JOURNAL OF TROPICAL LIFE SCIENCE
2024, Vol. 14, No. 3, 593 – 602
http://dx.doi.org/10.11594/jtls.14.03.15

Research Article

Authentication of *Barbonymus* spp. From Lake Singkarak Using DNA Barcoding

Viola Mutiara Salis, Dewi Imelda Roesma *, Djong Hon Tjong, Syaifullah, Aadrean, Dahelmi

Department of Biology, Universitas Andalas 25163, Indonesia

Article history:
Submission May 2024
Revised July 2024
Accepted August 2024

**Corresponding author:*
E-mail: dewiroesma@sci.unand.ac.id

ABSTRACT

The local community distinguishes between *Barbonymus belinka* (Balingka) and *Barbonymus schwanefeldii* (Kapiek) in Lake Singkarak based on size due to the morphological similarities between the two species. From previous reports, *B. belinka* (Balingka) is a fish endemic to Lake Singkarak, West Sumatra, while *B. schwanefeldii* has a wider distribution, including Sumatra, Kalimantan, and Java. Consequently, molecular identification is necessary to discern between the species and to understand the DNA barcode characteristics of fish belonging to the genus *Barbonymus* in Lake Singkarak. One molecular technique utilized for species identification is DNA barcoding, which focuses on the *COI* (*Cytochrome Oxidase Subunit I*) gene. Liver tissue samples from Balingka and Kapiek fish from Lake Singkarak were used in the study. Based on 585 bp of *COI* gene sequences and 30 comparison sequences from BOLD system and GenBank NCBI, seven samples from Lake Singkarak show a genetic distance of 0–1.2% from *B. schwanefeldii* populations elsewhere, with 15 differing nucleotide bases. Moreover, samples from Lake Singkarak show a genetic distance of 7.7–8.2% from *B. belinka* in the BOLD system from Aceh, with 42 differing nucleotide bases. Furthermore, two specific bases are present in *B. schwanefeldii* from Lake Singkarak. Based on the results of this research, it is known that all samples from Lake Singkarak, including Balingka and Kapiek, belong to the same species, namely *B. schwanefeldii*.

Keywords: BOLD system, Cytochrome oxidase 1, Endemic, Phylogenetic, Taxonomic status, West sumatra

Introduction

As the largest freshwater fish family, Cyprinidae has an important role in aquatic ecosystems, including as a source of consumption and economic income for humans. Examples of Cyprinidae fish that have this role are the Kapiek and Balingka fish from Lake Singkarak. Kapiek (*Barbonymus schwanefeldii*) and Balingka (*Barbonymus belinka)* fish belong to the genus *Barbonymus* [1-3]. Communities around Lake Singkarak call one type of fish with two different names based on size. Small fish are called Kapiek and large fish are called Balingka. According to a previous report [4], Kapiek and Balingka are different species. Kapiek (Kepiat) is the vername of *B. schwanefeldii* and Balingka is *B. belinka*. Morphologically, *B. belinka* is difficult to distinguish from *B. schwanefeldii*. Both have similarities, but there are differences in terms of meristics. It is necessary to identify using molecular markers to ascertain the species of *Barbonymus* fish in Lake Singkarak.

Molecular markers are needed to support results in species identification. One method for identifying species is using DNA barcoding. There is a large volume of published studies describing the role of DNA barcoding. DNA barcoding is used for species identification that uses short gene sequences of an organism [5-7]. It can also be used to authenticate species that have taxonomic conflicts and cryptic species [8,9]. DNA barcoding was first introduced, and the *COI* (*Cytochrome Oxidase Subunit I*) gene was used as a molecular marker [10]. The advantage of using *COI* is that the gene section is quite short. These genes make it quick and easy to identify variations between species, have very few deletions and insertions,

How to cite:
Salis VM, Roesma DI, Tjong DH, et al. (2024) Authentication of *Barbonymus* spp. From Lake Singkarak using DNA Barcoding. Journal of Tropical Life Science 14 (3): 593 – 602. doi: 10.11594/jtls.14.03.15.

and conserve many sites [10]. This method is universally proven in species identification, including freshwater fish [11-17].

Molecular studies using DNA barcoding of several fish from the genus *Barbonymus* have been reported. Report on the discovery of *B. gonionotus* in the Ing River, Thailand using DNA barcoding (*COI* gene) [15], identification of Lampam fish as *B. schwanefeldii* from Aceh waters using the *COI* gene. However, molecular research on *B. belinka* (Balingka) has not been investigated. *B. belinka* (Balingka) was distributed in Malacca and Lake Singkarak [4], with Lake Singkarak serving as the type locality [2]. There is currently no reported distribution of *B. belinka* (Balingka) in Malacca, and it only suggests a potential endemism to Sumatra [18,19]. *B. belinka* (Balingka) is identified as an endemic fish to the west coast of Sumatra [19], in contrast, *B. schwanefeldii* (Kapiek) has a broader distribution, encompassing Sumatra, Borneo, Malaya, and Indochina [4]. Additionally, the DNA barcode of *B. belinka* (Balingka) is available in the BOLD system (BOLD: AED2516), but upon verification, it is linked to GenBank under the name *B. gonionotus* (MK978151). The lack of data on *B. belinka* (Balingka), resulting in its status as Data Deficient (DD) on the IUCN Red List, further emphasizes the necessity of conducting this research. This paper aims to authenticate and clarify the misnaming of species that have persisted for years within the local community. It will be the first molecular study of (Balingka), the fish known as Balingka from Lake Singkarak, presenting DNA barcode characters and a phylogenetic tree. This research was carried out to determine the characteristics of the DNA barcode of *B. belinka*, which will be reported and registered in the BOLD system. Molecular data in the BOLD system can be used as a basis for identification, monitoring, and conservation.

