INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
24–0191/2025/34:340305
http://www.fspublishers.org

Full Length Article

Vouchered DNA Barcoding to Support Aquatic Biodiversity Exploration and Management: Case Study on Maros Karst Endemic *Marosatherina ladigesi*

Inayah Yasir[1*†], Andi Aliah Hidayani[2†], Asmi Citra Malina AR Tasakka[2], Nadiarti Nurdin Kadir[2], Jamaluddin Jompa[1], Widyastuti Umar[1], Magdalena Litaay[3], Andi Parenrengi[4], Samliok Ndobe[5], Abdul Gani[6], Amanda Pricella Putri[7], Muhammad Iqram[3], Sharifuddin Bin Andy Omar[2] and Abigail Mary Moore[8†]

[1]*Department of Marine Science, Faculty of Marine Science and Fisheries, Hasanuddin University, South Sulawesi 90245, Indonesia*
[2]*Department of Fisheries, Faculty of Marine Science and Fisheries, Hasanuddin University, South Sulawesi 90245, Indonesia*
[3]*Department of Biology, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Makassar, South Sulawesi 90245, Indonesia*
[4]*National Research and Innovation Agency (BRIN), Indonesia*
[5]*Fisheries and Marine Department, Tadulako University, Palu 94118, Central Sulawesi, Indonesia*
[6]*Aquaculture Study Program, Faculty of Agriculture and Fisheries, Universitas Muhammadiyah Luwuk, Jl. KH Ahmad Dahlan No. 79 Luwuk 94712, Banggai Regency, Central Sulawesi, Indonesia*
[7]*Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science, Berlin 10115, Germany*
[8]*Graduate School, Hasanuddin University, Makassar 90245, South Sulawesi, Indonesia*
[*]For correspondence: inayah.yasir@mar-sci.unhas.ac.id
[†]Contributed equally to this work and are co-first authors
Received 19 April 2024; Accepted 21 April 2025; Published 18 May 2025

Editor: Muhammad Farooq

Abstract

DNA barcoding is a powerful tool for biodiversity exploration and conservation, directly and to support other molecular methods such as environmental DNA (eDNA). However, the usefulness of DNA barcoding depends on the reliability of data in databases such as BOLD and GenBank. In addition to barcodes for around 8,000 fish species, the 5-year FISH-BOL global fish barcoding program produced guidance for reliable barcoding. This included the need for voucher specimens or other strong evidence to verify taxonomic identity and a recommendation for multiple barcodes from different populations of each species. Here we present the case of the Maros karst endemic Celebes rainbowfish *Marosatherina ladigesi* (Ahl 1936), local name "beseng-beseng", from collection to barcoding and voucher specimen preparation to meet such standards. The sole species in the genus *Marosatherina* Ivantsoff *et al.* (1998) (Atheriniformes: Telmatherinidae), *M. ladigesi* is listed as Vulnerable in the IUCN Red List and was represented by a single GenBank accession of an individual of unknown origin sourced from the aquarium trade in the USA. From two studies in the endemic distribution of this fish, we obtained vouchered *M. ladigesi* sequences from two Maros karst rivers. These data contribute to global DNA barcode databases for the Wallacea biodiversity and endemism hotspot while enriching the Hasanuddin University Wallacea collection.

Keywords: Celebes rainbowfish; COI mtDNA; Telmatherinidae; Sulawesi; Wallacea; DNA Barcoding

Introduction

DNA barcoding is a powerful tool for biodiversity exploration and conservation, directly and to support other molecular methods such as environmental DNA (eDNA). However, the usefulness of DNA barcoding depends on the reliability of data in databases such as BOLD and GenBank. In addition to barcodes for around 8000 fish species, the 5-year FISH-BOL global fish barcoding program produced guidance for reliable barcoding. This included the need for voucher specimens or other strong evidence to verify taxonomic identity and a recommendation for multiple barcodes, ideally sourced from different populations of each species (Becker *et al.* 2011; Steinke and Hanner 2011; Ward 2012).

To cite this paper: Yasir I, AA Hidayani, ACMAR Tasakka, NN Kadir, J Jompa, W Umar, M Litaay, A Parenrengi, S Ndobe, A Gani, AP Putri, M Iqram, SBA Omar, AM Moore (2025). Vouchered DNA barcoding to support aquatic biodiversity exploration and management: case study on maros karst endemic *Marosatherina ladigesi*. *Intl J Agric Biol* 34:340305. https://doi.org/10.17957/IJAB/15.2365

Sulawesi is the largest island in Wallacea, an acknowledged biodiversity and endemism hotspot. One of the many unique habitats in the geologically complex island of Sulawesi is the Maros karst formation (Waluyo *et al.* 2005), with many rivers and streams supporting a biodiverse riverine fauna, including native and introduced freshwater and diadromous species of fish and invertebrates (Omar *et al.* 2021; Maulidanti *et al.* 2023; Iqram *et al.* 2024). The native Maros karst ichthyofauna includes at least five restricted range endemic fishes (Omar *et al.* 2021), one of which is the Celebes rainbowfish *Marosatherina ladigesi* (Ahl 1936) with the local name "beseng-beseng" (Omar *et al.* 2021; Maulidanti *et al.* 2023). This fish of the family Telmatherinidae (sailfin silversides), order Atheriniformes, was assessed as Vulnerable (VU) under the International Union for Conservation of Nature (IUCN) Red List (IUCN 2023) in 1996 and 2019 (Daniels and Palmer-Newton 2019). Over-collection for the international aquarium trade and pollution, especially pesticide run off from paddy fields, given as the major threats responsible for both declining population abundance and skewed sex ratios.

