MJSAT 2 (4) 212-215

MJSAT

Malaysian Journal of Science
and
Advanced Technology

journal homepage: https://mjsat.com.my/

Phylogenetic of Sumatran *Clarias batrachus* based COI Gene

Cinthya Fitri Hardianti[1], Djong Hon Tjong[1], Mansyurdin[1], Syaifullah[1], Dewi Imelda Roesma*[1].

[1] *Department of Biology, Faculty of Mathematics and Natural Sciences, Andalas University, Padang, West Sumatra, Indonesia.*

KEYWORDS

Clarias batrachus
Gene CO1
Phylogenetic
Sumatra

ARTICLE HISTORY

Received 22 November 2022
Received in revised form
17 December 2022
Accepted 19 December 2022
Available online 21 December 2022

ABSTRACT

Clarias batrachus is a species of fish that can be found in Sundaland. The separation of Sundaland resulted in the separation of the islands, one of which was the island of Sumatra. The island of Sumatra is divided into west and east sides by the Bukit Barisan mountains. So it is necessary to study the phylogenetics of *C. batrachus.* The CO1 gene is one of the mtDNA genes that can be used for phylogenetic studies. This study was conducted to determine the geographical relationship to the phylogenetic relationship of *C. batrachus* populations in Sumatra and other populations in Southeast Asia using the CO1 gene and the relationship between geographic isolation in the eastern and western watersheds of Sumatra. Based on 504 bp of *C. batrachus* CO1 gene sequence and 19 comparison sequences, three sub-clusters with two subspecies were formed. Based on the results of this research, it is known that the isolation between West and East Sumatra has a genetic influence with variations in the subspecies level of *C. batrachus* and the separation of Sundaland affects the kinship relationships of these populations with variations in the subspecies level.

1. INTRODUCTION

The order Siluriformes is a group of catfish consisting of 3093 species in 478 genera and 36 families. Clariidae is one of the largest families in Siluriformes, consisting of 16 genera and 57 species [1]. One of the Clariidae genera is *Clarias,* commonly known as catfish. *Clarias* generally have a cylindrical body resembling an eel, a hard and flat head, a wide mouth with four pairs of barbels, a long dorsal fin, and another organ (suprabranchial organ) that allows them to live in oxygen-poor waters [2,3].

Currently, in Indonesia, as many as 16 species of *Clarias* have been identified [1,4]. One of them is *Clarias batrachus* Linnaeus (1758), has elongated shapes, no adipose fins, pectoral fin spine is serrated on the inner edge, head and back are gray, the abdomen is pale gray and white spots on the sides of the body [5,6]. Morphologically, there are variations in the shape and width of the supraoccipital process of *C. batrachus* from Javanese, Malaysian, and Thailand populations [7]. The shape of the supraoccipital process can be used as a marker for *Clarias* [8].

One of the distribution areas of *C. batrachus* is Sumatra, then confirmed that *C. batrachus* comes from Sundaland [6,9]. The land that connected the Malay Peninsula, Sumatra, Java, and Borneo during the last glacial maximum was called Sundaland. Sumatra Island is one of the islands formed by sea level rise (Andaman Sea, South China Sea, and Java Sea) in Sundaland [10,11]. Evolution can occur across a wide range of distributions, and geographical differences can lead to evolution. Adaptation to the environment causes the accumulated genes to differ from one population to another [12]. Diversity and specific characters, morphometrically and genetically, arise as part of the adaptation process of living things to the surrounding environment [13]. This study was conducted to determine the geographical relationship to the phylogenetic relationship of *C. batrachus* populations in Sumatra and other populations in Southeast Asia using the CO1 gene and the relationship between geographic isolation in the eastern and western watersheds of Sumatra.

---

*Corresponding author:
E-mail address: Dewi Imelda Roesma <dewiroesma@sci.unand.ac.id>.

2. MATERIALS AND METHODS

2.1 Sample collection

The samples used in this study are liver samples from nine (n-9) catfishes. The samples collected were from five populations in West Sumatra.