Material and Methods

Samples

Kapiek and Balingka were directly collected in the field. The samples that have been obtained are photographed and labeled, and their morphological characteristics are recorded, which are thought to change during storage. The samples consisted of liver tissue stored in microtubes containing absolute ethanol obtained from seven individual fish samples (Table 1). Liver tissue samples were taken to the laboratory for long-term storage in a refrigerator with a maximum temperature of 4^{o}C. In the laboratory, individual samples were stored for one week and then washed with running water for 4 hours to remove residual formalin. Then the samples were placed in a sample box containing 70% alcohol for long-term preservation. These samples were collected from Lake Singkarak (Paninggahan), West Sumatra. The DNA sequences obtained from this study were compared with 30 sequences (Table 2) retrieved from GenBank NCBI and the BOLD system.

Sample sources and DNA extraction

Individual samples were placed in sample boxes containing 10% formalin and transported to the Genetic and Biomolecular Laboratory of Andalas University. Liver tissue samples intended for DNA isolation were stored in micro-Eppendorf tubes that already contain absolute ethanol. The process of tissue DNA extraction followed the GeneAll Biotechnology (Korea). Visualization of isolated DNA was carried out by electrophoresis using 1.2% agarose gel.

Table 1. The samples used were individual samples and liver tissue from 3 individual Kapiek fish and 4 individual Balingka fish

No	Types of Fish (Based on size)	Sample Type	Total Sample	Sample Code
1.	Balingka (big)	Liver tissue	4	BL PG_1 BL PG_2 BL PG_4 BL PG_5
2.	Kapiek (small)	Liver tissue	3	KP PG_1 KP PG_5 KP PG_6

Notes: BL (Balingka), KP (Kapiek), PG (Paninggahan).

Table 2. The list of comparative species consists of the same species, the same genus, and outgroups obtained from the Genbank database and the BOLD system with location, accession number, and BIN ID

No.	Species	No. Accession / BIN ID
1	*B. belinka* BOLD system Aceh	AED2516
2	*B. schwanefeldii* complete genome	NC_024274.1
3	*B. schwanefeldii* Riau	AAU0688
4	*B. schwanefeldii* West Kalimantan	AAU0688
5	*B. schwanefeldii* South Sumatra	AAU0688
6	*B. schwanefeldii* Aceh	MK978150.1
7	*B. schwanefeldii* Malaysia 1	KT001006.1
8	*B. schwanefeldii* Jambi	AAU0688
9	*B. schwanefeldii* Malaysia 2	KT001008.1
10	*B. schwanefeldii* Laos	JQ346171.1
11	*B. schwanefeldii* Thailand 1	MK049364.1
12	*B. schwanefeldii* Thailand 2	MK448176.1
13	*B. gonionotus* Myanmar 1	LC189763.1
14	*B. gonionotus* Myanmar 2	AAD1940
15	*B. gonionotus* India 1	JX181874.1
16	*B. gonionotus* India 2	JX181878.1
17	*B. gonionotus* Fish market	KP856760.1
18	*B. gonionotus* South Sumatra	MZ636555.1
19	*B. gonionotus* South Kalimantan	AAD1940
20	*B. mahakkamensis* South Kalimantan	ADN2907
21	*B. balleroides* Java	KU692330.1
22	*B. altus* Malaysia	KT001010.1
23	*B. altus* Laos	JQ346154.1
24	*B. altus* Kamboja	JX066755.1
25	*Barbodes lateristriga* Thailand	MT483454.1
26	*Barbodes banksi* Jambi	ON668384.1
27	*Barbodes lateristriga* Malaysia	MW168729.1
28	*Barbodes banksi* West Sumatra	ON668382.1
29	*Clarias batrachus*	KU692432.1
30	*Channa striata*	KU692420.1

Polymerase Chain Reaction (PCR) and DNA sequencing

Mitochondrial DNA *COI* gene amplification using universal primers for fish [20]. The composition of PCR components includes 1 µl forward and reverse primers, 11 µl PCR Supermix Bioline (London), 3 µl DNA template, and 9 µl nuclease free water. The PCR step consisted of predenaturation at 95°C for 2.0 min, 35 cycles of denaturation at 94°C for 0.5 min, annealing at 53°C for 0.5 min, extension at 72°C for 1.0 min, and a final step at 72°C for 10.0 min. The PCR product was visualized on 2% agarose gel with SYBR-safe staining from Bioline (London). Visualization of amplification DNA was carried out by electrophoresis using 2% agarose gel. DNA amplification products from the PCR process were sent to Firstbase Company Malaysia for sequencing.

Data analysis

Partial *COI* gene sequences of 37 (7 samples from Singkarak and 30 comparison sequences from GenBank NCBI and BOLD system) were examined. The Basic Local Alignment Search Tool (BLAST) program to assess data similarity. DNA sequencing results were analyzed with the DNAstar program [21] to obtain complete DNA sequences. All sequences were aligned using Clustal X [22]. Haplotype analysis and variations in the nucleotide bases were compared and identified using the DnaSP 6 sequence polymorphism program [23]. The genetic distance was computed by applying the Kimura 2-parameter (K2P) model of sequence evolution [24]. The MEGA 11 program [25] was utilized to construct a phylogenetic tree representing species divergence through Neighbor-Joining (NJ), Maximum Likelihood (ML), Maximum Parsimony (MP), and Minimum Evolution (ME) methods, each with 1000 bootstraps.

Results and Discussion

DNA amplification

The results of the DNA amplification based on the *COI* gene visualized on 2% agarose gel. The length of the DNA sequence that has been amplified is around 700 bp. DNA bands were clear, although some appeared as smears. A clear DNA band indicates that the DNA isolation results are of good quality and are not contaminated by other substances that cause smears. Smears can be caused by contaminants such as proteins or residual solutions used in the DNA isolation stage. Smears on DNA bands could result from residual solutions carried out during the DNA isolation process, leading to DNA degradation [26].