Although *M. ladigesi* is represented in major databases such as FishBase, the Global Database of Fishes (Froese and Pauly 2025) and the World Register of Marine Species (WoRMS Editorial Board 2025) and has been assessed under the IUCN Red List (IUCN 2023) criteria, there are relatively few *M. ladigesi* references currently accessible on-line (although more may exist in the so-called "grey literature"). Some of these simply include *M. ladigesi* in species lists (Laan 2019; Hadiaty 2019), while several address various aspects of the biology or ecology of the species (Hadiaty 2007; Nasyrah *et al.* 2020; Omar *et al.* 2020; Thieme *et al.* 2021; Wasiljew *et al.* 2022; Maulidanti *et al.* 2023) and others address taxonomic and/or evolutionary aspects (Ivantsoff *et al.* 1998; Sparks and Smith 2004; Berra 2007; Hadiaty 2007; Stelbrink *et al.* 2014; Campanella *et al.* 2015; Omar *et al.* 2021; Thieme *et al.* 2021; Wasiljew *et al.* 2022). Although four of these studies address phylogenetic issues using molecular methods (Sparks and Smith 2004; Stelbrink *et al.* 2014; Campanella *et al.* 2015; Omar *et al.* 2021), surprisingly there is only one *M. ladigesi* DNA barcode present in the main public nucleotide databases (NCBI GenBank and BOLD). This sequence was obtained from a specimen sourced from the ornamental fish trade in the United States of America (Sparks and Smith 2004), meaning that there were no *M. ladigesi* DNA barcodes in these databases which can be traced to a known geographic origin within the species distribution as recommended by the FISH-BOL global fish barcoding program (Steinke and Hanner 2011).

A DNA barcoding study of Maros karst fish in 2020 (Omar *et al.* 2021) collected specimens in the Pattunuang and Bantimurung rivers. In March 2023, a pilot vouchered DNA barcode program in the Maros karst region collected specimens in the Pattunuang River with the aim of collecting, DNA barcoding and preserving voucher specimens obtained from georeferenced locations within their native habitat following international standards such as those of the FISH-BOL program. This study presents and describes the vouchered and georeferenced DNA barcodes and voucher specimens of the Maros karst endemic rainbowfish *M. ladigesi* from these two research programs.

Materials and Methods

Collection sites and procedures

Maros karst endemic rainbowfish *Marosatherina ladigesi* (Ahl 1936) were collected from two rivers (Bantimurung and Pattunuang) within the endemic distribution of this fish using fine-mesh gillnets. Specimens from Batubassi (n = 1) and Bantimurung (n = 1) rivers were collected in September 2020 (Omar *et al.* 2021). Specimens from the Pattunuang River (n = 3) were collected from a site known as Samanggi in Samangki Village, Simbang District, Maros Regency, South Sulawesi Province, Indonesia in March 2023 (Fig. 1).

The voucher specimens collected in March 2023 were euthanised in a 1% solution of clove oil and then photographed. Each specimen was labelled with the field code, location and date using waterproof paper attached to the specimen using button thread, with the thread entering under the operculum and passed through the gills to the mouth, enabling a removable loop to be formed. The right pectoral fin of each specimen was removed and placed in a 1.5 mL Eppendorf tube filled with 96% absolute ethanol for genetic analysis and marked with the same field code as the specimen. Each specimen (minus the right pectoral fin) was then preserved as a voucher specimen. Morphological identification of all specimens followed references, in particular (Kottelat *et al.* 1993; Ivantsoff *et al.* 1998; Hadiaty 2007).

DNA barcoding

The DNA barcoding process comprised DNA extraction, polymerase chain reaction (PCR), Sanger sequencing, sequence verification and editing. Genomic DNA extraction and PCR were performed at the Maros Brackishwater Research Station, South Sulawesi. Genomic DNA was extracted using Geneaid genomic DNA (Geneaid genomic DNA mini kits following manufacturers protocols) and the product was visualised through electrophoresis (Fig. 2). The target barcode region of the Cytochrome Oxidase I mitochondria DNA (CO1-mtDNA) gene was amplified through PCR with the FISH-BCL forward (5'-TCAACYAATCAYAAAGATATYGGCAC-3') and FISH BCH reverse (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') primer set (Baldwin *et al.* 2009; Omar *et al.* 2021) with 20 μL reaction volumes including 1 μL of each primer. The PCR profile was: initial denaturation at 95ºC for 5 min; 40 cycles of denaturation at 95ºC for 30s, annealing at 50ºC for 30 sec and extension at 72ºC for 45 sec and final extension at 72ºC for 5 min.