Table 1. Catfish Sampling data

No	Population	Sample Code
1	Batang Mungo	Bt. Mungo
2	Batang Sinamar	Bt. Sinamar 1 Bt. Sinamar 2
3	Batang Agam	Bt. Agam 1 Bt. Agam 2
4	Batang Anai	Bt. Anai 1 Bt. Anai 2
5	Lubuak Paraku River	Lb. Paraku 1 Lb. Paraku 2

(* Batang = River)

2.2 DNA Isolation, Amplification and Sequencing

The DNA isolation steps follow the Invitrogen PureLinkTM Genomic DNA Mini Kit (America) protocol. The results of DNA isolation were then electrophoresed using 1.2% agarose gel. Furthermore, DNA amplification using a PCR reaction mixture consisted of 11µl Supermix Bioline (USA), 9µl Nuclease freewater, 1µl forward primer (0,01 mM), 1µl reverse primer (0,01 mM), 3µl DNA template. Primers for amplification used FISH F1 and FISH R1 [14]. PCR was carried out for 35 cycles. The PCR process begins with predenaturation at 95˚C for 2 minutes, followed by denaturation at 94˚C for 30 seconds, annealing at 53˚C for 30 seconds, and elongation at 72˚C for 1 minute. The final extension was carried out at 72˚C for 10 minutes. The PCR product was electrophoresed using a 2% agarose gel. The PCR products were sent for sequencing at 1st Base Malaysia.

2.3 Data Analysis

The contiguous DNA sequences of *C. batrachus* were then BLAST and compared with 19 comparison sequences downloaded from the NCBI GenBank. DNA sequence alignment and phylogenetic tree creation using MEGA (Molecular Evolutionary Genetics Analysis) software v.7.0.

3. RESULTS AND DISCUSSIONS

3.1 Blast analysis

The result of BLAST analysis of the CO1 gene sequences of *C. batrachus* West Sumatra showed a similarity value of 98.5%-99.6% with the sequences found in Genbank. Based on the BLAST results, 19 CO1 gene sequences were used for comparison (Table 2). The sequences consisted of 18 sequences from the family Clariidae and an outgroup from the family Channidae. The total of all analyzed sequences, including the sample sequences of *C. batrachus*, was 28 sequences.

The length of the CO1 gene aligned to the reference sequence was 504 bp. The results of the *C. batrachus* alignment were 5635-6139 bases of the complete mitochondrial genome and 166-668 bases of the complete CO1 gene. Consists of 334 bp (66.2%) conserved sites (same base), 170 bp (33.7%) variable sites (different bases), 108 bp (21.4%) parsimony sites (more than one different base), and 62 bp (12.3%) singleton sites (one different base). [15] stated that the conserved site value indicated that nucleotide bases were conserved or invariant in all species of the same genus. The value of the point variable indicates the different nucleotide bases of each species, so these bases can be the differentiator of species in the same genus.

The composition of the nucleotide bases based on 504 bp of the CO1 gene of *C. batrachus* was 27.5% A (Adenine), 15.8% G (Guanine), 29.1% T (Thymine), and 27.5% C (Cytosine). The average percentage of nucleotide bases of adenine + thymine (A+T) was 56.7%, while guanine + cytosine (G+C) was 43.3%. In line with the research reported that the nucleotide composition COI gene of *C. batrachus* was 27.4% A (Adenine), 28.4% T (Thymine), 17.9% G (Guanine), and 26.2% C (Cytosine) [16].

Table 2. List of comparison species.

No	*Spesies*	*Accession number*	*Location*
1.	C. *batrachus*	HM345932.1	Aceh, Sumatra
2.	*C. batrachus*	KU692435.1	Banten, Java
3.	*C. batrachus*	KU692436.1	Banten, Java
4.	*C. batrachus*	JF292303.1	Thailand
5.	*C. batrachus*	JF292304.1	Thailand
6.	*C. batrachus*	JF292305.1	Thailand
7.	*C. batrachus*	KC789523.1	Philippines
8.	*C. batrachus*	KC789524.1	Philippines
9.	*C. batrachus*	MG407367.1	Philippines
10.	*C. batrachus*	JF280810.1	Serawak, Malaysia
11.	*C. batrachus*	JF280811.1	Serawak, Malaysia
12.	*C. batrachus*	MW591036.1	Pinang Island, Malaysia
13.	*C. batrachus*	MW591038.1	Pinang Island, Malaysia
14.	*C. fuscus*	MT884512.1	Yunnan, China
15.	*C. gabonensis*	MK074102.1	Africa
16.	*C. teijsmanni*	JN646093.1	Malaysia
17.	*C. pachynema*	KT193020.1	Africa
18.	*C. macrochepalus*	KU495729.1	Philippines
19.	*Channa sttriata*	MN992969.1	Sumatra

3.2 Phylogenetic Analysis

Based on the phylogenetic tree, it can be observed that there are two main clusters. The first cluster is divided into three sub-clusters, namely sub-cluster 1, which consists of the *C. batrachus* populations of Sumatra (Batang Sinamar, Batang Mungo, Batang Agam, Lubuak Paraku, Aceh) and Malaysia; sub-cluster 2, consists of the *C. batrachus* populations of Batang Anai, Sumatra and Banten, Java; and subcluster 3, consists of the *C. batrachus* from other Southeast Asia populations (Philippines and Thailand). The second cluster consists of other Clarias species from the same genus and *Channa striata* (Channidae) as the outgroup species.

