*T. mekongensis*" from Vietnam; and one identified as "*T. tambroides*" from Malaysia. Notably, the GenBank record KT001033, originally identified as *T. tambroides*, clustered within G1 together with *T. dongnaiensis* KT261298 and *T. mekongensis* KT261292. These three records were clearly separated from G3-2, provisionally designated as the *T.* cf. *tambroides* mitochondrial lineage based on the included Sumatran GenBank sequences. Within the G1, samples from central Thailand (Tor_Phet001–005 from Phetchaburi and Tor_Kan001–003 from Kanchanaburi) form a median-supported subclade (G1-*bis*). In contrast, all the samples from northern Thailand, collected from the Salween River basin, along with one GenBank sequence, identified as *T. tambroides* (accession no. HM536923; locality unknown), form a strongly supported clade, G2. G2 was recovered as sister to G1, although this relationship lacked statistical support and was not recovered by ASAP. A third clade, G3, contains two samples from Malaysia (Tor_PHG003 [locality: Endau, Pahang] and Tor_HUL005 [locality: Hulu Langat, Selangor]), one GenBank sequence identified as "*T. douronensis*" from Bogor, Indonesia, and four *T. tambroides* from Sumatra, Indonesia. The bPTP analysis further partitioned G3 into two OTUs: G3-1, comprising the two Malaysian samples and the Javan *T. douronensis* sequence, and G3-2, comprising the four Sumatran *T. tambroides* sequences. Finally, two GenBank sequences of *T. douronensis* from Malaysia form a separate clade (G4) in the *COX1* tree.

The reconstructed ML 16S rRNA gene tree is shown in figure 3, with most nodes lacking support. Nevertheless, the clustering patterns are similar to those observed in the *COX1*, with most of the *T.* cf. *tambra* and "*T.* cf. *tambroides"* samples from Thailand and Malaysia collected in this study and available in GenBank clustered together into a main G1 group, albeit without statistical support. Sequences of *Tor* from the Salween River in northern Thailand and "*T. tambroides*" sequences retrieved from GenBank (HM536781, EF588066, and AP011372) form a group, corresponding to *COX1* G2. Two Malaysian samples (Tor_PHG003 and Tor_HUL005) remain distinct from the main group and cluster with several "*T. douronensis*" sequences from GenBank (localities: China and Sarawak) as well as one "*T. tambroides*" from Vietnam. Although some topological differences exist between the *COX1* and 16S rRNA trees—likely due to variation in gene resolution and taxon sampling—both markers reveal broadly similar mitochondrial lineage structuring. However, the low nodal support and limited resolution of the 16S rRNA marker prevented confident discrimination among the nominal species. The incongruence between some database identifications and their phylogenetic placements may reflect inaccurate or uncertain GenBank annotations, insufficient phylogenetic resolution of the 16S rRNA gene, or unresolved taxonomy within the *Tor* complex.