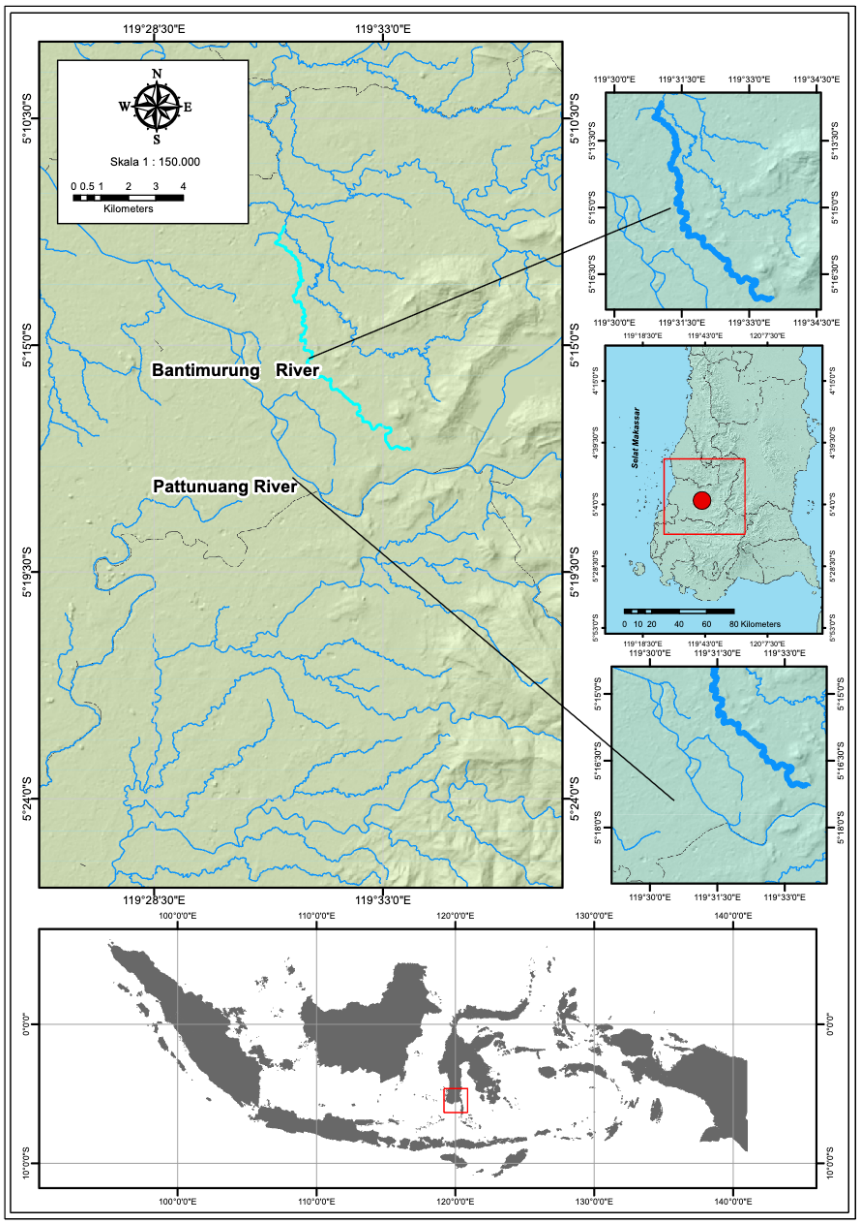


Fig. 1: Locations of sampling sites for *Marosatherina ladigesi* in the Karst Region Maros-Pangkep, South Sulawesi, Indonesia

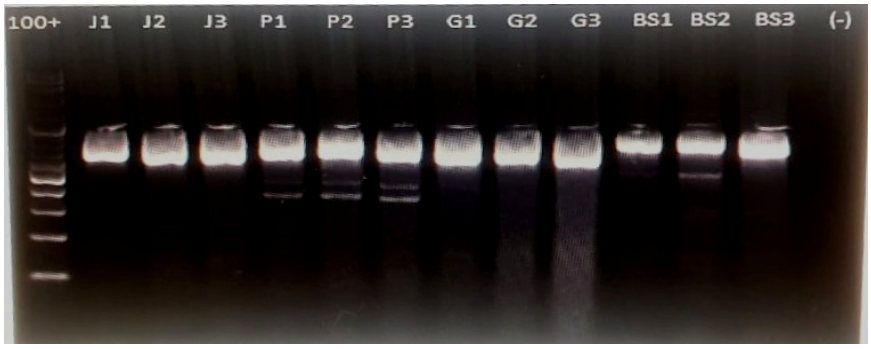


Fig. 2: Electrophoresis of genomic DNA extracted from *Marosatherina ladigesi* specimens (BS1, BS2, BS3) from the Pattunuang River in March 2023 (100 bp Plus DNA Ladder in left hand column)

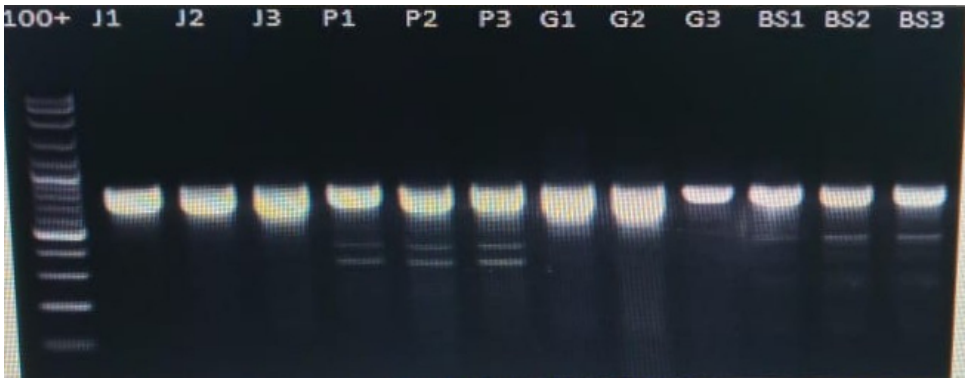


Fig. 3: PCR product (COI mtDNA fragment of *Marosatherina ladigesi* specimens (BS1, BS2, BS3) collected from the Pattunuang River in March 2023 (100 bp Plus DNA Ladder in left hand column)

The PCR amplification product was validated through electrophoresis on 1% agarose gel using a Gel Documentation System (Biometra) with 100 bp Plus DNA Ladder to verify the DNA fragment size (Fig. 3).

