

DNA Barcoding and Phylogenetics of Freshwater Ornamental Fish Using Mitochondrial COI Gene

Sang Ayu Made Putri Suryani^{1*}, I Gusti Ngurah Pramana², I Gede Putu Wirawan³

¹Aquatic Resources Management, Warmadewa University, Indonesia

²Research Center for Fishery, Research Organization for Earth Sciences and Maritime, National Research and Innovation Agency, Indonesia.

³Faculty of Agriculture, Udayana University

Abstract. The purpose of this study was to identify and conserve local fish that have the potential to be ornamental fish. The total sample was 16 fish. Methods employed were DNA isolation, polymerase chain reaction (PCR) and COI gene sequencing. The DNA sequences generated were compared with the sequences from NCBI using the MEGA 7 and BLASTn programs. The results showed that the size of the COI gene produced was 615 bp to 699 bp. As many as twelve species were identified to have the potential to be ornamental fish, namely *Rasbora lateristriata*, *Barbodes binotatus*, *Xiphophorus helleri*, *Sicyopterus microcephalus*, *Puntigrus tetrazona*, *Moenkhausia santaefilomenae*, *Hyphessobrycon eques*, *Poecilia reticulata*, *Poecilia latipinna*, *Xiphophorus maculatus*, *Pethia conchonius*, *Poecilia orri*. These results showed that the local Balinese fish tested had a 70-99% similarity. DNA barcoding and molecular identification can be used for the genetic conservation to prevent local fish populations from becoming extinct.

1 Introduction

DNA barcoding is a method of identifying species using short gene sequences from an organism's genome. The DNA barcode technique can be applied to identify fish species using fish DNA barcodes, and the technology has been widely used [1]. This technique aids in the conservation of biodiversity by identifying species from small or degraded samples, which is essential for the monitoring and management of fishery resources. The cytochrome oxidase subunit I (COI) gene is commonly used as a standard marker for DNA barcoding in animals, including fish [2]. The COI gene sequence of mitochondria has a length of 658 base pairs. The COI gene has several characteristics that can be used to identify species in different animal groups. DNA barcoding has been widely used for identifying freshwater fish diversity in Indonesia. DNA barcoding data provide genetic resources for conservation, studying evolutionary processes, cryptic species, phylogeography, and ecology [3]. Combining DNA barcoding with morphological characters such as body size and color for species identification and phylogenetic analysis can determine intraspecific variation and allow the formation of subspecies [4]. Genetic distance measures genetic differences between populations or species, often using allele frequency data from multiple loci. Smaller genetic distances indicate a higher degree of kinship, which means a closer genetic relationship [5].

* Corresponding author: suryanip@rocketmail.com

The smaller the genetic distance between species, the higher the degree of kinship, and vice versa, with the genetic distance in the range of 0.01000-0.09999 categorized as low [6]. Genetic distance is critical in genetics, forensics, and conservation breeding programs. In breeding, genome linkage estimates can complement populations with poor lineages and ensure better management of genetic diversity [5]. Conservation of biodiversity, such as gene variation, species, and ecosystems, is essential to maintain environmental balance [7] and as an attempt to determine the population origin and biology of invasive species [8]. Species identification by DNA barcoding is accurate and critical to understanding species diversity and distribution and formulating species conservation strategies [7]

Invasive alien species can impact biodiversity, ecosystems, and human well-being. Therefore, there are many national and international policies and regulations to prevent new invasive species and limit their geographical spread through the Biodiversity Conference [9]. In the development of technology, the aquatic environmental DNA (eDNA) method was used for the first time to detect invasive species in freshwater [10]. However, studies on the genetic structure and phylogenetics of local freshwater fish in Bali are still rare. Molecular species identification is an essential initial stage in efforts to develop and manage the wealth of biological resources, especially in the waters of Bali, where there are species whose species names are not known or have not been identified. The distribution of invasive fish is a concern because wild releases in public waters can damage the ecosystem. People often enjoy the beauty, uniqueness, and nature of ornamental fish initially, but they are often thrown away or released on purpose. Several cases of ornamental and exotic fish species have occurred in public aquatic [9]. The spread and control of invasive fish in public waters have not been implemented optimally, so it is necessary to suppress the development and expansion of fish species that are not native by utilizing these native fish as ornamental fish. This study aims to identify local fish species and native fish using the DNA barcoding method (COI gene) to obtain taxonomic certainty for conservation purposes and to obtain kinship relationships between species described by phylogenetic trees. The results of the research contribute to helping to conserve the germplasm of local Balinese freshwater ornamental fish.