BLAST and alignment analysis

The results of the BLAST analysis revealed that the seven samples from Lake Singkarak (consisting of Kapiek and Balingka) had a similarity value ranging from 98.88% - 99.69% with *B. schwanefeldii* in GenBank. Based on these findings, there is a suspicion that the seven samples belong to *B. schwanefeldii*. All individual sequences were obtained and compared with the *B. belinka* sequence in the BOLD system, and several species in the same genus were obtained from the BOLD system and GenBank. These include *B. schwanefeldii*, *B. gonionotus*, *B. balleroides*, *B. altus*, and *B. mahakkamensis*. Four sequences were added to the outgroup, namely *Channa striata*, *Barbodes lateristriga*, *Barbodes banksi*, and *Clarias batrachus*, so that the total number of sequences analyzed was 37.

The sequence length of 585 bp after alignment is located at positions 5.570 – 6.154 bp in mitochondrial DNA and 85-669 bp in the *COI* gene sequences. Sequence analysis indicated an average frequency of nucleotide bases as follows: Adenine (A) 27.1%, Guanine (G) 16.8%, Thymine (T) 27.5%, and Cytosine (C) 28.6%. The AT content (54.6%) in this study was higher than the GC content (45.4%). Furthermore, out of 585 nucleotide bases, there are 367 bp (63.73%) conserved sites, for 218 bp, 153 bp (26.15%) parsimony sites, and 65 bp (11.11%) singleton sites. The average nucleotide base composition in the vertebrate group has a higher number of AT bases than GC bases due to the relatively high mutation rate (transitions and transversions) in mitochondrial DNA [27] and is similar to previous studies [28, 29]. The number of GC bases in the *COI* gene of the fish group is 42.2% - 47.1% [20].

Variation of nucleotide bases and amino acids

The total sequences analyzed resulted in 27 haplotypes. Five sample individuals from Lake Singkarak (KP PG_6, BL PG_4, BL PG_1, BL PG_5, BL PG_2) grouped on the same haplotype, namely haplotype 1. Meanwhile, the other two sample individuals (KP PG_5, KP PG_1) were on haplotype 2 (Table 3). This difference is likely due to intrapopulation variation.

For the 585 bp partial *COI* gene, 42 nucleotide bases were different between the seven samples from Lake Singkarak and *B. belinka* in the BOLD system (from Aceh). Seven samples from Lake Singkarak and *B. schwanefeldii* from all populations have 15 different nucleotide bases. Transition mutations occurred at sequence 1st, 69th, 144th, 192nd, 231st, 252nd, 336th, 346th, 366th, 391st, 411st, 507th, and 582nd positions. Transversion mutations occur at sequences 3rd and 31st. In the 231st nucleotide base positions, two samples from Lake Singkarak (KP PG_1 and KP PG_5) had different nucleotide bases compared to other samples from Singkarak.

Based on 15 different bases between samples from Lake Singkarak and *B. schwanefeldii* from all populations, 13 were due to transition mutations, and two were due to transversion mutations. Transition mutations occur less frequently than transversion mutations (1:7). Transition mutations are more common than transversion mutations because they are related to the structure of purine or pyrimidine bases [30]. The GC nucleotide base (Guanine-Cytosine) has a triple bond, which requires greater energy to break the bond compared to the AT nucleotide base (Adenine-Thymine), which has a double bond. Based on the results of nucleotide base analysis, there were two different bases between the seven samples from Lake Singkarak and other populations of *B. schwanefeldii*. The specific sites located at the 3rd and 192nd nucleotide positions with changes of guanine to cytosine and adenine to guanine.

Transition mutations are changes in bases that are similar to purines or pyrimidines. One of the 13 bases that experienced a purine transition mutation was base sequence 192. Seven samples from Lake Singkarak had the base Adenine (A) and the other *B. schwanefeldii* had the base Guanine (G) (A-G). The 69th base sequence is an example of a transition mutation between pyrimidines, namely

Table 3. Haplotypes of Kapiek and Balingka fish from Lake Singkarak and other species

No	Haplotype	Species
1	Haplotype 1	BL PG_5, BL PG_4, BL PG_2, KP PG_6, BL PG_1
2	Haplotype 2	KP PG_5, KP PG_1
3	Haplotype 3	*B. belinka* BOLD *system* Aceh
4	Haplotype 4	*B. schwanefeldii complete genome*
5	Haplotype 5	*B. schwanefeldii* Riau
6	Haplotype 6	*B. schwanefeldii* West Kalimantan
7	Haplotype 7	*B. schwanefeldii* South Sumatra
8	Haplotype 8	*B. schwanefeldii* Aceh
9	Haplotype 9	*B. schwanefeldii* Malaysia 1
10	Haplotype 10	*B. schwanefeldii* Jambi
11	Haplotype 11	*B. schwanefeldii* Malaysia 2
12	Haplotype 12	*B. schwanefeldii* Laos
13	Haplotype 13	*B. schwanefeldii* Thailand 1, *B. schwanefeldii* Thailand 2
14	Haplotype 14	*B. gonionotus* Myanmar 1, *B. gonionotus* Myanmar 2
15	Haplotype 15	*B. gonionotus* India 1, *B. gonionotus* India 2
16	Haplotype 16	*B. gonionotus* fish market
17	Haplotype 17	*B. gonionotus* South Sumatra, *B. gonionotus* South Kalimantan
18	Haplotype 18	*B. mahakkamensis* South Kalimantan
19	Haplotype 19	*B. balleroides* Java
20	Haplotype 20	*B. altus* Malaysia
21	Haplotype 21	*B. altus* Laos, *B. altus* Kamboja
22	Haplotype 22	*Barbodes lateristriga* Thailand
23	Haplotype 23	*Barbodes banksi* Jambi
24	Haplotype 24	*Barbodes lateristriga* Malaysia
25	Haplotype 25	*Barbodes banksi* West Sumatra
26	Haplotype 26	*Clarias batrachus*
27	Haplotype 27	*Channa striata*