PCR product purification and DNA sequencing were performed at 1st Base in Malaysia through PT Genetika Science in Jakarta. The resulting DNA barcode nucleotide sequences in the form of chromatogram trace (.abi) files were checked for quality and the forward and reverse nucleotide sequences cleaned, trimmed and combined in MEGA 11 (Tamura *et al.* 2021) to produce a DNA barcode sequence for each specimen. The nucleotide composition and protein transcription were also computed for each *M. ladigesi* DNA barcode sequence in MEGA 11. Genetic distances between sequences, nucleotide composition and codon protein composition were calculated.

Phylogenetic analysis

Homologous DNA barcode sequences were obtained from the NCBI GenBank nucleotide sequence repository using the Basic Local Alignment Search Tool (Altschul *et al.* 1990) BLASTn routine with default settings and through a search of the NCMBI GenBank nucleotide database using the terms "Telmatherinidae" and "COI". This yielded a total of 20 homologous sequences: one *M. ladigesi* sequence, 12 sequences of the genus *Telmatherina*. A sequence of the Sulawesi endemic halfbeak *Dermogenys orientalis* from the Pattunuang River (Hidayani *et al.* 2024) was added to the data set as an outgroup. The vouchered *M. ladigesi* sequences from two Maros karst rivers and the GenBank accessions were aligned and trimmed, giving a data set of 17 sequences with 606 nucleotide positions. The nucleotide composition and protein transcription for *M. ladigesi* DNA barcode sequences from this study and the homologous GenBank accession were computed in MEGA 11 (Tamura *et al.* 2021) and compared to identify polymorphic sites and determine whether the affected codons represented a change in transcribed proteins.

Phylogenetic analysis was performed in MEGA 11 (Tamura *et al.* 2021) using the Maximum Likelihood method and Kimura 2-parameter model (Kimura 1980) with default vertebrate mitochondrial DNA settings and 1,000 bootstrap replicates to infer the evolutionary history. Initial tree(s) for the heuristic search were obtained through applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach and selecting the topology with the highest log likelihood value. All codon positions were included in the analysis. Based on this trimmed data set, a phylogenetic tree was constructed using the Maximum Likelihood (ML) Kimura-2 parameter method. The tree was exported in Newick format and edited in the online interactive Tree of Life (iToL) (Letunic and Bork 2019, 2021).

Voucher specimen preservation

Voucher specimens were euthanized and preserved following the guidelines of the American Fisheries Society for the use of fish in research (Nickum *et al.* 2004). The labelled voucher specimens were fixed by soaking in 4% formalin and were then rinsed under clean running water until the formalin could no longer be detected by smell. The fixed specimens were placed in 96% ethanol to draw out the excess water. The ethanol was changed twice, after which the specimens were transferred to 70% ethanol which was changed after colouration had occurred until the ethanol remained colourless. The voucher specimens were then registered and stored in a controlled environment as a contribution to the ichthyological section of the recently founded Hasanuddin University Wallacea Centre Natural History Collection.

Results

DNA barcodes and phylogenetic analysis

Combined cleaned DNA barcodes 659 bp in length were obtained for all three Maros karst *M. ladigesi* specimens and were submitted to the NCBI GenBank with accession numbers OR708698 – OR708700. The genetic distances

Table 1: Genetic distance between *Marosatherina ladigesi* COI mtDNA barcode sequences from this study and the NCBI GenBank with outgroup *Telmatherina sarasinorum*

Marosatherina ladigesi source	GenBank accession	*T. sarasinorum*	*Marosatherina ladigesi*		
		MH568800.1	AY290808.1	Pattunuang	Bantimurung
USA Aquarium trade*	AY290808.1	0.254263			
Pattunuang River	OR708700.1	0.253982	0.006400		
Bantimurung River	OR708698.1	0.256168	0.006400	0.001519	
Bantimurung River	OR708699.1	0.256168	0.006400	0.001519	0.000000

*(Sparks and Smith 2004)

Table 2: Nucleotide composition of *Marosatherina ladigesi* COI mtDNA barcode sequences from this study and the NCBI GenBank

Marosatherina ladigesi sequence	Number of nucleotides	Percentage of nucleotides				Undetermined
		T	C	A	G	N
AY290808.1 Aquarium (Sparks and Smith, 2004)	631	31.70	25.83	25.04	17.43	3
Maros, Pattunuang River	659	32.32	25.64	24.58	17.45	0
Maros, Bantimurung River	659	32.47	25.49	24.58	17.45	0
Maros, Bantimurung River	659	32.47	25.49	24.58	17.45	0
Mean		32.24	25.62	24.70	17.45	

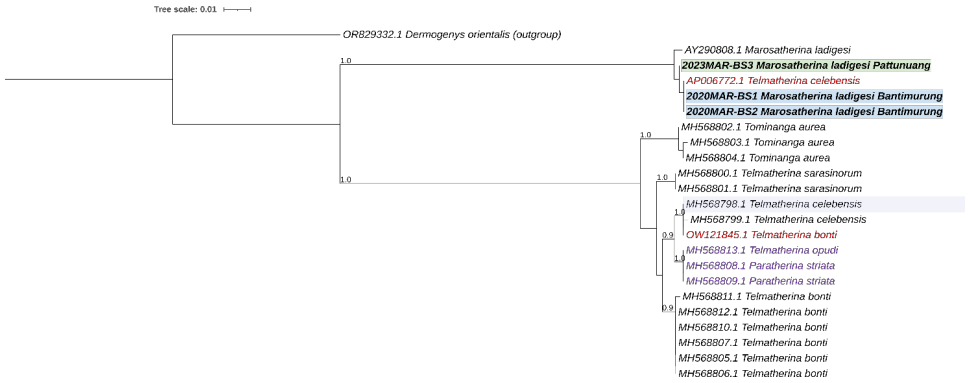


Fig. 4: Maximum Likelihood phylogenetic tree based on COI mtDNA barcodes of *Marosatherina ladigesi* from two Maros karst rivers (this study, bold text, highlight colour by river) and other Telmatherinid NCBI GenBank accessions with outgroup *Dermogenys orientalis*. Text colour: red = suspected misidentified source specimen; purple = identical accession sequences assigned to different taxa