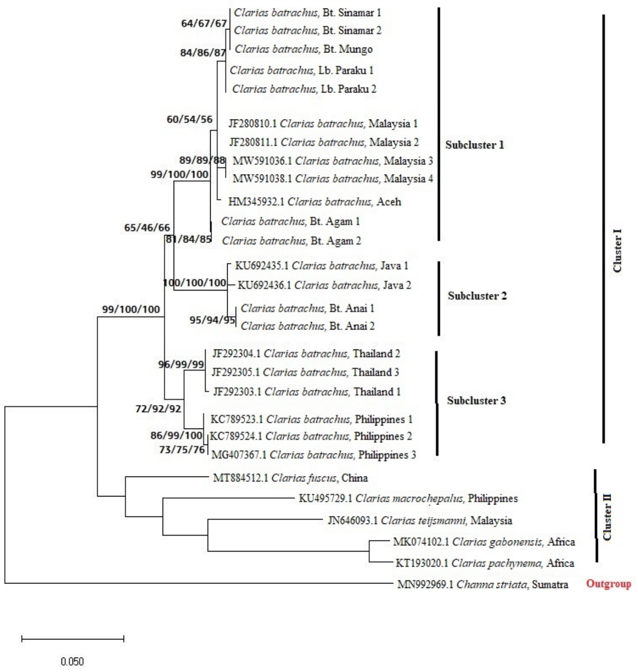

Fig 1. Reconstruction of the phylogenetic tree (COI gene) of Sumatran *C. batrachus* with other *C. batrachus* presented using the Maximum Likelihood (ML) with bootstrap 1000 (ML/NJ/ME).

The value of sequence divergence in subcluster 1 is 0-1%. Based on the CO1 gene, the same species has a genetic distance of 0.89±0.16% [17]. The value indicates that this subcluster has a close relationship. This is because the islands of Sumatra and Malaysia were one landmass in the past. The Malay Peninsula, Sumatra, Java, and Borneo had a geographical connection in the past, namely Sundaland [10]. The eastern watershed of Sumatra shared the same flow as Malaysia and Kalimantan, while the western watershed of Sumatra shared the same flow as Java in the past [18]. Thus, based on the geological record, it supports that *C. batrachus* in the populations of Batang Sinamar, Batang Agam, Batang Mungo, Lubuak Paraku, Aceh, and Malaysia originated from the same species that is still conserved.

The island of Sumatra is divided into the west and east sides by the Bukit Barisan mountains, which stretch from the north of Sumatra island to the south of Sumatra Island [19]. The presence of Bukit Barisan mountains as a barrier between the northern and southern regions of Sumatra Island allows the differences in environmental conditions, which trigger the appearance of genetic variation. The existence of a barrier will cause morphological and genetic variations in a species [20]. The rivers in the western region of Sumatra island are relatively short and flow into the Indian Ocean. While the rivers in the eastern region of Sumatra island are relatively gentle, wide, and long that flow into the South China Sea [21]. Populations originating from watersheds east of Sumatra (Batang Mungo, Batang Agam, Batang Sinamar) and populations originating from watersheds west of Sumatra (Lubuak Paraku, and Batang Anai) have sequence divergence values of 5.0–5.6%. This value indicates that there is variation in the population in the western and eastern river flows of the island of Sumatra. This shows that the geographical conditions of the rivers affect the genetic variation in populations originating from the eastern and western regions of the island of Sumatra.

The second subcluster consisted of four *C. batrachus* individuals from the Batang Anai and Java populations. The value of sequence divergence between the Batang Anai and Java populations is 0.6-0.8%. The sequence divergence-based COI gene has the same species level with a value range of 0.89 ± 0.16% [17]. The phylogenetic tree shows that the Batang Anai population does not cluster with other Sumatra populations. The first subcluster and the second subcluster have a sequence divergence value of 4.3-5.6%. These values indicate *C. batrachus* in the first subcluster is different at the sub-species level from *C. batrachus* in the second subcluster populations.