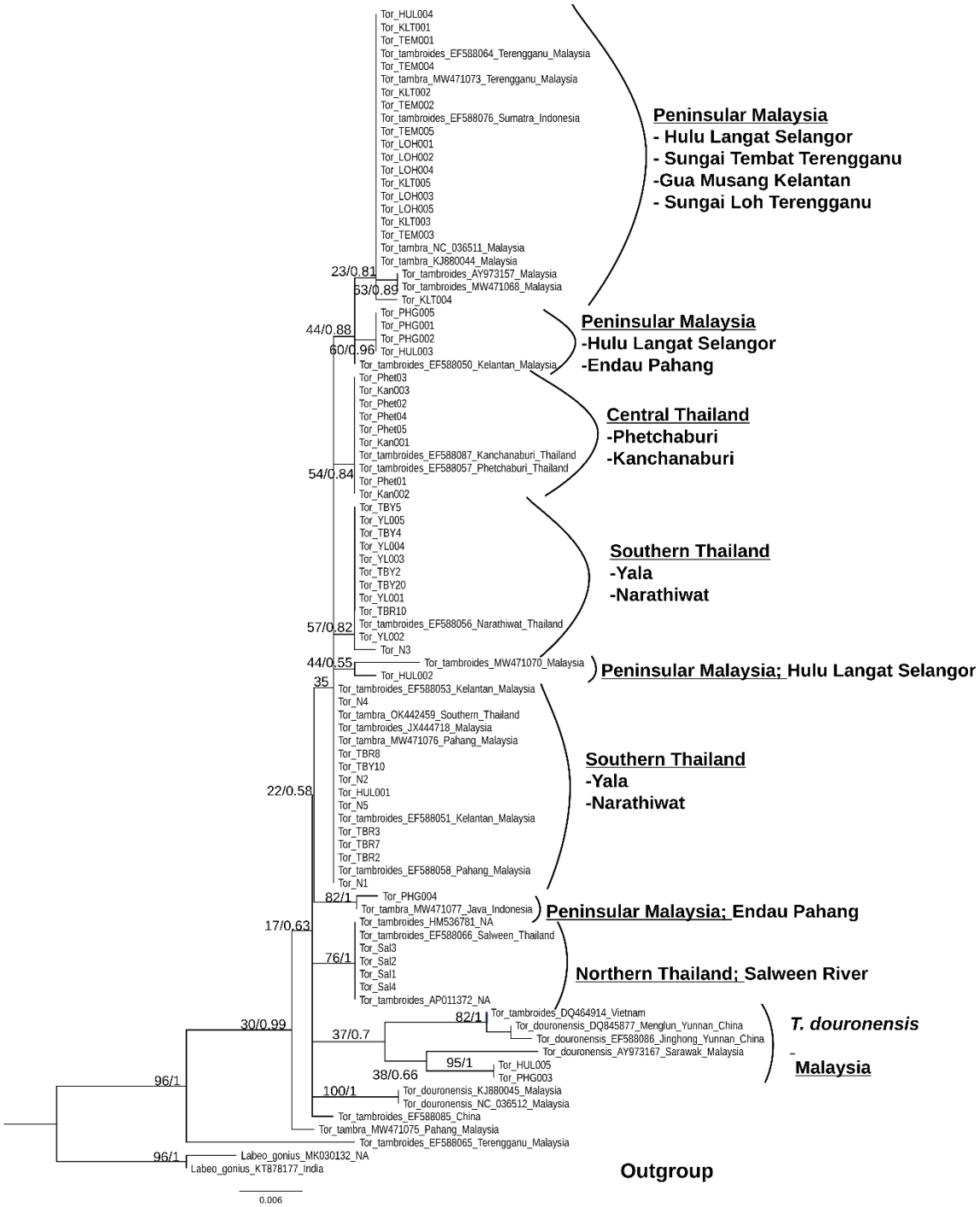

Fig. 3. Maximum-likelihood phylogenetic tree of *Tor* species based on 16S rRNA sequences, with *Labeo gonius* as the outgroup. Numbers at nodes represent ML bootstrap values (BS) and Bayesian inference posterior probabilities (PP). Taxon information is provided in tables 1 and S4. Taxa labelled with accession numbers represent sequences retrieved from GenBank.

To further evaluate mitochondrial lineage divergence and explore potential cryptic diversity, average pairwise p-distances based on *COX1* sequences were calculated among the delimited groups. As shown in table 2, G1 was differentiated from G3-1, G3-2, and G4 by average *p*-distances of 3.6%, 3.0%, and 3.8%, respectively. These values are comparable to the approximately 3%

COX1 divergence sometimes used as a preliminary benchmark for identifying deeply differentiated fish lineages (Ward et al. 2009), although such thresholds are taxon-dependent and should not be interpreted as definitive evidence of species boundaries. G1-bis showed minimal divergence from G1 (0.3%), supporting its placement within G1. The comparatively low divergences between G1 and G2 (1.7%) and between G3-1 and G3-2 (1.4%) indicate close mitochondrial relationships but do not, by themselves, establish conspecificity. Additional evidence, particularly from nuclear genomic data, reproductive biology, and voucher-based morphology, will be required to confirm their taxonomic relationships. Based on the barcode-gap pattern, we tentatively recognize three major mitochondrial lineage clusters: G1/G2, G3, and G4.

Table 2. Genetic divergences as estimated by average *p*-distance based on *COX1* sequences among identified *Tor* groups from phylogenetic analysis and species-delimitation results. G1 represents the major *Tor* clade; G1-*bis* includes samples from Central Thailand (Phetchaburi and Kanchanaburi); G2 corresponds to the Salween group from Northern Thailand; G3-1 includes two samples (Tor_PHG003 and Tor_HUL005) from Malaysia; G3-2 includes *T.* cf. *tambroides* (GenBank sequences); and G4 represents *T.* cf. *douronensis* (GenBank sequences)

	G1	G1-bis	G2	G3-1	G3-2	G4
G1	–					
G1-bis	0.003	–				
G2	0.017	0.019	–			
G3-1	0.036	0.033	0.038	–		
G3-2	0.030	0.028	0.033	0.014	–	
G4	0.038	0.040	0.040	0.040	0.035	–

The haplotype network, nucleotide and haplotype diversities, and population differentiation

Figure 4 presents a haplotype network constructed based on D-loop sequences of samples assigned to G1, G2, and G3-1. The main haplotype groups recovered here correspond closely to those revealed by the *COX1* phylogeny. The main group (G1) consists of haplotypes from the Central–Southern Thailand and Peninsular Malaysia, with one internal haplogroup from central Thailand individuals, separated from the rest by seven mutation steps. Haplotypes from the Salween River, northern Thailand, formed a second haplogroup (equivalent to G2 in *COX1* tree), which is separated from the G1 by 17 mutation steps. Meanwhile, the PHG003 and HUL005 samples from Malaysia (Hap_1 and Hap_2), which clustered with the Indonesian *T. tambroides* (GenBank sequences) in the *COX1* tree (G3), formed a distinct haplogroup separated from main haplogroup by 33 mutation steps.

For genetic diversities, most populations exhibited low nucleotide and haplotype diversities (π = 0.00–0.01; h =0.00–0.46), except those from Hulu Langat, Selangor and Endau Pahang in Malaysia, which showed markedly high nucleotide and haplotype diversities (π = 0.0294–0.0313; h

= 0.8444–0.9167). This pattern suggests a potential admixture of individuals originating from different sources (lineages) within the latter two aquaculture ponds. In contrast, the populations Salween (northern Thailand) (π = 0.0194; h = 0.8333) and Gua Musang Kelantan (π = 0.0068; h = 0.6889) populations displayed low nucleotide but high haplotype diversities, suggesting a natural genetic structure with historical bottleneck/expansion or broad maternal lineage representation (Table 3).

