World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

Haplotype diversity of yellow-tail Rasbora population in Sundaland Hotspot based on COI Gene

Furqan Dwiki Lintang Prawira, Syaifullah, Djong Hon Tjong, Aadrean, Tesri Maideliza and Dewi Imelda Roesma *

Department of Biology, Faculty of Mathematics and Natural Science, Andalas University, 25163 Padang, West Sumatra, Indonesia.

World Journal of Advanced Research and Reviews, 2023, 20(01), 708–714

Publication history: Received on 01 August 2023; revised on 13 September 2023; accepted on 15 September 2023

Article DOI: https://doi.org/10.30574/wjarr.2023.20.1.1864

Abstract

Rasbora tornieri is known as a yellow-tail *Rasbora* because of its dark red tail. In Sundaland, *R. tornieri* is distributed in Sumatra and Kalimantan. The barrier caused by the life history between Sumatra and Kalimantan island can influence the genetic variation of yellow-tail *Rasbora*. Therefore, the necessary to conduct the haplotype diversity study of *R. tornieri* using the CO1 gene. Tissue samples of *R. tornieri* were collected from two populations representing upstream and downstream of Musi River in South Sumatra. A total of 633 bp of the CO1 gene were obtained for analysis. Three haplotypes were obtained from *R. tornieri* individuals in Sundaland, with the low haplotype diversity value of 0.28571. The value of nucleotide diversity (π) *R. tornieri* was 0,00045. Overall, the values indicate the low genetic diversity and necessary to maintain also increase the genetic diversity of *R. tornieri*.

Keywords: Haplotype diversity; Barrier; COI gene; Sundaland; *Rasbora tornieri*

1. Introduction

Indonesia has the Sumatra, Kalimantan, and Java island, which are included in the Southeast Asia's five zoogeographical regions. This group is established based on the degree of endemic freshwater fish species found in the area [1]. The three islands are also part of the Sundaland Hotspot, which has the highest freshwater fish diversity in the world [2]. Furthermore, many freshwater fish species in Sundaland cannot be found in other regions. In the Sundaland area, [3] reported as many as 79 species of the Rasboninae group, with 61 Rasbora species.

One of them is *R. tornieri,* which is is known as the yellow-tailed Rasbora because of its dark red tail [4]. Additionally, this species has a sharply demarcated, dark brown mid-lateral stripe which extends the length of the body from the gill covers to the base of the caudal fin. *R. tornieri* is found around the main Sunda and Indochina basins. In Sundaland, *R. tornieri* was spread across the Sumatra and Kalimantan islands, where is two islands are separated by a barrier in the form of sea [3].

The separation of the Sumatra and Kalimantan Islands was influenced by historical events within the Sundaland hotspot, resulting in the formation of a barrier between the two islands [5]. This barrier is considered one of the factors contributing to the morphological and genetic variations observed in species [6]. A study on *Rasbora lateristriata*, showed the difference in molecular and morphological of *Rasbora lateristriata* populations between Java, Lombok, and Sumbawa island [7]. Other studies, also used mitochondrial DNA markers to know the genetic variation of *R. tornieri* and the Rasbora group [7, 8, 9, 10]. The relationship between genetic variation of *R. tornieri* and past events can be

* Corresponding author: Dewi Imelda Roesma

identified using haplotype analysis. DNA isolation , amplification and sequencing can analyze the haplotype diversity to know genetic diversity of *R. tornieri*.

2. Materials and methods

2.1. Sample collection

The samples were collected from two populations in South Sumatra (upstream and downstream of the Musi River) with around 40 km distance (Table 1). The individual fish samples were collected using fishing nets, next labeled and photographed before preserved in 10% formalin solution. Tissue samples were collected and put into a tube containing 96% PA alcohol. The fish samples were deposited in the Genetic and Biomolecular Laboratory, Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Andalas, Padang.