between these sequences (Table 1) were low compared to the distance with the GenBank accession AY290808.1 (Sparks and Smith 2004) from an *M. ladigesi* specimen obtained through the aquarium trade and with the second closest accession (MH568800.1) submitted as *Telmatherina sarasinorum*, also from Family Telmatherinidae.

The Maximum Likelihood phylogenetic tree, with the highest log likelihood (-1751.93) was drawn to scale with branch lengths measured in the number of substitutions per site, was based on a 605 bp aligned dataset obtained after trimming to complete codons present in all sequences at each end (Fig. 4). The *Marosatherina ladigesi* DNA barcodes obtained from the two Maros karst rivers cluster with the sole reference sequence for this species in public databases, accession AY290808 (Sparks and Smith 2004) and are deeply separated from other Telmatherinidae which clustered in a single sister clade. The sequences with the least genetic distance to the *M. ladigesi* sequnce were two *Telmatherina sarasinorum* accessions (MH568801.1 and MH568800.1) obtained from specimens collected from Towuti Lake (Jayadi *et al.* 2019). This tree structure provides further support for the creation of a new genus for this species (Ivantsoff *et al.* 1998).

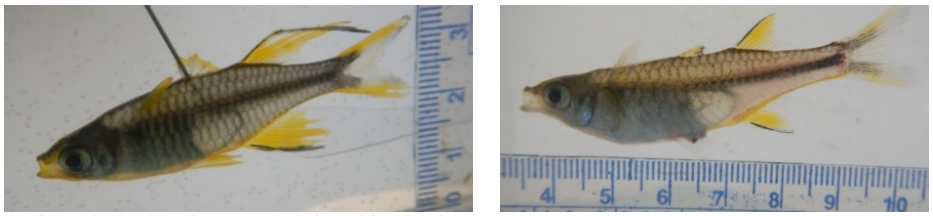
The alignment of all four *M. ladigesi* sequences revealed three alleles with five polymorphic sites. Nucleotide percentages also varied (Table 2), although the percentages of adenine (A) and guanine (G) were the same for all three sequences from specimens collected in the Maros karst region. The barcodes for the two specimens from the Bantimurung River were identical, while the barcode for the specimen from the Pattunuang River had higher cytosine (C) and lower thymine (T) percentages. These differences were due to one polymorphic site, at which the GenBank accession AY290808.1 (Sparks and Smith 2004) had an unidentified nucleotide (N). The GenBank accession AY290808.1 (Sparks and Smith 2004) had higher C and A and lower T and G than any of the Maros karst river specimen sequences due to four polymorphisms, the presence of three unidentified bases (N) in the sequence and slightly shorter sequence length.

The codon protein transcription composition (Table 3) shows differences in the percentages of 18 out of the 19 proteins transcribed between the GenBank accession AY290808.1 (Sparks and Smith 2004) representing a specimen sourced from the aquarium trade in the USA and

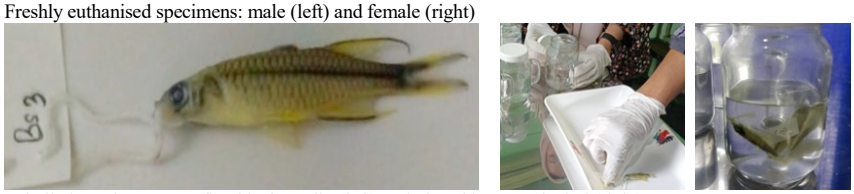
Table 3: Codon protein transcription composition of *Marosatherina ladigesi* COI mtDNA barcode sequences from this study and the NCBI GenBank

Marosatherina ladigesi	Number of codons		Percentage protein transcription composition								
sequence			Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys
AY290808.1*	209		10.05	0.00	3.35	0.96	6.22	9.57	1.91	9.09	0.48
Pattunuang River	219		10.05	0.00	3.20	1.37	6.85	9.59	2.28	8.68	0.46
Bantimurung River	219		10.05	0.00	3.20	1.37	6.85	9.59	2.28	8.68	0.46
Bantimurung River	219		10.05	0.00	3.20	1.37	6.85	9.59	2.28	8.68	0.46
Mean			10.05	0.00	3.23	1.27	6.69	9.58	2.19	8.78	0.46
Marosatherina ladigesi			Percentage protein composition								
sequence	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr
AY290808.1 Aquarium	15.31	5.26	4.78	7.18	2.39	1.44	5.74	6.22	5.74	2.39	1.91
Pattunuang River	15.07	5.02	4.57	7.31	2.28	1.37	5.48	5.94	5.94	2.74	1.83
Bantimurung River	15.07	5.02	4.57	7.31	2.28	1.37	5.48	5.94	5.94	2.74	1.83
Bantimurung River	15.07	5.02	4.57	7.31	2.28	1.37	5.48	5.94	5.94	2.74	1.83
Mean	15.13	5.08	4.62	7.27	2.31	1.39	5.54	6.01	5.89	2.65	1.85

*Three undetermined proteins due to undetermined nucleotides (N)