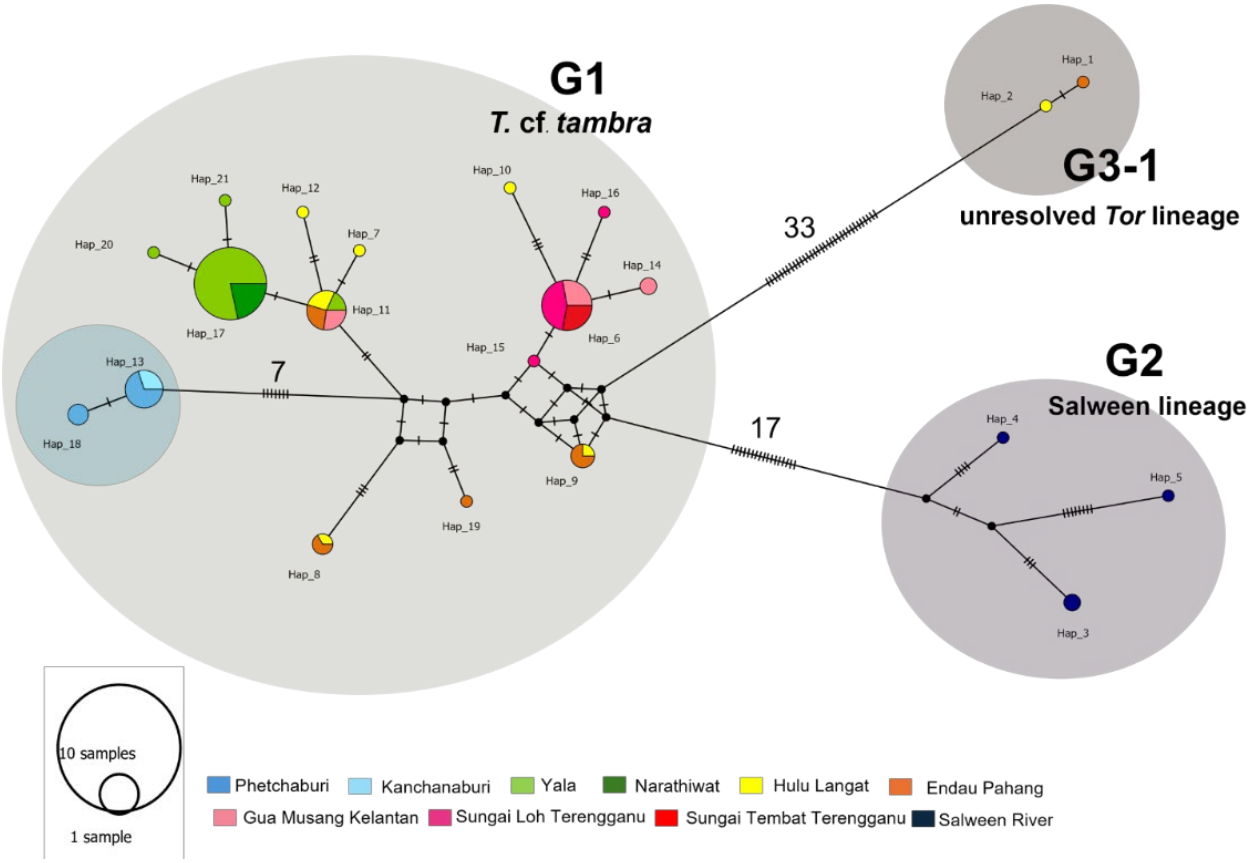

Fig. 4. Median-joining haplotype network of 21 haplotypes generated from 102 newly obtained *Tor* D-loop sequences. The number of mutational steps is indicated by crossbars and numbers on the connecting lines. Circle sizes are proportional to haplotype frequency, and small black dots represent inferred missing haplotypes. Sample and taxon information is provided in table 1. Group G3-2 and G4 are not represented because they comprised only GenBank *COX1* reference sequences for which corresponding D-loop sequences were unavailable.

Table 3. Summary of haplotype and nucleotide diversity based on D-loop sequences of *Tor* spp. across sampling localities in Thailand and Peninsular Malaysia. The number of samples is indicated by *n*. An asterisk denotes a wild population

Locality	n	Haplotypes	Haplotype Designation (Individual Number)	Nucleotide diversity (π)	Haplotype diversity (h)
Salween River, Thailand*	4	3	Hap_3(2), Hap_4(1), Hap_5(1)	0.0194 ± 0.0135	0.8333 ± 0.2224
Kanchanaburi, Thailand*	3	1	Hap_13(3)	0.0000 ± 0.0000	0.0000 ± 0.0000
Phetchaburi, Thailand	10	2	Hap_13(7), Hap_18(3)	0.0010 ± 0.0010	0.4667 ± 0.1318
Yala, Thailand	33	4	Hap_11(2), Hap_17(29), Hap_20(1), Hap_21(1)	0.0006 ± 0.0008	0.2292 ± 0.0950
Narathiwat, Thailand*	8	1	Hap_17(8)	0.0000 ± 0.0000	0.0000 ± 0.0000
Sungai Tembat, Malaysia	5	1	Hap_6(5)	0.0000 ± 0.0000	0.0000 ± 0.0000
Gua Musang, Malaysia	10	3	Hap_6(5), Hap_11(3), Hap_14(2)	0.0068 ± 0.0043	0.6889 ± 0.1038
Sungai Loh, Malaysia	10	3	Hap_6(8), Hap_15(1), Hap_16(1)	0.0017 ± 0.0015	0.3778 ± 0.1813
Hulu Langat, Selangor, Malaysia	9	7	Hap_2(1), Hap_7(1), Hap_8(1), Hap_9(1), Hap_10(1), Hap_11(3), Hap_12(1)	0.0313 ± 0.0176	0.9167 ± 0.0920
Endau, Malaysia	10	5	Hap_1(1), Hap_8(2), Hap_9(3), Hap_11(3), Hap_19(1)	0.0294 ± 0.0164	0.8444 ± 0.0796