B. schwanefeldii from West Kalimantan, South Sumatra, Jambi, Laos, Thailand 1, and Thailand 2 have the base Thymine (T) while seven samples from Lake Singkarak and *B. schwanefeldii* from Riau, Aceh, Malaysia has the base Cytosine (C), so the change from Thymine to Cytosine (T-C). Transversion mutations are changes in bases from purine to pyrimidine or vice versa. Transversion mutations occur in the 3rd order base, namely seven samples from Lake Singkarak have the base Cytosine (C) while the other *B. schwanefeldii* have the base Guanine (G). The second transversion mutation occurred in the 31st base sequence, namely *B. schwanefeldii* Malaysia 2 has base C (Cytosine) while the other seven samples and *B. schwanefeldii* sequences have base G (Guanine), resulting in a change from Cytosine (C) to Guanine (G). Based on the results of nucleotide base analysis, there were two different bases between the seven samples from Lake Singkarak and other *B. schwanefeldii* populations. These two bases can be used as specific bases from samples from Lake Singkarak, namely in the 3rd order with a base change from G (Guanine) to C (Cytosine) and the 192nd order from A (Adenine) to G (Guanine).

This research analyzed 195 amino acids located at positions 8 – 224 in the *COI* gene. There are three different amino acids in the sequences analyzed (1st, 11th, and 116th) and 11 amino acids that are silent mutation (23rd, 48th, 64th, 77th, 84th, 112th, 122nd, 131st, 137th, 169th, and 194th). Although there are variations in nucleotide bases between Lake Singkarak samples, the amino acids produced remain the same. The first amino acid in the sequence is a product of mutation. According to the analysis results, seven samples from Lake Singkarak and *B. schwanefeldii* from other populations exhibited the amino acid type Valine (V), whereas *B. schwanefeldii* Malaysia 2 displayed a different amino acid type, namely Methionine (M). This difference is caused by different base compositions. *B. schwanefeldii* Malaysia 2 has a Methionine (M) base type with an ATG base composition. Seven samples from Lake Singkarak and other *B. schwanefeldii* have the same type of amino acid (Valine) with different base compositions. Seven samples from Lake Singkarak had the GTC base sequence, while the other *B. schwanefeldii* had the GTG base sequence.

The second amino acid variation occurs in the 11th amino acid sequence. *B. schwanefeldii* Malaysia 2 with the base composition CCT produces the amino acid type Proline (P), while seven samples from Lake Singkarak and other sequences have the base sequence GCT and produce the amino acid type Alanine (A). The third amino acid variation occurs in the 116*th* amino acid sequence. *B. schwanefeldii* from Laos has the AAC base sequence, which produces the amino acid type Asparagine (N), while seven samples from Lake Singkarak and other sequences have the base sequence GAC and produce the amino acid type Aspartate (D).

The amino acids produced by the seven samples from Lake Singkarak analyzed are of the same type, although there are variations in their constituent bases. This variation is a silent mutation. A silent mutation is a mutation that occurs due to a change in base that does not change the amino acid produced [30]. Two samples from Lake Singkarak (KP PG_5 and KP PG_1) have the CTA nucleotide base sequence, while the other samples have the CTG base sequence but produce the same type of amino acid, Leucine (L). It is a transition mutation (CTG → CTA) which is a silent mutation, so it does not change amino acids.

Phylogenetic tree

The topology of the phylogenetic tree shows two main clades: clade 1 includes a *Barbonymus* group (*B. schwanefeldii, B. belinka, B. altus, B. gonionotus, B. mahakkamensis, B. balleroides*), and clade 2 consists of outgroup species (Figure 1). The phylogenetic tree was constructed to illustrate the relationships between species. The first cluster consists of two subclusters. In the first subcluster, seven samples from Lake Singkarak reside in the same branch and exhibit a genetic distance of 0% - 0.2% (Table 4). This value indicates a low genetic variation among the seven samples from Lake Singkarak. Seven samples from Lake Singkarak were in the same group as *B. schwanefeldii* from various populations with a genetic distance of 0% – 1.2%. Meanwhile, *B. belinka* is in the same subcluster but on a different branch and has a genetic distance of 7.7% – 8%.

The first subcluster is also occupied by *B. belinka, B. mahakkamensis* and *B. altus* which are in different branches for each species. seven samples from Lake Singkarak had a genetic distance of 7.7% - 8% with *B. belinka*, 6.7% - 6.9% with *B. mahakkamensis*, and 7.4% - 8.2% with *B. altus*. Based on this genetic distance, the seven samples from Lake Singkarak are classified as different species and are still in the same genus as *B. belinka, B. mahakkamensis* and *B. altus*. In the second subcluster, seven samples from Lake Singkarak had a genetic distance of 11% - 11.2% with *B. balleroides* and 11.2% - 12.7% with *B. gonionotus*. Based on this genetic distance, the seven samples from Lake Singkarak are classified as different species and are still in the same genus as *B. balleroides* and *B. gonionotus*. Based on the phylogenetic tree (Figure 1), the seven samples from Lake Singkarak still come from the same ancestor as other *Barbonymus* (monophyletic).