Freshly euthanised specimens: male (left) and female (right)



Labelled specimen BS3 (fixed in formalin (left) and placed in 70% alcohol (right)

Fig. 5: Voucher specimens of *Marosatherina ladigesi* from Pattunuang River in the Maros Karst region

the three specimens from two rivers in the native Maros karst area to which *M. ladigesi* is considered endemic. Alanine (10.05%) was the same for all and cysteine was zero for all sequences. Of these differences, 10 corresponded to higher percentages and 8 to lower percentages, indicating that this is not simply related to the slightly shorter sequence length. The two alleles recovered from the two Maros karst *M. ladigesi* specimens correspond to the same proteins. This means that the mutations are silent and should not have an effect on the phenotype, although recent research indicates that some such "silent" mutations can affect fitness and come under selection (McVean and Charlesworth 1999; Hunt *et al.* 2014; Bali and Bebok 2015; Liu *et al.* 2021). One of the four polymorphic sites resulted in a different translated protein for accession AY290808.1 (valine) compared to the three Maros sequences (leucine). As the DNA barcode sequences are from a coding region of the mitochondrial genome, the mutations present in the specimen from an unknown population could, therefore, affect phenotype and fitness.

Voucher specimens and scientific collection

Voucher specimens are important to inform future research, for example, enabling taxonomic revisions and validating specimen identification (Baldwin *et al.* 2009; Becker *et al.* 2011; Ward 2012), as was demonstrated in the assignment of *Marosatherina ladigesi* to a new genus, *Marosatherina* (Ivantsoff *et al.* 1998). In addition to the physical specimen, photographic records are also important (Baldwin *et al.* 2009; Becker *et al.* 2011). Photographs of the voucher specimens at different stages of the collection and preservation processes are presented in (Fig. 5).

Discussion

Originally named *Telmatherina ladigesi* by Ahl in 1936, the genus *Marosatherina* Aarn, Ivantsoff & Kottelat was described in 1998 (Ivantsoff *et al.* 1998), with the former *T. ladigesi* as the type species based on a phylogenetic analysis of Telmatherinidae analysis involving 26 coded character states and three additional characters. *Marosatherina ladigesi* is still the sole species described in this genus (Kottelat 2013; Froese and Pauly 2025). Studies on the phylogeny of the order Atheriniformes that include *Marosatherina* differ in their interpretations of morphological, genetic or combined data. According to Sparks and Smith (2004), *Marosatherina* is sister to the

genus *Pseudomugil*, nested within Pseudomugilidae, while within the order Atheriniformes the *Marosatherina-Pseudomugil* clade is more closely related to clades containing rainbowfishes from Australasia (Melanotaeniidae) and Madagascar (Bedotiidae) than to the Atherinidae clade; however, they did not include other Telmatherinidae in their analysis.

The family Telmatherinidae includes 18 recognised species from 5 recognised genera, most of which are considered endemic to Sulawesi (Fricke *et al.* 2025; Froese and Pauly 2025). The most speciose genus is *Telmatherina* (10 species), followed by *Paratherina* (4 species) and *Tominanga* (2 species), while *Kalyptatherina* and *Marosatherina* each comprise one species. Of the 20 homologous telmatherinid sequences obtained from the NCBI GenBank and included in the phylogenetic analysis, 16 (MH568798-MH568813) come from telmatherinid specimens collected in 2018 from Towuti Lake in the Malili ancient lake complex, South Sulawesi (Jayadi *et al.* 2019). The only sequence previously deposited in the NCBI GenBank repository under the name *M. ladigesi* was accession AY290808.1 (Sparks and Smith 2004). This sequence has three indeterminate (N) nucleotide positions and was sourced from a captive fish obtained from the aquarium trade. Therefore, the barcodes provided are the first unambiguous COX1 nucleotide sequences for *M. ladigesi*, as well as the first sequences sources from wild populations in the endemic distribution of this species.

Accession AP006772 also resolved into the *M. ladigesi* clade. This accession was submitted as *Telmatherina celebensis* mitochondrial DNA, complete genome except for D-loop. Submitted in 2004, accession AP006772 appears to be unpublished to date, with no information provided in the GenBank metadata on the provenance of the specimen from which the DNA was isolated. *Marosatherina ladigesi* was originally described as *Telmatherina ladigesi* Ahl (1936) and both *T. celebensis* and *M. ladigesi* are commonly referred to as Sulawesi rainbowfish and it is likely that accession AP006772 is misnamed. This example underscores the need for further barcoding of Sulawesi freshwater fishes, as well as the importance of vouchered sequences with preserved reference specimens, the need for accurate metadata and the risk of species misidentification when relying on sequence databases alone.

The telmatherinid accessions from Towuti Lake present a challenge, as the taxonomic assignments in the published paper (Jayadi *et al.* 2019) represent seven species from three genera (*Telmatherina*, *Paratherina* and *Tominanga*), but the taxonomic assignments differ between Jayadi *et al.* (2019) and the GenBank accession metadata for 13 out of the 16 sequences. Of the three sequences with matching assignments, two *Telmatherina celebensis* accessions (MH568798 and MH568799) form a clade with accession OW121845 submitted as *Telmatherina bonti,* a mitochondrial genome assembled under the Wellcome Sanger Tree of Life Programme with no information on the provenance of the specimen from which DNA was extracted. The third matching accession (MH568813) forms a clade with two accessions (MH568808 and MH568809) submitted as *Paratherina striata* but given as *T. bonti* in Jayadi *et al.* (2019). The other mis-matches between the GenBank metadata (GB) and Jayadi *et al.* (2019) (J) are: *T. sarasinorum* (GB) *vs. P. striata* (J) (MH568800 and MH568801); *Tominanga aurea* (GB) *vs. T. sanguicauda* (J) (MH568802 - MH568804); *T. bonti vs. Paratherina wolterecki* (J) (MH568805 - MH568807) and *T. bonti vs. P. cyanea* (MH568810 - MH568812).