Table 4 presents the pairwise Φ_{ST} comparisons among populations based on mitochondrial D-loop sequences, excluding two samples (Tor_PHG003 and Tor_HUL005) because they belonged to the taxonomically unresolved G3-1 mitochondrial lineage rather than the principal G1 assemblage (see discussion). The pairwise Φ_{ST} comparisons revealed low differentiation between populations within the same geographic regions but strong differentiation among regions (Table 4). The two central Thailand populations (Kanchanaburi and Phetchaburi) showed very low differentiation (Φ_{ST} = 0.0187), while the southern Thailand populations (Yala and Narathiwat) showed no differentiation (Φ_{ST} = 0), indicating that each region represents a cohesive maternal population. However, comparisons between central Thailand and southern Thailand yielded extremely high Φ_{ST} values (0.9673–1.0000), indicating strong genetic differentiation and limited maternal gene flow between these regions. In addition, populations from Salween also produced high Φ_{ST} values (Φ_{ST} = 0.8409 – 0.9565) when compared with other populations, indicating pronounced genetic isolation and strong population structuring. Overall, Thai populations were clearly separated into three genetic lineages: (i) the Salween river, (ii) central Thailand, and (iii) southern Thailand.

Table 4. The pairwise Φ_{ST} comparison among populations within the broader G1/G2 mitochondrial assemblage based on D-loop sequences. An asterisk denotes wild populations

Population	Salween*	Kanchanaburi*	Phetchaburi	Yala	Narathiwat
Salween*	–				
Kanchanaburi*	0.8409	–			
Phetchaburi	0.9197	0.0187	–		
Yala	0.9565	0.9722	0.9673	–	
Narathiwat	0.8934	1	0.9746	0	–
Sungai Tembat Terengganu	0.8603	1	0.9688	0.9626	1
Gua Musang Kelantan	0.8294	0.7542	0.8224	0.8042	0.6805
Sungai Loh Terengganu	0.8925	0.9374	0.9403	0.9408	0.9361
Hulu Langat Selangor	0.7693	0.6294	0.7502	0.5453	0.3304
Endau Pahang	0.7692	0.6219	0.7382	0.6557	0.4526

Population	Sungai Tembat Terengganu	Gua Musang Kelantan	Sungai Loh Terengganu	Hulu Langat Selangor	Endau Pahang
Salween*					
Kanchanaburi*					
Phetchaburi					
Yala					
Narathiwat					
Sungai Tembat Terengganu	–				
Gua Musang Kelantan	0.1035	–			
Sungai Loh Terengganu	0	0.1326	–		
Hulu Langat Selangor	0.4764	0.2067	0.5245	–	
Endau Pahang	0.4496	0.2356	0.4996	0	–

Within Peninsular Malaysia, regional substructure was also observed. Sungai Tembat and Sungai Loh both from Terengganu showed no differentiation ($\Phi_{ST} = 0$), forming a distinct east-coast cluster, while Hulu Langat and Endau Pahang also showed no differentiation ($\Phi_{ST} = 0$), representing a separate west–south cluster. Gua Musang Kelantan exhibited moderate differentiation from the Terengganu populations ($\Phi_{ST} = 0.1035–0.1326$) and greater differentiation from Hulu Langat and Endau ($\Phi_{ST} = 0.2067–0.2356$), suggesting an intermediate but distinct position among Malaysian populations.

Collectively, these results demonstrate a well-defined geographic structure across *Tor* populations in Thailand and Peninsular Malaysia, pointing to restricted gene flow, regional isolation, and potentially long-term evolutionary divergence among river systems. In contrast, the high genetic diversity observed in Malaysia's *Tor* populations might have different origins, most likely reflecting admixture resulting from the translocation of mahseer among aquaculture facilities.

DISCUSSION

This study provides valuable insights into the genetic identity and population structure of *Tor* fishes in Thailand and Peninsular Malaysia, contributing to the ongoing clarification of

taxonomic ambiguities between *T. tambra* and *T. tambroides*. Species identification within the genus *Tor* has long been problematic due to overlapping morphological characters and inconsistent diagnostic traits reported in previous taxonomic studies. Traditional morphological traits such as snout shape, fin coloration, and the lower median lobe (Jaafar et al. 2021) were unreliable in our juvenile specimens, highlighting the importance of molecular approaches for accurate species identification in both aquaculture and conservation management. The meristic characters such as gill raker counts, vertebral counts, and scale patterns have been used to distinguish *Tor* species; however, several studies have reported substantial overlap in these characters between *T. tambra* and *T. tambroides*. For example, Roberts (1999) noted that key diagnostic characters, including gill raker and vertebral counts, often overlap between these species, making reliable identification based solely on morphology difficult. In addition, morphological variation associated with ontogenetic stages and environmental conditions may further complicate species identification in mahseer. These challenges have contributed to the long-standing taxonomic uncertainty within the genus *Tor* and highlight the importance of integrating molecular approaches with traditional morphological observations when investigating species boundaries. Morphological assessment in the present study was further limited by the predominance of juvenile specimens, in which diagnostic external characters are often not yet fully developed or are difficult to assess consistently. Although adults may exhibit clearer interspecific differences, previous studies have reported overlap in several key meristic traits between *T. tambra* and *T. tambroides*. Therefore, morphology alone was insufficient to resolve species boundaries in our dataset, underscoring the value of molecular data for provisional lineage assignment and the need for future studies integrating adult voucher specimens with nuclear genomic evidence.