In the first subcluster, it is assumed that there are 3 groups between sequences from Singkarak and *B. schwanefeldii* based on river flow. First, a sample of Lake Singkarak with *B. schwanefeldii* from Aceh, Malaysia, and Riau has a genetic distance of 0% - 0.9% (Table 4). Second, the *B. schwanefeldii* sequences from Thailand are closer to Laos, which has a genetic distance of 0.2%. Third, *B. schwanefeldii* sequences from South Sumatra, Jambi, and West Kalimantan with a genetic distance of 0.2% - 0.3%. This is in line with flows in Southeast Asia during the Pleistocene [31]. Lake Singkarak, Aceh, Malaysia, and Riau end in the Strait of Melaka. The mainland of Thailand and Laos are closer together, and then South Sumatra, Jambi, and West Kalimantan end in the Karimata Strait.

B. belinka sequences available in the BOLD system from Aceh are not in the same branch as any sequences. However, it cannot be confirmed that this sequence is correct because, based on reports by Weber et al. [4] *B. belinka* has the type locality in Lake Singkarak [2]. Besides that, previously classified *B. belinka* as *B. gonionotus* based on morphological characters [32]. Furthermore, the sample was identified as *B. gonionotus* by molecular analysis using the *COI* gene and was reported as a cryptic species because it had a genetic distance of 8.4% from *B. gonionotus* in GenBank [16]. So, it is assumed that the existence of *B. belinka* in Aceh cannot be confirmed.

The second cluster consists of two sub-clusters: The third sub-cluster is two species of *Barbodes lateristriga* with a sequence divergence of 16.4% - 16.8% with seven samples from Lake Singkarak. The fourth subcluster is two species of *Barbodes banksi* which have a sequence divergen-

Table 4. The genetic divergence of each species based on the Kimura 2 Parameter (K2P) model using MEGA11