In addition to ambiguities with respect to the taxonomic assignment of several accessions, the BLAST and targeted search results reveal gaps in COI/COX1 DNA barcode sequence coverage for the Telmatherinidae. Assigned GenBank accessions are currently available for four (Fig. 4) out of the 10 currently accepted species in the genus *Telmatherina*, although most could be misnamed at species or genus level, while six species (*T. abendanoni*, *T. albolabiosa*, *T. antoniae*, *T. obscura*, *T. prognatha*, *T. wahjui*) lack accessions. Of the remaining eight recognised species, two species (*Kalyptatherina helodes*, *Paratherina labiosa*) lack accessions, while the taxonomic uncertainties described above make the situation ambiguous for four species (*P. cyanea*, *P. wolterecki*, *T. aurea* and *T. sanguicauda*). The high level of uncertainty regarding the identity of the source material and the accuracy of the taxonomic assignments of currently available GenBank accessions for the family Telmatherinidae underscore the need for vouchered DNA barcode accessions where the source material has a well-defined provenance, has been identified by experts in the relevant taxa and is maintained in a scientific collection for future reference.

Conclusion

From two studies in Maros karst rivers within the endemic distribution of *Marosatherina ladigesi* we obtained the first vouchered *Marosatherina ladigesi* COI mtDNA nucleotide sequences. These reference DNA barcodes contribute to global public DNA barcode databases for the Wallacea biodiversity and endemism hotspot. The study also underscores the need for further vouchered DNA barcoding of Sulawesi freshwater fishes, in particular the Telmatherinidae. The voucher specimens will contribute to the ichthyofaunal collection in the Hasanuddin University Wallacea Centre. The skills and lessons learned from this pilot project will support the expansion of reliable vouchered barcoding for other freshwater and diadromous taxa in Wallacea, with a special focus on the aquatic fauna living in the many rivers and streams of Sulawesi.

Acknowledgments

The authors gratefully acknowledge support from the Hasanuddin University 2023 Research Grant scheme with

contract number: 00323/UN4.22/PT.01.03/2023. They also thank all who assisted in the field and the laboratory.

Author Contributions

All authors contributed to the study's conception and design. IY, AAH, NNK, JJ, ACMART, ML, WU, SN, AG and AMM planned the study and finalized the manuscript. Data collection and specimen preservation were performed by IY, NNK, WU, AG, APP, MI and SBAO. Data analysis and visualisation were carried out by AAH, AP and AMM. Data validation was performed by IY, AAH, ML, SN and SBAO. The first draft of the manuscript was written by IY and AMM. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflict of Interest

All authors declare no conflict of interest.

Data Availability

The nucleotide sequence data presented in this study have been deposited in the NCBI GenBank repository (accessions OR708698.1, OR708699.1 and OR708700.1) and are publicly available. Further data are available on reasonable request to the corresponding author.

Ethics Approval

This study was approved by the Hasanuddin University Research and Community Service Institute under contract Number 00323/UN4.22/PT.01.03/2023

References

Altschul SF, W Gish, W Miller, EW Myers, DJ Lipman (1990). Basic local alignment search tool. *J Mol Biol* 215:403–410

Baldwin CC, JH Mounts, DG Smith, LA Weigt (2009). Genetic identification and color descriptions of early life-history stages of Belizean Phaeoptyx and Astrapogon (Teleostei: Apogonidae) with Comments on identification of adult Phaeoptyx. *Zootaxa* 2008:1–22

Bali V, Z Bebok (2015). Decoding mechanisms by which silent codon changes influence protein biogenesis and function. *Intl J Biochem Cell Biol* 64:58–74

Becker S, R Hanner, D Steinke (2011). Five years of FISH-BOL: Brief status report. *Mitochondr DNA* 22:3–9

Berra TM (2007). *Freshwater Fish Distribution*. Chicago and London, The University of Chicago Press

Campanella D, LC Hughes, PJ Unmack, DD Bloom, KR Piller, G Ortí (2015). Multi-locus fossil-calibrated phylogeny of Atheriniformes (Teleostei, Ovalentaria). *Mol Phylogenet Evol* 86:8–23

Daniels A, A Palmer-Newton (2019). *Marosatherina ladigesi, Celebes Rainbow*. The IUCN Red List of Threatened Species 2019. Accessible athttps://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T21574A126023418.en.(Captured: 3 April 2024)

Fricke R, WN Eschmeyer, RVD Laan (2025). *Eschmeyer's Catalog of Fishes: Genera, Species, Reference, Electronic Version*. Accessible athttp://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp.(Captured: 8 April 2025)

Froese R, D Pauly (2025). *FishBase, The Global Database of Fishes*. Accessible athttp://www.fishbase.org. (Captured: 20 January 2025)

Hadiaty RK (2019). Taxonomical status of endemic freshwater ichthyofauna of Sulawesi. *J Iktiol Indo* 18:175–190

Hadiaty RK (2007). Scientific review of a rainbow fish {Marosatherina ladigesi (Ahl 1936)} an endemic fauna of Sulawesi. *Ber Biol* 8:473–479

Hidayani AA, ACMAR Tassaka, NN Kadir, I Yasir, J Jompa, W Umar, M Litaay, A Parenrengi, S Ndobe, A Gani, AP Putri, M Iqram, SBA Omar, AM Moore (2024). First DNA barcodes of two endemic fishes from the Maros karst region, South Sulawesi, Indonesia. *Cybium* 48:247–250