Genetic analyses conducted in this study revealed that the majority of samples from Thailand and Peninsular Malaysia clustered within a major lineage together with reference sequences identified as *T. tambra* from Java and Malaysia in GenBank. This dominant clade, designated G1, included multiple sequences identified as *T. tambra*, including the Javan record KT354850, and is therefore provisionally interpreted as a *T.* cf. *tambra* mitochondrial lineage. However, this interpretation remains provisional because the available GenBank sequences cannot be regarded as confirmed topotypic reference specimens. Notably, the GenBank record KT001033, originally identified as *T. tambroides*, was also placed within G1 together with sequences assigned to *T. dongnaiensis* (KT261298) and *T. mekongensis* (KT261292). Their placement within the principal *T.* cf. *tambra* lineage was incongruent with their nominal species assignments, and these records were clearly separated from the Sumatran *T.* cf. *tambroides* lineage recovered in G3-2. This pattern is congruent with Phanklam et al. (2026), who also highlighted taxonomic ambiguity among mitochondrial sequences assigned to *T.* cf. *tambra*, *T.* cf. *tambroides*, and related nominal species.

Accordingly, these GenBank records were treated as taxonomically uncertain and were not considered definitive species references in species-level interpretations. Nevertheless, they were retained in the phylogenetic tree under their original database annotations to illustrate inconsistencies within publicly available sequence records.

G3 was subdivided into G3-1 and G3-2. G3-2 contained GenBank sequences identified as *T. tambroides* from Sumatra, including JQ665791 and JQ665794, and is therefore provisionally designated as the *T.* cf. *tambroides* mitochondrial lineage. In contrast, G3-1 comprised Tor_PHG003 and Tor_HUL005 from Peninsular Malaysia together with a GenBank sequence identified as *T. douronensis* from Bogor, Java. G3-1 showed mitochondrial affinity to G3-2 but remained taxonomically unresolved. The average *COX1* p-distance between G3-1 and G3-2 was 1.4%, indicating that they are closely related but genetically differentiated mitochondrial sublineages. Nevertheless, the available *COX1* data do not permit the definitive assignment of Tor_PHG003 and Tor_HUL005 to *T.* cf. *tambroides*. Voucher-based morphological comparisons and nuclear genomic data will be required to resolve the taxonomic identity of G3-1. The Salween population from northern Thailand formed a differentiated mitochondrial group that may represent a cryptic lineage although its *COX1* divergence from G1 remains relatively limited (1.7%). Similarly, the central Thailand populations (Phetchaburi and Kanchanaburi) emerged as a subgroup within the broader G1 assemblage.

These findings suggest that the names *T. tambra* and *T. tambroides* may have been applied inconsistently to populations from northern Thailand to Peninsular Malaysia. Most specimens examined in the present study clustered within the G1 *T.* cf. *tambra* mitochondrial lineage, despite often being identified as *T. tambroides* by farmers and collectors. This pattern is broadly consistent with Chew et al. (2021), who regarded most *Tor* stocks sampled across Malaysia as *T. tambra*, whereas a genetically differentiated Hulu Langat stock reportedly imported from Sumatra was considered most likely *T. tambroides*. Collectively, these findings indicate that the name *T. tambroides* may have been applied too broadly to Malaysian mahseer stocks. However, mitochondrial evidence alone cannot confirm the extent of historical misidentification or establish the precise geographical distributions of these species. The fine-scale mitochondrial structuring observed within the broader *T.* cf. *tambra* assemblage may reflect historical isolation and restricted maternal gene flow among drainage systems. In particular, the Isthmus of Kra, which is widely recognized as an important biogeographic transition for freshwater and terrestrial taxa in Southeast Asia (Dejtaradol et al. 2016) (Fig. 1), may have contributed to lineage differentiation by restricting historical connectivity among river basins across the peninsula. Nevertheless, this interpretation remains tentative because several populations were obtained from aquaculture facilities, where broodstock translocation and mixing may obscure natural phylogeographic patterns. The observed

genetic structuring emphasizes the importance of region-specific broodstock management and careful documentation of stock origins. Avoiding indiscriminate mixing or translocation of genetically differentiated stocks may help maintain regional genetic diversity and support sustainable aquaculture and conservation programs.

The divergent samples Tor_HUL005 from Hulu Langat, Selangor, and Tor_PHG003 from Endau, Pahang, formed the G3-1 mitochondrial lineage, which was closely related to, but differentiated from, the Sumatran *T.* cf. *tambroides* G3-2 lineage. This phylogenetic affinity suggests that these specimens may possess mitochondrial affinity to the *T.* cf. *tambroides* lineage; however, the available evidence does not permit their definitive assignment to *T. tambroides*. Several explanations may account for this pattern, including distinct species membership, mitochondrial introgression following historical hybridization, incomplete lineage sorting, or anthropogenic translocation of broodstock. Because all markers used in the present study are maternally inherited, they cannot distinguish among these alternative scenarios or determine whether the mitochondrial placement reflects the overall genomic ancestry of these individuals. Resolving the taxonomic identity of G3-1 will therefore require evidence from independent nuclear markers, such as genome-wide SNPs or whole-genome sequencing, together with voucher-based morphological examination. From an applied perspective, the occurrence of genetically differentiated mitochondrial lineages in aquaculture stocks emphasizes the importance of traceable broodstock records and molecular screening, because inconsistent species identification or unintentional mixing of divergent stocks may affect breeding management, genetic integrity, and market confidence.