2 Methods

This study was exploratory research with molecular identification using the Cytochrome c Oxidase I (COI) gene marker as a marker. Sample and population: The population in this study was fish caught from rivers in Bali in April to May 2024. The total number of samples was 16 fish. Samples consisted of 8 fish from nature and eight fish that were traded, which were stored in 96% ethanol to maintain the durability of the samples. Data collection and instrument: Data collection method uses the PCR, electrophoresis, and sequencing methods.

2.1 DNA extraction and PCR amplification, sequencing

The sample was taken from tissue of fish (approximately 10 grams in size), and an extraction process was performed to isolate the DNA. The extraction method used was following the 10% Chelex protocol. The extraction results were then analyzed for the next stage, namely PCR. The PCR process used Indonesian Biodiversity laboratory protocols. The primers used in the amplification process for fish samples were FISH F1 (5'-TCA ACC AAC CAC AAA GAC ATT GGC AC-3') dan FISH R1 (5'-TAG ACT TCT GGG TGG CCA AAG AAT CA -3' [11] [12]. The total solution of PCR reaction was 26 μ L consisting of a mixture: 2 μ L of DNA template from the extraction result, 1.25 μ L of each primer in a concentration of 10 mM, nine μ L of ddH₂O, and 12.5 μ L of Ready-mix. The reaction mixture was amplified using the Applied Biosystems™ 2720 Thermal Cycler machine. The temperature profile and time of the PCR protocol used were as follows: Initial denaturation: 94 °C for 3 minutes,

then the denaturation stage: 94 °C for 30 seconds, annealing 55-58 °C for 30 seconds, and the extension stage 72 °C for 60 seconds, the denaturation stage to extension was carried out for 38 cycles, the last stage was the final extension: 72 °C for 2 minutes. The PCR results were then visualized with electrophoresis used 1% Agarose gel with Nucleic Acid Gel Stain (GelRed®) staining. Positive samples (flattening DNA bands) were then carried out in a DNA reading process (sequencing) using the Sanger method detected at PT. Genetics Science Jakarta.

2.2 Data analysis

The DNA sequencing results in the form of data sequences (Ab1 files) were then analyzed. The sequence data obtained were edited and aligned using the ClustalW method in the MEGA 7 program. Each base arrangement was checked manually, and all the data used were ensured to be of good quality. PCR and re-sequencing were carried out for data with poor sequencing results. The data were then matched with the database in the data bank (Genbank NCBI) through the Basic Local Alignment Search Tools (BLAST) method on the NCBI website (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Each piece of data was recorded to determine the level of similarity and accuracy. The data were then analyzed using a phylogenetic tree to determine the kinship relationship between the fish samples and confirm the results of BLAST in identifying fish species up to the species level. The phylogenetic tree was created using the Neighbor-joining (NJ) method with 1000 times bootstrap replication on the MEGA 7.

3 Results and Discussion

The collected ornamental fish samples are shown in Fig. 1. The results show that molecular identification using DNA barcoding is an effective way to differentiate between fish species and to determine the diversity of fish species. DNA barcoding has been shown to accurately identify freshwater ornamental fish species, even among related species with similar morphological traits. Twelve types of local ornamental fish were obtained using the DNA barcoding method using the COI gene (Table 1). Four species of samples were obtained in nature, and eight species of samples were traded. These results state the effectiveness of DNA barcodes for studying fish species diversity in Bali. DNA barcoding data are used as a conservation effort to identify species and understand their genetic diversity. This is important for managing and protecting fish populations, as shown in research from the Ili River Valley and Way Kambas National Park [13].