Table 1 List of *R. tornieri* samples and their origin

No.	Population	Sample Code
1	Kertapati River, Palembang, South Sumatra	SK
2	Kenten Laut River, Banyuasin, South Sumatra	KL

2.2. DNA Isolation, Amplification and Sequencing

The DNA isolation steps follow the GeneAll Exgene Genomic DNA micro protocol. DNA isolate was electrophoresed using 1.2% agarose gel. Furthermore, DNA amplification using a PCR Supermix Bioline 11µl, 9µl Nuclease free water, 1µl forward primer, 1µl reverse primer. Primers for amplification used FISH F1 and FISH R1 [11]. PCR was carried out for 35 cycles, beginning with predenaturation at 95℃ for 2 minutes followed by denaturation at 94℃ for 30 seconds, annealing at 53℃ for 30 seconds, elongation at 72℃ for 1 minute and the last final extension at 72℃ for 10 minutes. Proceed to the results of PCR electrophoresed using 2% agarose gel. The PCR products were sent for sequencing at 1st Base Malaysia.

2.3. Data Analysis

The results of *R. tornieri* sequences (forward and reverse) were contig using DNA STAR software [12]. Then, contig sequences were BLAST using the blast tool to check the similarity with the GenBank sequence. Five sequences from GenBank were downloaded and compared with nine *R. tornieri* sequences. The alignment, nucleotide composition, and p-distance information were obtained using MEGA (Molecular Evolutionary Genetics Analysis) software v.11.0 [13].

3. Results and Discussion

3.1. Blast Analysis

The results of BLAST analysis of the approximately 700 bp CO1 gene of *R. tornieri* is shown in Table 2.

Table 2 BLAST analysis result of *R. tornieri*

Samples	Similarity Index		
	R. tornieri **(Sarawak) MK955877**	***R. tornieri*** **(West Kalimantan) MN869459**	***R. tornieri*** **(South Kalimantan) LC130783**
Sk_Rt_001	99.43%	100%	98.78%
Sk_Rt_002	99.29%	99.84%	98.63%
Sk_Rt_003	99.29%	100%	98.77%
Sk_Rt_004	99.43%	100%	98.78%

Sk_Rt_005	99.43%	100%	98.78%
Sk_Rt_006	99.43%	100%	98.78%
Sk_Rt_007	99.43%	100%	98.78%
Kl_Rt_001	99.43%	100%	98.78%
Kl_Rt_002	99.29%	99.84%	98.63%

The results of BLAST showed a similarity value of *R. tornieri* with sequence GenBank ranging from 98.63-100%. According to the BLAST results, five accession numbers were downloaded as comparison species (Table 3). One COI gene data was downloaded from the complete mitochondrial genome, and four COI gene data from the partial mitochondrial genome of *R. tornieri*, approximately 650 bp. A total of 14 *R. tornieri* sequences were analyzed with the length of the CO1 gene of 633 bp.

Table 3 List of accession number comparison species downloaded from NCBI GenBank

No.	Accession number	Location
l	MK955877	Sarawak, Malaysia
2	MN869615	Central Kalimantan 1
3	MN869528	Central Kalimantan 2
4	MN869459	West Kalimantan
5	LC130783	South Kalimantan

3.2. Sequence Analysis

The analysis of the 633 bp of the CO1 gene provides compelling evidence to support its connection to the fascinating narrative of the separation of Sundaland. The nucleotide composition of *R. tornieri* sequences is shown in Table 4. These differences in nucleotide bases cause transversion and translation mutations. Analysis of nucleotide bases will give information on haplotype and nucleotide diversity.