No	Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1	BL PG_5																																				
2	BL PG_4	0.0																																			
3	BL PG_2	0.0	0.0																																		
4	KP PG_6	0.0	0.0	0.0																																	
5	KP PG_5	0.2	0.2	0.2	0.2																																
6	BL PG_1	0.0	0.0	0.0	0.0	0.2																															
7	KP PG_1	0.2	0.2	0.2	0.2	0.0	0.2																														
8	*B. belinka* BOLD SYSTEM Aceh	8.0	8.0	8.0	8.0	7.7	8.0	7.7																													
9	*B. schwanefeldii* complete genome	0.5	0.5	0.5	0.5	0.7	0.5	0.7	8.2																												
10	*B. schwanefeldii* Riau	0.3	0.3	0.3	0.3	0.5	0.3	0.5	8.0	0.2																											
11	*B. schwanefeldii* West Kalimantan	0.9	0.9	0.9	0.9	1.0	0.9	1.0	7.3	0.7	0.5																										
12	*B. schwanefeldii* South Sumatra	1.0	1.0	1.0	1.0	1.2	1.0	1.2	7.6	0.9	0.7	0.2																									
13	*B. schwanefeldii* Aceh	0.7	0.7	0.7	0.7	0.9	0.7	0.9	8.0	0.5	0.3	0.9	1.0																								
14	*B. schwanefeldii* Malaysia 1	0.5	0.5	0.5	0.5	0.7	0.5	0.7	8.2	0.3	0.2	0.7	0.9	0.5																							
15	*B. schwanefeldii* Jambi	1.0	1.0	1.0	1.0	1.2	1.0	1.2	7.1	0.9	0.7	0.2	0.3	1.0	0.9																						
16	*B. schwanefeldii* Malaysia 2	0.9	0.9	0.9	0.9	0.7	0.9	0.7	8.2	0.7	0.5	1.0	1.2	0.9	0.7	1.2																					
17	*B. schwanefeldii* Laos	0.9	0.9	0.9	0.9	1.0	0.9	1.0	7.8	0.7	0.5	0.3	0.5	0.9	0.7	0.5	1.0																				
18	*B. schwanefeldii* Thailand 1	0.7	0.7	0.7	0.7	0.9	0.7	0.9	7.6	0.5	0.3	0.2	0.3	0.7	0.5	0.3	0.9	0.2																			
19	*B. schwanefeldii* Thailand 2	0.7	0.7	0.7	0.7	0.9	0.7	0.9	7.6	0.5	0.3	0.2	0.3	0.7	0.5	0.3	0.9	0.2	0.0																		
20	*B. gonionotus* Myanmar 1	11.4	11.4	11.4	11.4	11.2	11.4	11.2	13.3	11.7	11.4	11.2	11.4	11.0	11.2	11.4	11.6	11.2	11.0	11.0																	
21	*B. gonionotus* Myanmar 2	11.4	11.4	11.4	11.4	11.2	11.4	11.2	13.3	11.7	11.4	11.2	11.4	11.0	11.2	11.4	11.6	11.2	11.0	11.0	0.0																
22	B. *gonionotus* India 2	11.4	11.4	11.4	11.4	11.2	11.4	11.2	12.8	11.6	11.4	11.2	11.4	11.0	11.2	11.4	11.6	11.2	11.0	11.0	0.9	0.9															
23	B. *gonionotus* India 1	11.4	11.4	11.4	11.4	11.2	11.4	11.2	12.8	11.6	11.4	11.2	11.4	11.0	11.2	11.4	11.6	11.2	11.0	11.0	0.9	0.9	0.0														
24	B. *gonionotus* (from fish maarket in Indonesia)	12.7	12.7	12.7	12.7	12.5	12.7	12.5	13.9	13.0	12.7	12.5	12.7	12.2	12.5	12.7	12.9	12.5	12.2	12.2	1.7	1.7	1.6	1.6													
25	B. *gonionotus* South Sumatra	11.4	11.4	11.4	11.4	11.2	11.4	11.2	12.8	11.7	11.4	11.2	11.4	11.0	11.2	11.4	11.6	11.2	11.0	11.0	0.5	0.5	0.3	0.3	1.2												
26	B. *gonionotus* South Kalimantan	11.4	11.4	11.4	11.4	11.2	11.4	11.2	12.8	11.7	11.4	11.2	11.4	11.0	11.2	11.4	11.6	11.2	11.0	11.0	0.5	0.5	0.3	0.3	1.2	0.0											
27	B. *mahakkamensis* South Kalimantan	6.9	6.9	6.9	6.9	6.7	6.9	6.7	9.9	6.9	6.9	6.7	6.9	6.9	7.1	6.9	7.1	7.1	6.9	6.9	12.8	12.8	12.1	12.1	13.9	12.4	12.4										
28	B. *balleroides* Java	11.2	11.2	11.2	11.2	11.0	11.2	11.0	14.0	11.4	11.2	11.4	11.7	10.8	11.0	11.7	11.4	11.4	11.2	11.2	1.0	1.0	1.2	1.2	2.1	0.9	0.9	12.1									
29	B. *altus* Malaysia	8.2	8.2	8.2	8.2	8.0	8.2	8.0	9.5	8.4	8.2	8.0	8.2	8.2	8.0	7.8	8.4	8.4	8.2	8.2	11.0	11.0	10.9	10.9	12.2	11.0	11.0	8.0	10.3								
30	B. *altus* Laos	7.6	7.6	7.6	7.6	7.4	7.6	7.4	8.8	7.8	7.6	7.4	7.6	7.6	7.8	7.2	7.8	7.8	7.6	7.6	10.8	10.8	10.7	10.7	12.0	10.8	10.8	7.8	10.5	0.9							
31	B. *altus* Kamboja	7.6	7.6	7.6	7.6	7.4	7.6	7.4	8.8	7.8	7.6	7.4	7.6	7.6	7.8	7.2	7.8	7.8	7.6	7.6	10.8	10.8	10.7	10.7	12.0	10.8	10.8	7.8	10.5	0.9	0.0						
32	*Barbodes lateristriga* Thailand	16.8	16.8	16.8	16.8	16.5	16.8	16.5	16.9	16.6	16.8	16.3	16.6	16.8	16.6	16.0	17.1	16.6	16.3	16.3	15.9	15.9	16.5	16.5	17.3	16.2	16.2	18.0	15.9	15.7	16.2	16.2					
33	*Barbodes banksi* Jambi	18.6	18.6	18.6	18.6	18.3	18.6	18.3	17.7	18.6	18.9	18.9	19.2	18.6	18.6	18.6	19.1	19.2	18.9	18.9	18.0	18.0	17.0	17.0	18.4	17.2	17.2	19.9	18.3	19.9	19.9	19.9	17.1				
34	*Barbodes lateristriga* Malaysia	16.7	16.7	16.7	16.7	16.4	16.7	16.4	17.4	17.0	16.7	16.7	17.0	16.7	17.0	16.5	16.9	17.0	16.7	16.7	16.1	16.1	16.7	16.7	17.8	16.4	16.4	18.0	16.4	16.0	15.9	15.9	5.7	15.4			
35	*Barbodes banksi* West Sumatra	18.1	18.1	18.1	18.1	17.8	18.1	17.8	18.3	18.1	18.4	18.4	18.7	17.9	18.1	18.1	18.6	18.7	18.4	18.4	17.6	17.6	17.6	17.6	19.0	17.9	17.9	17.9	18.4	18.3	17.9	17.9	16.7	6.9	15.8		
36	*Clarias batrachus*	25.2	25.2	25.2	25.2	24.9	25.2	24.9	25.5	26.0	25.7	25.7	26.0	25.1	25.7	26.0	26.0	26.0	25.7	25.7	24.1	24.1	23.3	23.3	25.1	23.5	23.5	23.2	24.1	25.1	24.6	24.6	23.7	26.3	23.7	24.9	
37	*Channa striata*	27.5	27.5	27.5	27.5	27.5	27.5	27.5	29.6	27.2	26.9	27.5	27.8	26.6	26.9	27.2	27.5	27.8	27.5	27.5	26.4	26.4	26.6	26.6	27.8	26.6	26.6	29.3	26.4	29.3	29.0	29.0	29.6	28.8	27.8	27.5	27.7