Table 1. Barcode results using BLAST

Field ID	Species	bp	Gene	Accession Number	Query Cover (%)	Identity (%)
1	<i>Rasbora lateristriata</i>	687	COI	LC130775.1	95%	100%
2	<i>Barbodes binotatus</i>	694		MG699678.1	93%	100%
3	<i>Barbodes binotatus</i>	687		MG699648.1	94%	100%
4	<i>Rasbora lateristriata</i>	694		LC130755.1	94%	100%
5	<i>Rasbora lateristriata</i>	687		LC130755.1	95%	100%
6	<i>Xiphophorus hellerii</i>	694		KJ669651.1	94%	100%
7	<i>Xiphophorus hellerii</i>	687		KU693081.1	94%	100%
8	<i>Sicyopterus microcephalus</i>	692		KU693055.1	94%	100%
9	<i>Puntigrus tetrazona</i>	699		OP604403.1	93%	100%
10	<i>Moenkhausia sanctaefilomenae</i>	685		FJ749056.1	96%	100%
11	<i>Hypessobrycon eques</i>	685		OR237777.1	96%	100%
12	<i>Poecilia reticulata</i>	685		PP059110.1	100%	100%
13	<i>Poecilia latipinna</i>	685		KU692773.1	91%	100%
14	<i>Xiphophorus maculatus</i>	685		OM003594.1	95%	100%
15	<i>Pethia conchonius</i>	685		OR148154.1	95%	100%
16	<i>Poecilia orri</i>	615		JQ840648.1	100%	100%

The results of the phylogenetic tree analysis are presented in Figure 2. It shows the evolutionary relationship of one species with another species. Species in different clades, indicating that they have a genetic relationship that is not close to each other. Two species are considered more closely related to each other than to a third species if they share a more recent common ancestor. As shown in Fig.2, the species *Poecilia latipinna* has a close genetic relationship with *Poecilia reticulata*. This single clade or group signifies that individuals between species have a common evolutionary history. *Rasbora lateristriata* and *Barbodes binotatus* are in a single clade, indicating that this species has a close relationship. *Xiphophorus hellerii* and *Scyopus auxilimentus* have two clades, indicating a distinct evolutionary history between their groups. Ten of the sixteen individuals of the species form separate groups. The query cover value of all fish samples is in the range of 91%-100%. This percentage indicates that the length of the sample sequence matches the species database in GenBank. The higher the query cover percentage value, the higher the homology rate [14].

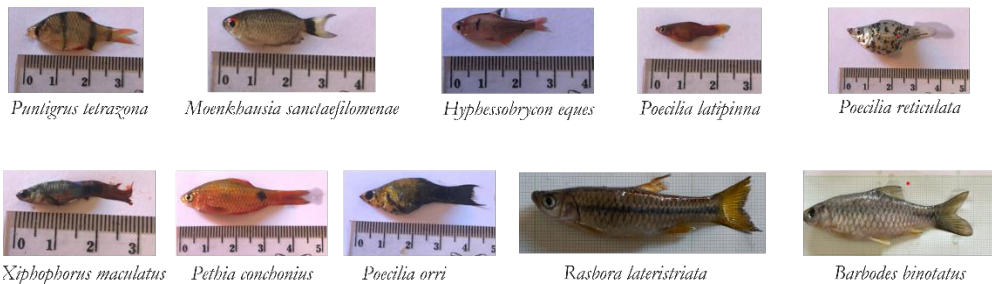


Fig. 1. Ornamental fish samples

Bootstrap analysis was used to test the validity of sequence data for branching in phylogenetic trees. The bootstrap method has a specific value to determine the number of repetitions in the sequence alignment so that the possible kinship of species in the phylogeny tree can be known with more accurate data [14]. In phylogenetic studies, the use of bootstrap analysis is a crucial method with high accuracy and efficiency and is continuously related to the evolution between species [14]

The bootstrap values show the branches' stability in the phylogenetic tree obtained. Higher branch bootstrap values can indicate more substantial branching in the phylogenetic tree of the species. The results of phylogenetic tree analysis support blast and bold system analyses that analyze species based on similarity and homology. It appears that the individuals of each species are separate with outgroups in both a genus and a different class. A phylogenetic tree describes the evolutionary lineage of a species, organism, or from one different ancestor. The kinship between species is seen using phylogenetic tree topology, which can be used to determine kinship relationships in a taxa or sequence. This study detected eight fish species with a query cover of 93% to 95% at the species level. This percentage shows the sample sequence length matching the species database in GenBank. The higher the query cover percentage value, the higher the homology level [15]