Table 4 Nucleotide composition sequences of *R. tornieri*

No.	Samples	T(U)%	C%	A%	G%
1	Sk_Rt_001	29.9	25.1	26.5	18.5
2	Sk_Rt_002	30	25	26.5	18.5
3	Sk_Rt_003	29.9	25.1	26.5	18.5
4	Sk_Rt_004	29.9	25.1	26.5	18.5
5	Sk_Rt_005	29.9	25.1	26.5	18.5
6	Sk_Rt_006	29.9	25.1	26.5	18.5
7	Sk_Rt_007	29.9	25.1	26.5	18.5
8	Kl_Rt_001	29.9	25.1	26.5	18.5
9	Kl_Rt_002	29.9	25.1	26.5	18.5
Mean		29.91	25.08	26.5	18.5

Table 5 shows the sequences consisting of 624 base (98.5%) as conserved sites, nine base (1.4%) as variable sites, eight base (1.2%) as parsimony site, and one base (0.1%) as singleton site. The analysis of conserved and variable site regions can give information to explain the genetic diversity and evolutionary processes [14]. The conserved and variable site

regions in the COI gene also play a role in determining the organisms up to the species level, making it used as a DNA barcodes [15,16].

Table 5 Nucleotide variation of *R. tornieri* and their nucleotide site

Samples	**Nucleotide Sites**								
	1	1	2	2	3	3	4	4	5
	5	9	0	4	1	9	0	2	5
	4	6	2	4	9	5	6	4	6
SK *R. tornieri* 001*	A	G	C	T	A	C	A	T	G
SK *R. tornieri* 002*	.	.	T	.	.	.	.	.	.
SK *R. tornieri* 003*	.	.	.	.	.	.	.	.	.
SK *R. tornieri* 004*	.	.	.	.	.	.	.	.	.
SK *R. tornieri* 005*	.	.	.	.	.	.	.	.	.
SK *R. tornieri* 006*	.	.	.	.	.	.	.	.	.
SK *R. tornieri* 007*	.	.	.	.	.	.	.	.	.
KL *R. tornieri* 001*	.	.	.	.	.	.	.	.	.
KL *R. tornieri* 002*	.	.	.	.	.	.	.	.	.
MK955877 *R. tornieri* Sarawak	.	.	.	.	.	.	.	.	.
MN869615 *R. tornieri* Central Kalimantan	G	A	.	C	G	T	G	C	A
MN869528 *R. tornieri* Central Kalimantan	G	A	.	C	G	T	G	C	A
MN869459 *R. tornieri* West Kalimantan	.	.	.	.	.	.	.	.	.
LC130783 *R. tornieri* South Kalimantan	G	A	.	C	G	T	G	C	A

Note: Sample*: Sample collected in South Sumatra

The variation in nucleotide bases among individual samples of *R. tornieri* can be related to historical events. The nucleotide base composition of *R. tornieri* population in South Sumatra, West Kalimantan, and Sarawak are identical. While between Central Kalimantan and South Kalimantan, there are differences in nucleotide base composition (Table 5). The nucleotide variation is caused by the distribution of *R. tornieri*, which is affected by past events [17]. The results showed that past events and the geographical distance have contributed to the nucleotide bases diversity in *R. tornieri*.

In SK_R. tornieri_002 sequence showed a different base from other individuals in South Sumatra at base position 202nd (C → T). The variation contributes to the diversity of South Sumatra populations. Furthermore, *R. tornieri* in South and Central Kalimantan populations different in eight nucleotide bases from South Sumatra populations (Table 5). The genetic distance value between individuals from South Sumatra was 0.001579, indicates low genetic variation in *R. tornieri* from South Sumatra. The study [18] reported a low genetic variation (0%) in *Rasbora* sp. from Beratan Lake, Bali.