The value of the sequence divergence between *C. batrachus* Sumatra populations (subcluster 1 and 2) with Thailand and Philippines populations (subcluster 3) was 5.2-5.6%. These values indicate that *C. batrachus* in Thailand and the Philippines (subcluster 3) is different at the sub-species level from *C. batrachus* Batang Anai and Java (subcluster 2). The type locality of *C. batrachus* came from Java [7]. The CO1 gene sequence divergence value of 3.78 ± 1.18% was at the subspecies level [17]. The difference in sequences between populations is due to the physical separation factor between the populations in the Sundaland. Thus, *C. batrachus* in Sumatra populations and *C. batrachus* in Thailand and Philippines populations have genetic variation but are still at the subspecies level. The fragmentation and physical separation of the populations allows a species to have a different genetic structure due to random mutations and genetic drift, resulting in high genetic differentiation, especially in freshwater species [22]. Another study also showed the genetic differentiation of *Channa marulius* due to the differences in river flow in India (Yamuna, Mahandi, and Teesta rivers) [23].

The two mutation types (silent and missense mutations) were observed in all the *C. batrachus* individuals. The silent mutation only changes the base arrangement but does not change the resulting amino acid. While the missense mutation occurs when the base change, and also changes the amino acid into another amino acid [24]. Among the 28 sequences of the COI gene, 22 of them had mutations in their amino acids.

There were 15 silent mutations, one of which occurred in the 17th amino acid sequence in the Batang Anai population had CCC nucleotide bases, while other populations had CCA nucleotide base sequences. However, they have still produced the amino acid proline (P). As many as seven missense mutations occurred. The 2nd amino acid is the first amino acid to undergo a mutation. The results showed almost all *C. batrachus* populations have amino acid mutated, except for the Batang Agam population. In the Sumatra populations, there are four amino acids mutated. One of the mutations occurs in the Batang Mungo, Batang Sinamar, and Lubuak Paraku river populations, where the amino acids change from Isoleusin (I) to Valine (V) because the base change from ATT to GTT.

C. batrachus West Sumatra populations have 46 different nucleotide bases with *C. batrachus* from Java, Aceh, and Southeast Asia, but the amino acids that undergo missense mutation are seven amino acids. If associated with the separation of Sundaland about 12.000 years ago, can be explained that the process of changing the CO1 gene amino acid from *C. batrachus* is slow. In line with the opinion of [25], who states that in the CO1 gene, amino acid changes occur slowly.

4. CONCLUSIONS

Based on the results of this research, it is known that the isolation between West and East Sumatra has a genetic influence with variations in the subspecies level on C. batrachus. Based on phylogeographic studies on C. batrachus in Sumatra and other Southeast Asia, it can be seen that the separation of Sundaland affects the kinship relationships of these populations with variations in the subspecies level.

ACKNOWLEDGMENT

We would like to express our sincere thanks to Biology Department of Andalas University for Laboratory work permit. We also would like to thank our team in Geneticist who helped us in sample collection and laboratory works in Genetic and Biomolecular Laboratory, Faculty of Mathematics and Sciences, Andalas University, Padang, Indonesia

REFERENCES

[1] Ferraris, C. J. 2007. Checklist of catfishes, recent and fossil (Osteichthyes: Siluriformes), and catalogue of siluriform primary types. Zootaxa. 1418: 136-152.

[2] Kottelat, A. J., Whitten, K. S. Nurani., dan W. Soetikno. 1993. Freshwater fishes of western Indonesia and Sulawesi. Pub. Periplus Editions Ltd, Hong Kong. p.107-108.

[3] Teugels, G., D. Adriens. 2003. Taxonomy and phylogeny of clariidae. Science Publisher, Inc. Pp. 465-487.

[4] Ng, H. H dan Hadiay. 2011. Clarias microspilus, a new walking catfish (Teleostei: Clariidae) from Northern Sumatra, Indonesia. Journal of Threatened Taxa. 3(3): 1577-1584.

[5] Więcaszek B., S. Krzykawski. 2010. Morphometric characteristics of asian walking catfish C. batrachus (L., 1758) (Actinopterygii: Clariidae) From The Laboratory Culture. Ejpau. 13(4), 01.

[6] Kottelat, M. 2013. The fishes of the inland waters of Southeast Asia: a catalogue and core bibiography of the fishes known to occur in freshwaters, mangroves and estuaries. Raffles Bulletin of Zoology Supplement No. 27: 242-244.