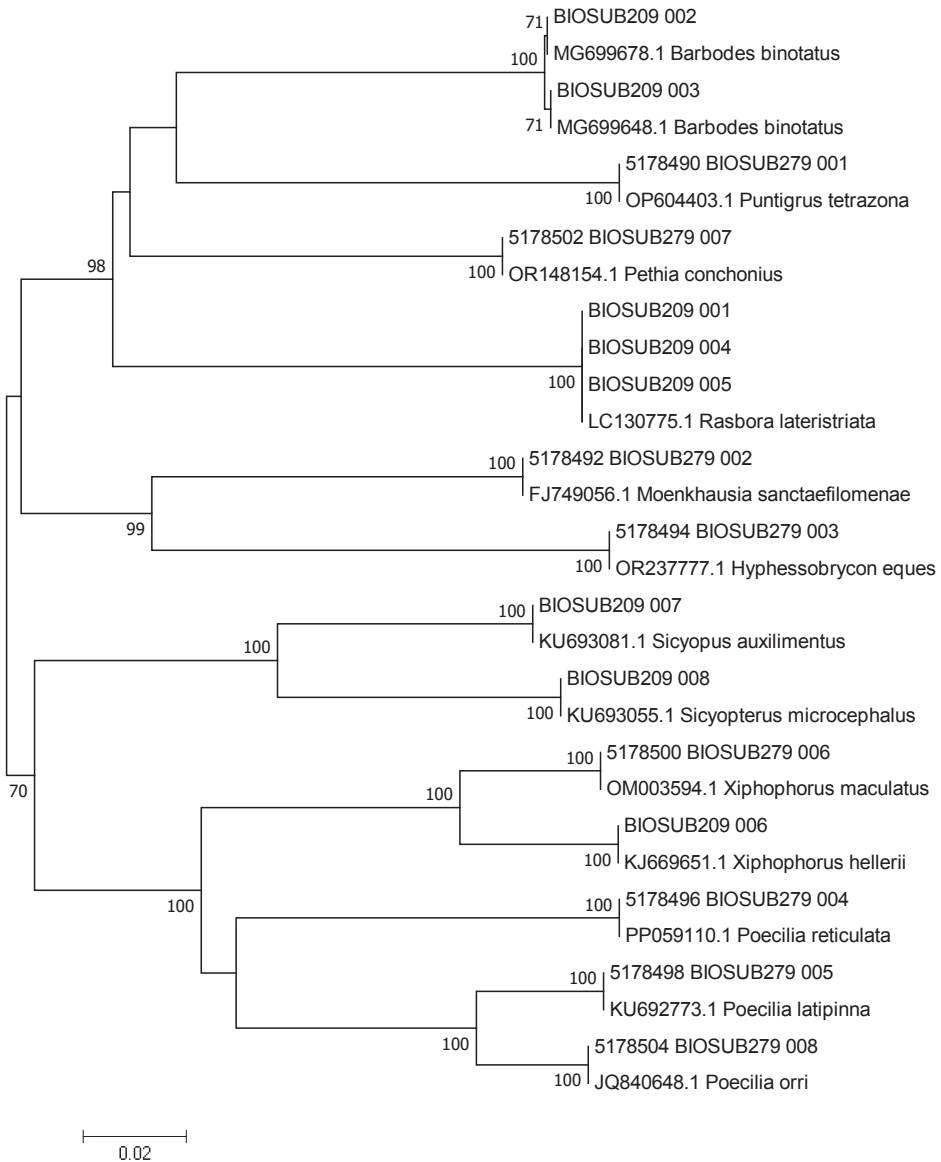


Fig. 2. Phylogenetics of fish samples

4 Conclusion

DNA barcodes have a high consistency of up to 89% in taxonomic groups compared to traditional morphological identification, and their effectiveness can vary depending on taxonomic groups and geographical distance. The COI Gene Phylogenetic Tree is not only used in identifying and measuring genetic distances but can also be used in reconstructing the phylogenies of a species, including the evolutionary process