Hunt RC, VL Simhadri, M Iandoli, ZE Sauna, C Kimchi-Sarfaty (2014). Exposing synonymous mutations. *Trends Genet* 30:308–321

Iqram M, M Litaay, Z Hasyim, WJC Moka, GM Tinungki (2024). Combining eDNA metabarcoding and conventional method in rapid screening of fish diversity in Maros-Pangkep Global Geopark, South Sulawesi, Indonesia. *Biodiversitas* 25:819–828

IUCN (2023). *The IUCN Red List of Threatened Species*. Version 2023-1. Accessible athttps://www.iucnredlist.org.(Captured: 5 March 2024)

Ivantsoff A, W Kottelat, M Ivantsoff (1998). Phylogenetic analysis of Telmatherinidae (Teleostei: Atherinomorpha) with descriptions of Marosatherina: A new genus from Sulawesi. *Ichthyol Explor Freshwaters 9*: 311–323

Jayadi J, I Ilmiah, S Hadijah, M Kasnir, DI Roslim (2019). DNA barcoding of Telmatherinidae family in Lake Towuti, South Sulawesi, Indonesia. *AACL Bioflux* 12:1208–1215.

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

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 Suppl* 27:1–663

Kottelat M, AJ Whitten, S Kartikasari, S Wirjoatmodjo (1993). *Freshwater fishes of Western Indonesia and Sulawesi*. Singapore. Periplus TD

Laan RVD (2019). *Freshwater Fish List: an alphabetic scientific name list of the world's freshwater fishes and an overview of the scientific names used in the aquarium literature*, 26th edn. Almere. Richard van der Laan

Letunic I, P Bork (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucl Acids Res* 47:256–259

Letunic I, P Bork (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucl Acids Res* 49:293–296

Liu Y, Q Yang, F Zhao (2021). Synonymous but Not Silent: The Codon Usage Code for Gene Expression and Protein Folding. *Ann Rev Biochem* 90:375–401

Maulidanti S, SBA Omar, N Nurdin, D Yanuarita, MT Umar, AA Hidayani (2023). Biodiversitas Iktiofauna Sungai Leang-Leang, Kawasan Karst Maros, Sulawesi Selatan. *In: Seminar Nasional Ikan XI, Bogor, 11 June 2022*, pp:62–83. Masyarakat Iktiologi Indonesia

McVean GAT, B Charlesworth (1999). A population genetic model for the evolution of synonymous codon usage: patterns and predictions. *Genet Res* 74:145–158

Nasyrah AFA, MF Rahardjo, CPH Simanjuntak (2020). PELANGI DARI SULAWESI: Marosatherina ladigesi Ahl, 1936 (The rainbow from Sulawesi: Marosatherina ladigesi Ahl, 1936). *Warta Iktiol* 4:21–26

Nickum JG, HLJ Bart, PR Bowser, IE Greer, C Hubbs, JA Jenkins, JR MacMillan, JW Rachlin, JD Rose, PW Sorensen, JR Tomasso (2004). *Guidelines for the Use of Fishes in Research*. Bethesda. American Fisheries Society

Omar, SBA, AA Hidayani, D Yanuarita, MT Umar, S Andriyono (2021). Phylogenetic Analysis of Endemic Fish from the Maros Karst Region, South Sulawesi, Indonesia. *Intl J Agric Biol* 26:661–666

Omar SBA, Kariyanti, D Yanuarita, MT Umar, YSA Lawi (2020). Length-weight relationship and condition factor of the celebes rainbowfish Marosatherina ladigesi, endemic to the Maros karst region, South Sulawesi, Indonesia. *AACL Bioflux* 13:3384–3396

Sparks JS, WL Smith (2004). Phylogeny and biogeography of the Malagasy and Australasian rainbowfishes (Teleostei: Melanotaenioidei): Gondwanan vicariance and evolution in freshwater. *Mol Phylogenet Evolut* 33:719–734

Steinke D, R Hanner (2011). The FISH-BOL collaborators' protocol. *Mitochondr DNA* 22:10–14

Stelbrink B, I Stöger, RK Hadiaty, UK Schliewen, F Herder (2014). Age estimates for an adaptive lake fish radiation, its mitochondrial introgression and an unexpected sister group: Sailfin silversides of the Malili Lakes system in Sulawesi. *BMC Evol Biol* 14:94-107

Tamura K, G Stecher, S Kumar (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol* 38:3022–3027

Thieme P, P Warth, T Moritz (2021). Development of the caudal-fin skeleton reveals multiple convergent fusions within Atherinomorpha. *Front Zool* 18:20-38

Waluyo H, SR Sadikin, Gustami, P Whiting (2005). An economic valuation of biodiversity in the karst area of Maros, south Sulawesi, Indonesia. *Biodiversity* 6:24–26

Ward RD (2012). FISH-BOL, A Case Study for DNA Barcodes. *In: DNA Barcodes: Methods and Protocols*, pp:423–439. Kress WJ, DL Erickson (Eds.). Totowa, NJ Humana Press

Wasiljew BD, J Pfaender, B Wipfler, LL Wantania, F Herder (2022). Functional morphology of prey capture in stream-dwelling sailfin silversides (Telmatherinidae) based on high-speed video recordings. *Zoomorphology* 141:347–359

WoRMS Editorial Board (2025). *World Register of Marine Species.* Accessible athttp://www.marinespecies.org. (Captured: 18 February 2025)