[7] Ng, H.H dan M. Kottelat. 2008. The identity of C. batrachus (Linnaeus, 1758), with the designation of a neotype (Teleostei: Clariidae). Zoological Journal of the Linnean Society. 153 : 725–732.

[8] Teugels, G. G., R.C. Diego., L. Pouyaud., dan M. Legendre. 1999. Redescription of Clarias macrocephalus (Siluriformes: Clariidae) from Southeast Asia. Cybium, 23: 285–295.

[9] Pouyaud, L., W. Hadie, dan Sudarto. 1998. Mitochondrial DNA differentiation of populations C. batrachus from South East Asia. Proc Midterm meeting of ORSTOM-Catfish Asia Project, Cantho University, Vietnam. 5pp.

[10] Hanebuth T., K. Stattegger., P. M. Grootes. 2000. Rapid Flooding of the Sunda Shelf: a Late-Glacial Sea-Level Record. Science 288(5468):1033-1035.

[11] Sathiamurthy E., H. K. Voris. 2006. Maps of Holocene Sea Level Transgression and Submerged Lakes on the Sunda Shelf. The Natural History Journal of Chulalongkorn University. 2:1-44.

[12] Avise, J.C., J, Arnol., R. M. Ball., Jr, Bermingham., T. Lamb., J.E. Neigel., C.A. Reeb., N. C. Saunders. 1987. Intraspecific Phylogeography: The Mitochondrial Dna Bridge Between Population Genetics and Systematics. Annual Review of Ecology and Systematics, 18, 489–522.

[13] Nguyen, N. H., A. Hamzah., dan T. P. Ngo. 2017. Effects of Genotype by Environment Interaction on Genetic Gain and Population Genetic Parameters in Red tilapia (Oreochromis spp). Front. Genet. 8:82.

[14] Ward. R. D., T. S. Zemlak., B. H. Innes., P. R. Last., dan P. D. N. Hebert. 2005. DNA Barcoding Australia's Fish Species. Phil. Trans. R. Soc. B. 360:1847–1857.

[15] Margoliash, E .1963. Primary Structure and Evolution of Cytochrome C. Proc Natl Acad Sci U S A. 50 (4): 672–679.

[16] Devaasy, A., K. Raj., P. P. Shajita., R. John., K.G. Padmakumar., V.S. Basheer., A. Gopalakrishnan., M. Linu. 2015. Genetic Identification and Phylogenetic Relationships of Indian Clariids Based on Mitochondrial CO1 Sequences. Mitochondrial DNA, Early Online. Hal 1-4.

[17] Kartavtsev, Y. P. H. 2011. Divergence at Cyt-b and Co-1 mtDNA Genes of Different Taxonomic Level and Genetics of Speciation in Animals. Mithocondrial DNA 22(3): 55-56.

[18] Voris, H. K. 2000. Maps of Pleistocene Sea Levels In Southeast Asia: Shorelines, River Systems And Time Durations. Journal of Biogeography. 27:1153- 1167.

[19] Whitten, T., S, J. Damanik, J. Anwar, N. Hisyam. 2000. The ecology of Sumatra. The Ecology of Indonesia Series Vol. 1. Oxford University Pres. 3-8.

[20] Frankham, R., J. D. Ballou, dan D. A. Briscoe. 2002. Introduction to Concervation Genetics. Cambridge University Press.Pp. 635-636.

[21] Colombijn, F.2005. A Moving History of Middle Sumatra 1600-1870. Modern Asia Studies. 39: 1- 38.

[22] Ward, R.D., M. Woodmark, D.O.F. Skibbinski. 1994. A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. Journal of Fish Biology. 44 pp. 213-232

[23] Habib,M., W.S. Lakra, V. Mohindra, P. Khare, A.S. Barman, A. Singh, K.K. Lal, P. Punia, A.A. Khan. 2010. Evaluation of cytochrome b mtDNA sequences in genetic diversity studies of Channa marulius (Channidae: Perciformes). Molecular Biology Reports. 38 (2) (2010), pp. 841-846.

[24] Lodish H., A. Berk., S. L. Zipursky. 2000. Molecular Cell Biology. 4th edition. Newyork. 168-171.

[25] Lynch, M. and P.E. Jarrell. 1993. A Method for Calibrating Molecular Clocks and Its Aplication to Animal Mithocondrial DNA. Genetics 135:1197-1208.