References

1. Tang Q, Deng L, Luo Q, Duan Q, Wang X, Zhang R (2023) DNA Barcoding of Fish Species Diversity in Guizhou, China. *Diversity* 15:203. <https://doi.org/10.3390/d15020203>
2. Wang L, Wu Z, Sun G, Chen H, Liu H, You F (2022) The preliminary DNA barcoding dataset of fishes from Chinese coastal waters. *Reg Stud Mar Sci* 56:102699. <https://doi.org/10.1016/j.rsm.2022.102699>
3. Dincă V, Dapporto L, Somervuo P, Vodă R, Cuvelier S, Gascoigne-Pees M, Huemer P, Mutanen M, Hebert PDN, Vila R (2021) High resolution DNA barcode library for European butterflies reveals continental patterns of mitochondrial genetic diversity. *Commun Biol* 4:315. <https://doi.org/10.1038/s42003-021-01834-7>
4. Ador K, Gobilik J, Benedick S (2023) Phylogenetic and Morphological Characteristics Reveal Cryptic Speciation in Stingless Bee, *Tetragonula laeviceps* s.l. Smith 1857 (Hymenoptera; Meliponinae). *Insects* 14:438. <https://doi.org/10.3390/insects14050438>
5. Hauser SS, Galla SJ, Putnam AS, Steeves TE, Latch EK (2022) Comparing genome-based estimates of relatedness for use in pedigree-based conservation management. *Mol Ecol Resour* 22:2546–2558. <https://doi.org/10.1111/1755-0998.13630>
6. Achmad DS, Nurdin MS, Taslim I, Moore AM (2023) Population structure of commercially important groupers in the coral triangle Gorontalo, Indonesia to support conservation. *Biodiversitas* 24:6592–6601. <https://doi.org/10.13057/biodiv/d241221>
7. Afu LOA, Chamidah A, Yanuhar U, Maftuch (2023) Species Identification of Sea Bamboo (*Isis hippuris*) Using COI-based DNA Barcoding. *Pertanika J Trop Agric Sci* 46:1347–1357. <https://doi.org/10.47836/pjtas.46.4.17>
8. Blekhman A, Goryacheva I, Schepetov D, Zakharov I (2020) Variability of the mitochondrial COI gene in native and invasive populations of *Harmonia axyridis* Pall. comparative analysis. *PLoS One* 15:1–19. <https://doi.org/10.1371/journal.pone.0231009>
9. Cuthbert RN, Pattison Z, Taylor NG, et al (2021) Global economic costs of aquatic invasive alien species. *Sci Total Environ* 775:145238. <https://doi.org/10.1016/j.scitotenv.2021.145238>
10. Takahashi M, Saccò M, Kestel JH, et al (2023) Aquatic environmental DNA: A review of the macro-organismal biomonitoring revolution. *Sci Total Environ* 873:162322. <https://doi.org/10.1016/j.scitotenv.2023.162322>
11. Roesma DI, Tjong DH, Syaifullah, Nofrita, Janra MN, Prawira FDL, Salis VM, Aidil DR (2023) The importance of DNA barcode reference libraries and selection primer

- pair in monitoring fish diversity using environmental DNA metabarcoding. *Biodiversitas* 24:2251–2260. <https://doi.org/10.13057/biodiv/d240438>
12. Ayu IP, Mashar A, Wardiatno Y, et al (2024) DNA Barcoding and Phylogenetic Analysis of Commercially Important Groupers (Serranidae) in Raja Ampat using gene marker Cytochrome c Oxidase I (COI). *Ilmu Kelaut Indones J Mar Sci* 29:321–328. <https://doi.org/10.14710/ik.ijms.29.3.321-328>
 13. Ariyanti Y, Rini IA, Oktaviani I, Leksikowati SS (2021) Dna barcoding for selected mangrove-based estuary fishes from way kambas national park, Lampung province, Indonesia. *J Trop Life Sci* 11:151–160. <https://doi.org/10.11594/jtls.11.02.04>
 14. Huang J, Liu Y, Zhu T, Yang Z (2021) The Asymptotic Behavior of Bootstrap Support Values in Molecular Phylogenetics. *Syst Biol* 70:774–785. <https://doi.org/10.1093/sysbio/syaa100>
 15. Hall BG (2013) Building phylogenetic trees from molecular data with MEGA. *Mol Biol Evol* 30:1229–1235. <https://doi.org/10.1093/molbev/mst012>