CONCLUSIONS

In conclusion, this study shows that combining multiple mitochondrial markers is effective for clarifying genetic structure and informing species-boundary hypotheses within *Tor* populations from Thailand and Peninsular Malaysia. The *COX1* data supported lineage-level divergence, whereas D-loop sequences provided finer resolution of population structure, revealing differentiation between central and southern Thai populations and additional subdivision among Peninsular Malaysian localities. Together, these results highlight the presence of geographically structured mitochondrial diversity within the *T.* cf. *tambra* assemblage and provide new insight into the taxonomic complexity of mahseer in the region.

Beyond taxonomic clarification, these findings provide a useful genetic baseline for conservation management, broodstock selection, and future aquaculture development of this

culturally and economically important fish group. Rather than resolving all taxonomic uncertainty, the present results narrow the range of plausible interpretations and identify priority questions for subsequent work. Integrating nuclear markers, broader geographic sampling, and detailed morphological analyses will be important for testing lineage boundaries more rigorously.

Our results also suggest that the currently assumed distribution of *T.* cf. *tambroides* in parts of Thailand and Peninsular Malaysia should be interpreted cautiously. Given the close relationship among some populations and the long-standing taxonomic confusion within *Tor*, records attributed to *T. tambroides* may require re-evaluation using integrative evidence from mitochondrial and nuclear datasets together with morphology. Such an approach will be essential for improving taxonomic accuracy, strengthening regional conservation strategies, and supporting reliable species identification in fisheries and aquaculture.

Acknowledgments: This research was supported by the National Science, Research, and Innovation Fund (NSRF) and Prince of Songkla University (Ref. No. SCI6701294S) of Thailand. We thank all members of Wei-Jen Chen's laboratory, Institute of Oceanography, National Taiwan University for their suggestions and discussions on phylogenetic analyses knowledge. Dr. Meng-Shin Shiao of Mahidol University kindly advised on primer design for D-loop amplification, and Professor Alex Hon-Tsen Yu provided substantial assistance.

Authors' contributions: Krittaporn Lempan: Writing, Investigation, Methodology, Visualization, Data curation. Hong-Chiun Lim: Writing, Validation, Data curation, Formal analysis, Visualization, review & editing. Wei-Jen Chen: Writing, Validation, review & editing. Monwadee Wonglapsuwan: Conceptualization, Resources, Data curation, Formal analysis, Investigation, Validation, Writing, Visualization, review & editing.

Competing interests: All authors declare that they have no competing interests.

Availability of data and materials: All sequences generated in this study have been deposited in GenBank. Individual accession numbers are provided in table S5.

Consent for publication: All the authors consent to the publication of this manuscript.

Ethics approval consent to participate: All procedures involving fish were conducted in accordance with the Animal Research for Scientific Work guidelines of Prince of Songkla University (PSU Animal Standard), Thailand (approval no. AQ082/2025).

REFERENCES

Al-Khafaji FM, Al-Jubouri MO, Al-Faisal AJ, Aljubouri TRS, Al-Shuhaib MBS. 2025. Genetic insights into the endangered yellowfin barbel in Iraqi rivers: a *COX1*—based approach. Conserv Genet Resour **17:**165–174. doi:10.1007/s12686-025-01388-4.

Alyamani N. 2024. Mitochondrial 16S rRNA gene as a molecular marker in the phylogenetic relationships of some Rabbitfishes species (Siganidae: Perciformes). Open Vet J **14:**1936. doi:10.5455/ovj.2024.v14.i8.22.

Ardura A, Linde AR, Moreira JC, Garcia-Vazquez E. 2010. DNA barcoding for conservation and management of Amazonian commercial fish. Biol Conserv **143:**1438–1443. doi:10.1016/j.biocon.2010.03.019.

Bhattacharya M, Sharma AR, Patra BC, Sharma G, Seo E-M et al. 2016. DNA barcoding to fishes: current status and future directions. Mitochondrial DNA. Mitochondrial DNA Part A **27:**2744–2752. doi:10.3109/19401736.2015.1046175.

Bingpeng X, Heshan L, Zhilan Z, Chunguang W, Yanguo W et al. 2018. DNA barcoding for identification of fish species in the Taiwan Strait. PLoS ONE **13:**e0198109. doi:10.1371/journal.pone.0198109.

Biun H, Sade A, Robert R, Rodrigues KF. 2021. Phylogeographic Structure of Freshwater *Tor* sp. in River Basins of Sabah, Malaysia. Fishes **6:**44. doi:10.3390/fishes6040044.

Brown GG, Gadaleta G, Pepe G, Saccone C, Sbisà E. 1986. Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. J Mol Biol **192:**503–511. doi:10.1016/0022-2836(86)90272-X.

Cawthorn D-M, Steinman HA, Witthuhn RC. 2012. Evaluation of the 16S and 12S rRNA genes as universal markers for the identification of commercial fish species in South Africa. Gene **491:**40–48. doi:10.1016/j.gene.2011.09.009.

Chew PC, Christianus A, Zudaidy JM, Ina-Salwany MY, Chong CM et al. 2021. Microsatellite Characterization of Malaysian Mahseer (*Tor* spp.) for Improvement of Broodstock Management and Utilization. Animals **11:**2633. doi:10.3390/ani11092633.