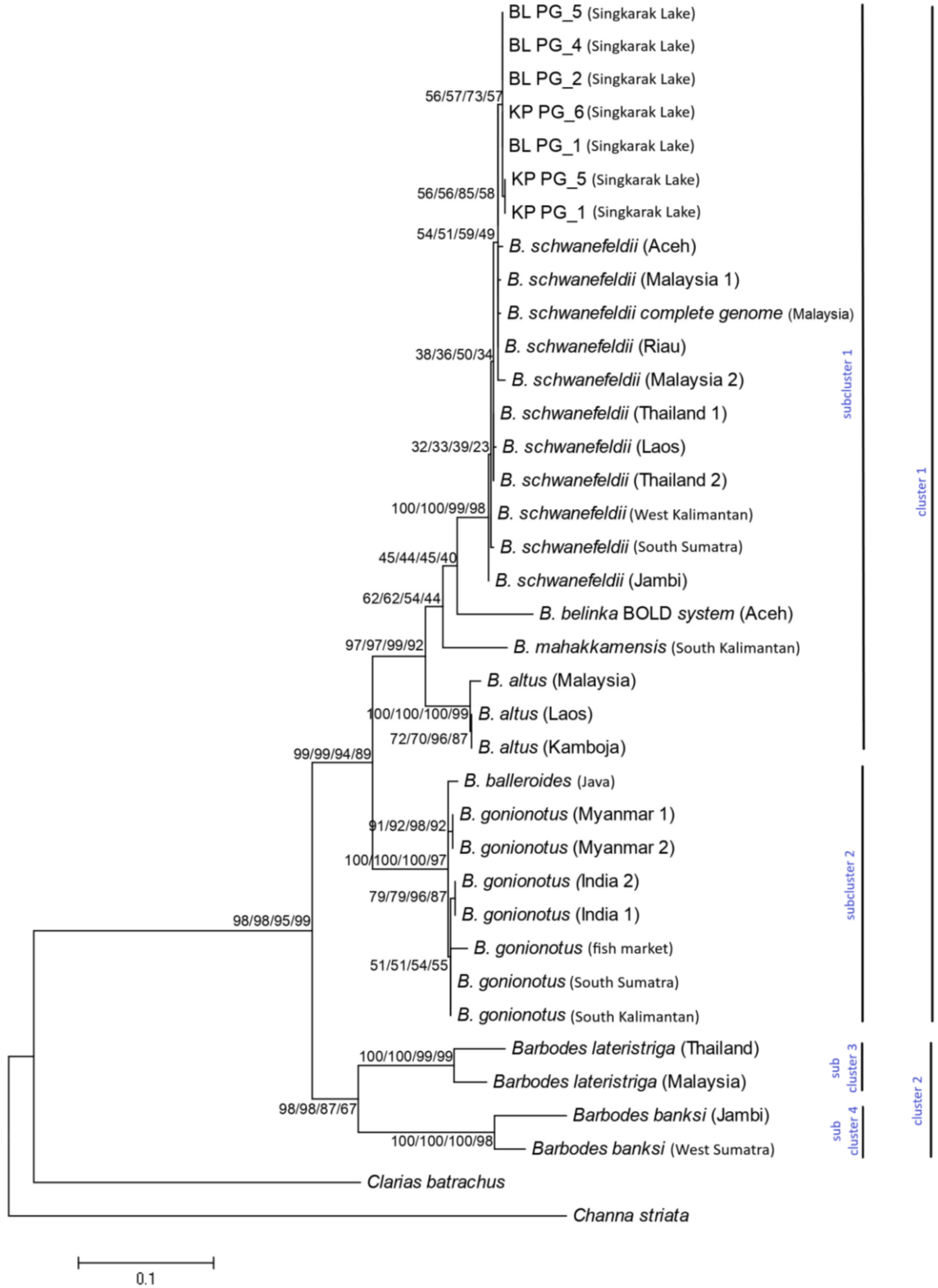

Figure 1. The ML phylogenetic tree is based on K2P genetic distance of COI sequences with 1000 replicates. Bootstrapping (NJ/ME/MP/ML). The ML tree was constructed from sequences found in Lake Singkarak (samples marked with BL PG and KP PG codes), and sequences from the Genbank database and BOLD system.

ce of 17.8% - 18.6% with seven samples from Lake Singkarak. According to [20, 33], different genera within the same family. have a genetic distance of 16.6% ± 0.69.

Conclusion

The DNA barcode characterization of seven samples, Kapiek and Balingka from Lake Singkarak, is *Barbonymus schwanefeldii*. The 585-bp *COI* gene analyzed was located between sequences 85 - 669 in the complete *COI* gene and had two specific bases at the 3rd and 192nd sequences, which were formed due to transversion and transition mutations. The findings of this research about taxonomic status provide insights for future conservation management and support for achieving SDGs zero hunger and life underwater.

Acknowledgment

The authors would like to express our sincere thanks to Universitas Andalas and to the Head of the Biology Department at Universitas Andalas for the laboratory work permit. We also would like to thank our team of Geneticists who helped us with sample collection and laboratory work in the Genetic and Biomolecular Laboratory, Faculty of Mathematics and Sciences, Universitas Andalas, Padang, Indonesia.