Dejtaradol A, Renner SC, Karapan S, Bates PJJ, Moyle RG et al. 2016. Indochinese-Sundaic faunal transition and phylogeographical divides north of the Isthmus of Kra in Southeast Asian Bulbuls (Aves: Pycnonotidae). J Biogeogr **43:**471–483. doi:10.1111/jbi.12662.

Dotmatics. 2025. SnapGene® software. Available at: https://www.snapgene.com. (Accessed 19 June 2026)

Duangjai E, Punroob J. 2018. Growth performance of Mahseer Fish (*Tor tambroides*) in different types of cultured environments. Eau Heritage Journal **12:**225–235.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:**1792–1797. doi:10.1093/nar/gkh340.

Edler D, Klein J, Antonelli A, Silvestro D. 2021. raxmlGUI 2.0: A graphical interface and toolkit for phylogenetic analyses using RAxML. Methods Ecol Evol **12:**373–377. doi:10.1111/2041-210x.13512.

Esa YB, Siraj SS, Daud SK, Rahim A, Tan SG. 2008. Mitochondrial DNA Diversity of *Tor* tambroides Valenciennes (Cyprinidae) from Five Natural Populations in Malaysia. Zool Stud **47:**360–367.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour **10:**564–567. doi:10.1111/j.1755-0998.2010.02847.x.

Fang D, Luo H, He M, Mao C, Kuang Z et al. 2022. Genetic Diversity and Population Differentiation of Naked Carp (*Gymnocypris przewalskii*) Revealed by Cytochrome Oxidase Subunit I and D-Loop. Front Ecol Evol **10:**827654. doi:10.3389/fevo.2022.827654.

Fricke R, Eschmeyer WN, Fong JD. 2025a. Species by Family/Subfamily. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp. (Accessed 24 Nov 2025)

Fricke R, Eschmeyer WN, Van der Laan R (eds). 2025b. Eschmeyer's Catalog of Fishes: Genera, Species. References. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp. (Accessed 24 Nov 2025)

Haque MA, Rashid J, Mia ML, Rahman MK, Ali MA et al. 2023. The first report on a new *Tor* species, *Tor barakae* Arunkumar & Basudha 2003, from Bangladesh using DNA barcoding technique. Heliyon **9:**e21764. doi:10.1016/j.heliyon.2023.e21764.

Ingram B, Sungan S, Tinggi D, Sim SY, De Silva SS. 2007. Breeding performance of Malaysian mahseer, *Tor tambroides* and *T. douronensis* broodfish in captivity. Aquac Res **38:**809–818. doi:10.1111/j.1365-2109.2007.01716.x.

Jaafar F, Uthairat NN, Prapansak S, Thumronk A, Duong T-Y et al. 2021. A Current Update on the Distribution, Morphological Features, and Genetic Identity of the Southeast Asian Mahseers, *Tor* Species. Biology **10:**286. doi:10.3390/biology10040286.

Karim A, Iqbal A, Akhtar R, Rizwan M, Amar A et al. 2016. Barcoding of fresh water fishes from Pakistan. Mitochondrial DNA Part A **27:**2685–2688. doi:10.3109/19401736.2015.1043544.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through

comparative studies of nucleotide sequences. J Mol Evol **16:**111–120. doi:10.1007/bf01731581.

Kottelat M. 2013. The fishes of the inland waters of Southeast Asia: a catalogue and core bibliography of the fishes known to occur in freshwaters, mangroves and estuaries. Raffles Bull Zool **27:**1–663.

Lim HC, Abidin MZ, Pulungan CP, Bruyn MD, Nor SAM. 2016. DNA Barcoding Reveals High Cryptic Diversity of the Freshwater Halfbeak Genus *Hemirhamphodon* from Sundaland. PLoS ONE **11:**e0163596. doi:10.1371/journal.pone.0163596

Lim HC, Habib A, Chen WJ. 2021. Comparative Phylogeography and Phylogeny of Pennah Croakers (Teleostei: Sciaenidae) in Southeast Asian Waters. Genes **12:**1926. doi:10.3390/genes12121926.

Nguyen TTT, Ingram B, Sungan S, Gooley G, Sim SY et al. 2006. Mitochondrial DNA diversity of broodstock of two indigenous mahseer species, *Tor tambroides* and *T. douronensis* (Cyprinidae) cultured in Sarawak, Malaysia. Aquaculture **253:**259–269. doi:10.1016/j.aquaculture.2005.09.014.

Nguyen TTT, Na-Nakorn U, Sukmanomon S, ZiMing C. 2008. A study on phylogeny and biogeography of mahseer species (Pisces: Cyprinidae) using sequences of three mitochondrial DNA gene regions. Mol Phylogenet Evol **48:**1223–1231. doi:10.1016/j.ympev.2008.01.006.

Palumbi SR. 1996. Nucleic acids II: The polymerase chain reaction. In Molecular Systematics. Edited by Hillis D, Moritz C, Mable B. Massachusetts: Sinauer Associates Inc. 205-247.

Phanklam V, Janekitkarn S, Seesan K, Muangmai N, Zeng C et al. 2026. Phylogeographic Structure Reveals Hidden Diversity Patterns in *Tor tambra* (Cyprinidae) Across Thai River Systems. Animals **16:**517. doi:10.3390/ani16030517.

Pinder AC, Britton JR, Harrison AJ, Nautiyal P, Bower SD et al. 2019. Mahseer (*Tor* spp.) fishes of the world: status, challenges and opportunities for conservation. Rev Fish Biol Fish **29:**417–452. doi:10.1007/s11160-019-09566-y.

Puillandre N, Lambert A, Brouillet S, Achaz G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol Ecol **21:**1864–1877. doi:10.1111/j.1365-294x.2011.05239.x.

Puillandre N, Brouillet S, Achaz G. 2021. ASAP: Assemble Species by Automatic Partitioning. Mol Ecol Resour **21:**609–620. doi:10.1111/1755-0998.13281.

QGIS.org. 2024. QGIS Geographic Information System. QGIS Association. Available at: https://www.qgis.org/. (Accessed 12 Jul. 2025)

Quraishia SF, Panneerchelvam S, Zainuddin Z, Abd Rashid NH. 2015. Molecular Characterization

of Malaysian Marine Fish Species using Partial Sequence of Mitochondrial DNA 12S and 16S rRNA Markers. Sains Malaysiana **44:**1119–1123. doi:10.17576/jsm-2015-4408-06.

Ramezani-Fard E, Kamarudin MS. 2012. Malaysian mahseer: New candidate for Asian aquaculture? Responsible Seafood Advocate. (Accessed 04 Sep. 2025)

Roberts TR. 1999. Fishes of The Cyprinid Genus *Tor* in The Nam Theun Watershed (Mekong Basin) of Laos, with Description of a New Species. Raffles Bull Zool **47:**225–236.

Ronquist F, Huelsenbeck JP, Teslenko M, Nylander J. 2019. MrBayes version 3.2 Manual: Tutorials and Model Summaries. Available at: https://gensoft.pasteur.fr/docs/mrbayes/3.2.7/Manual_MrBayes_v3.2.pdf. (Accessed 22 Oct. 2025)

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P et al. 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Mol Biol Evol **34:**3299–3302. doi:10.1093/molbev/msx248.

Sbisà E, Tanzariello F, Reyes A, Pesole G, Saccone C. 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene **205:**125–140. doi:10.1016/s0378-1119(97)00404-6.

Surachat K, Deachamag P, Wonglapsuwan M. 2022. The first de novo genome assembly and sex marker identification of Pluang Chomphu fish (*Tor tambra*) from Southern Thailand. Comput Struct Biotechnol J **20:**1470–1480. doi:10.1016/j.csbj.2022.03.021.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol **38:**3022–3027. doi:10.1093/molbev/msab120.

Torgunakova OA, Egorova TA, Semenova SK. 2012. Comparative analysis of variability of three mitochondrial genes of cytochrome oxidase complex (*cox1*, *cox2*, and *cox3*) in wild and domestic carp (*Cyprinus carpio* L.). Russ J Genet **48:**1218–1226. doi:10.1134/s1022795412120149.

Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res **44:**W232–W235. doi:10.1093/nar/gkw256.

Vences M, Miralles A, Brouillet S, Ducasse J, Fedosov A et al. 2021. iTaxoTools 0.1: Kickstarting a specimen-based software toolkit for taxonomists. Megataxa **6:**1–16. doi:10.11646/megataxa.6.2.1.

Walton SE, Gan HM, Raghavan R, Pinder AC, Ahmad A. 2017. Disentangling the taxonomy of the mahseers (*Tor* spp.) of Malaysia: an integrated approach using morphology, genetics and historical records. Rev Fish Sci Aquac **25:**171–183. doi:10.1080/23308249.2016.1251391.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish

species. Philos Trans R Soc, B **360:**1847–1857. doi:10.1098/rstb.2005.1716.

Ward RD. 2009. DNA barcode divergence among species and genera of birds and fishes. Mol Ecol Resour **9:**1077–1085. doi:10.1111/j.1755-0998.2009.02541.x.

Zhang J, Kapli P, Pavlidis P, Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics **29:**2869–2876. doi:10.1093/bioinformatics/btt499.

Supplementary materials

Fig. S1. Representative photographs of *Tor* specimens examined in this study. The specimens include juveniles from Kanchanaburi, Thailand (A, Tor_Kan003); Yala, Thailand (B, Tor_YL004; C, Tor_YLS16); Phetchaburi, Thailand (D, Tor_Phet09); the Salween River, Mae Hong Son, Thailand (F, Tor_Sal1); Hulu Langat, Selangor, Malaysia (G, Tor_HUL005); Gua Musang, Kelantan, Malaysia (H, Tor_KLT005); Sungai Tembat, Terengganu, Malaysia (I, Tor_TEM003); Endau, Pahang, Malaysia (J, Tor_PHG003); and Sungai Loh, Terengganu, Malaysia (K, Tor_LOH003), together with an adult specimen from the Sai Buri River, Narathiwat, Thailand (E, Tor_N7). The corresponding meristic measurements are provided in Table S1, except for Tor_Phet09, which was included as a representative photograph but excluded from the meristic analysis because its small body size prevented reliable examination of the gill rakers. (download)

Table S1. The samples were counted as predorsal scales (Pred. S), scales above the lateral line (TRa), scales below the lateral line (TRb), and gill rakers. (download)

Table S2. Primers of genes with annealing temperature and time, product size, and references.

Table S3. Sequences of the *COX1* gene of *Tor* species obtained from the GenBank database. (download)

Table S4. Sequences of the 16S rRNA gene of *Tor* species obtained from the GenBank database. (download)

Table S5. GenBank Number of sequences based on the *COX1*, 16S rRNA gene, and D-loop control region of *Tor* species obtained from this study. (download)