References

1. Kottelat M, Kartikasari SN, Wirjoatmodjo S, Whitten T (1993) Freshwater Fishes of Western Indonesia and Sulawesi. Periplus Edition.
2. Kottelat M, Raffles Museum of Biodiversity Research (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. National University of Singapore.
3. Fishbase: Barbonymus Genus (2018) http://www.fishbase.org/Nomenclature/ScientificNameSearchList.php?. Accessed date: 11 December 2022.
4. Weber M, de Beaufort LF (1916) The fishes of the Indo-Australian Archipelago. III. Ostariophysi: II Cyprinoidea Apodes, Synbranchi. Netherlands, Brill Leiden.
5. Kress WJ, García-Robledo C, Uriarte M, Erickson DL (2015) DNA barcodes for ecology, evolution, and conservation. In Trends in Ecology and Evolution 30 (1): 25 – 35. doi: 10.1016/j.tree.2014.10.008.
6. Imtiaz A, Nor SA, Naim DM (2017) Progress and potential of DNA barcoding for species identification of fish species. Biodiversitas Journal of Biological Diversity 18 (4): 1394–1405. doi: 10.13057/biodiv/d180415.
7. Tsoupas A, Papavasileiou S, Minoudi S et al. (2022) DNA barcoding identification of Greek freshwater fishes. PLoS ONE 17(1): e0263118. doi: 10.1371/journal.pone.0263118.
8. Hubert N, Hanner R (2015) DNA barcoding, species delineation and taxonomy: a historical perspective. DNA Barcodes 3 (1): 1–6. doi: 10.1515/dna-2015-0006.
9. Roesma DI, Tjong DH, Munir W, Aidil, DR (2018) New Record Species of Puntius (Pisces: Cyprinidae) from West Sumatra based on Cytochrome Oxidase1Gene. International Journal on Advanced Sciences, Engineering, and Information Technology 8 (1). doi: 10.18517/IJASEIT.8.1.4170.
10. Hebert PDN, Ratnasingham S, DeWaard JR (2003) Barcoding animal life: Cytochrome oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society: Biological Sciences 270 (1). doi: 10.1098/rsbl.2003.0025.
11. Decru E, Moelants T, De Gelas K, Vreven E, Verheyen E, Snoeks J (2016) Taxonomic challenges in freshwater fishes: A mismatch between morphology and DNA barcoding in fish of the north-eastern part of the Congo basin. Molecular Ecology Resources 16 (1): 342–352. doi: 10.1111/1755-0998.12445.
12. Farhana NS, Muchlisin ZA, Duong TY et al. (2018) Exploring hidden diversity in Southeast Asia's *Dermogenys* spp. (Beloniformes: Zenarchopteridae) through DNA barcoding. Scientific Reports 8 (1). doi: 10.1038/s41598-018-29049-7.
13. Roesma DI, Tjong DH, Karlina W, Aidil DR (2019) Taxonomy confirmation of *Puntius cf. Binotatus* from Gunung Tujuh lake, Jambi, Indonesia based on cytochrome oxidase-I (COI) gene. Biodiversitas 20 (1): 54 – 60. doi: 10.13057/biodiv/d200107.
14. Fadli N, Mohd nor SA, Othman AS et al. (2020) DNA barcoding of commercially important reef fishes in Weh Island, Aceh, Indonesia. PeerJ (8). doi: 10.7717/peerj.9641.
15. Panprommin D, Iamchuen N, Soontornprasit K, Tuncharoen S (2020) The utility of DNA barcoding for the species identification of larval fish in the lowering river, Thailand. Turkish Journal of Fisheries and Aquatic Sciences 20 (9): 671–679. doi: 10.4194/1303-2712-v20_9_02.
16. Batubara AS, Muchlisin ZA, Efizon D et al. (2021) DNA barcoding (COI genetic marker) revealed hidden diversity of Cyprinid fish (*Barbonymus* spp.) from Aceh Waters, Indonesia. BIHAREAN BIOLOGIST 15 (1). doi: http://biozoojournals.ro/bihbiol/index.html.
17. Roesma DI, Tjong DH, Janra MN, Aidil DR (2022) DNA barcoding of freshwater fish in Siberut Island, Mentawai Archipelago, Indonesia. Biodiversitas 23 (4): 1795 – 1806. doi: 10.13057/biodiv/d230411.
18. IUCN (2019) The IUCN Red List of Threatened Species. Version 2019-2. www.iucnredlist.org . Accessed date: 1 December 2022.
19. Dahruddin H, Sholihah A, Sukmono T et al. (2021) Revisiting the diversity of barbonymus (Cypriniformes, Cyprinidae) in Sundaland using DNA-based species delimitation methods. Diversity 13 (7). doi: 10.3390/d13070283
20. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society. Biological Sciences 360 (1462): 1847 – 1857. doi: 10.1098/rstb.2005.1716.
21. Burland TG (2000) Dnastar's laser gene sequence analysis software. Methods in Molecular Biology 132: 71 - 91. doi: 10.1385/1-59259-192-2:71.

22. Larkin MA, Blackshields G, Brown NP et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23 (21): 2947 – 2948. doi: 10.1093/bioinformatics/btm404.
23. Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC et al. (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. Molecular Biology and Evolution 34 (12): 3299 – 3302. doi: 10.1093/molbev/msx248.
24. Kimura M (1980) Journal of Molecular Evolution a Simple Method for Estimating Evolutionary Rates of Base Substitutions Through Comparative Studies of Nucleotide Sequences. Journal of Molecular Evolution 16. doi: 10.1007/bf01731581.
25. Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution 38 (7): 3022 – 3027. doi: 10.1093/molbev/msab120.
26. Domingues L (2017) PCR: Methods and Protocols. New York, Springer-Verlag.
27. Yan C, Duanmu X, Zeng L, Liu B, Song Z (2019) Mitochondrial DNA: distribution, mutations, and elimination. Cells 8 (4): 379. doi: 10.3390/cells8040379.
28. Bingpeng X, Heshan L, Zhilan Z et al. (2018) DNA barcoding for identification of fish species in the Taiwan Strait. PLoS ONE 13 (6). doi: 10.1371/journal.pone.0198109.
29. Roesma DI, Tjong DH, Aidil DR (2020) Phylogenetic analysis of transparent gobies in three Sumatran lakes, inferred from mitochondrial cytochrome oxidase I (COI) gene. Biodiversitas 21(1): 43 – 48. doi: 10.13057/biodiv/d210107.
30. Aziz R, Sen P, Beura PK, Das S, Tula D, Dash M, Namsa ND, Deka RC, Feil EJ, Satapathy SS, Ray SK (2022) Incorporation of transition to transversion ratio and nonsense mutations improves the estimation of the number of synonymous and non-synonymous sites in codons. DNA Research 29 (4): dsac023. doi: 10.1093/dnares/dsac023.
31. Voris HK (2000) Maps of Pleistocene Sea Levels in Southeast Asia: Shorelines, River Systems And Time Durations. Journal of Biogeography 27: 1153 - 1167. doi: 10.1046/j.1365-2699.2000.00489.x.
32. Batubara AS, Muchlisin ZA, Efizon D et al. (2018) Morphometric variations of the Genus Barbonymus (Pisces, Cyprinidae) harvested from Aceh Waters, Indonesia. Fisheries and Aquatic Life 26 (4): 231 – 237. doi: 10.2478/aopf-2018-0026.
33. Kartavtsev YP (2011) Divergence at Cyt-b and Co-1 mtDNA genes on different taxonomic levels and genetics of speciation in animals. Mitochondrial DNA 22(3): 55 – 65. doi: 10.3109/19401736.2011.588215.