3.3. Haplotype and haplotype network analysis

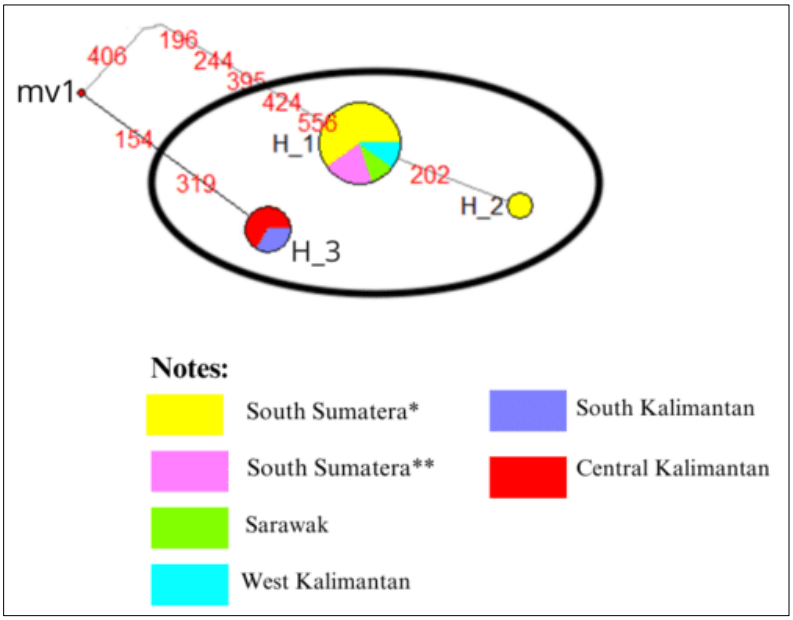
Among the 14 sequences analysis, were obtained three haplotypes of *R. tornieri* (Table 5 and Figure 1). The haplotype diversity is influenced by nucleotide base variation in different populations of *R. tornieri* (Tables 3 and 4). The study [8] shows that the haplotype diversity of *Rasbora* sp. in Lakes Indonesia is due to the differences in origin populations.

Table 6 Haplotypes of *R. tornieri* with CO1 gene

Haplotype	Species	Locations
Haplotype 1	*R. tornieri*	South Sumatra*, South Sumatra**, MK955877 Sarawak, MN869459 West Kalimantan
Haplotype 2	*R. tornieri*	South Sumatra*
Haplotype 3	*R. tornieri*	MN869615 Central Kalimantan, MN869528 Central Kalimantan, LC130783 South Kalimantan

Notes: South Sumatra*: SK samples, South Sumatra**: KL samples

Analysis of the haplotype network of COI gene sequences in nine *R. tornieri* individuals and five comparison sequences (Table 3) is shown in Figure 1.



Notes: South Sumatra*: SK samples, South Sumatra**: KL samples

Figure 1 Haplotype network of *R. tornieri* based on CO1 gene using Network 10.0 software

Haplotype 1 (H1) consisted of *R. tornieri* populations from South Sumatra, West Kalimantan, and Sarawak. Haplotype 3 (H3) consisted of *R. tornieri* populations from Central Kalimantan and South Kalimantan. The two haplotypes (H1, H3) can be connected due to a transgression event that causes the separation of the Sundaland. The impact of the past history between the islands of Sumatra and Kalimantan becomes one of the reasons for haplotype diversity [17, 19]. The movement direction of the tectonic plates from south to north also explains the *R. tornieri* distribution [17]. Thus, it can be seen that past history and geographic distances cause haplotype diversity.

The Musi River flow has an outlet in the Bangka Strait [20]. [19] stated that ancient River flows in Bangka Island, Belitung Island (East Sumatra), and Karimata Island (West Kalimantan) were connected in the past. Thus, it can be the reason for sharing the same haplotype (H1) between Sumatra and Kalimantan. Besides, the relatively closer geographical distance between South Sumatra, West Kalimantan, and Sarawak also causes the sharing of the same haplotype. The ancient rivers in Sumatra (Indragiri River, Batang Hari River, and Musi River) are also connected with several river flows in Central Kalimantan and South Kalimantan. Thus, it can be the reason for sharing the same haplotype (H3) between Sumatra and Kalimantan. Besides, the relatively closer geographical distance between Central Kalimantan and South Kalimantan also causes sharing of the same haplotype. These two reasons can emphasize the separation between H1 and H3.

The variety of one nucleotide base of *R. tornieri* from South Sumatra causes a variety of haplotypes, namely H1 and H2 (Table 5). The variety of haplotype diversity in one population have previously been reported in research [21] that there is haplotype diversity in the same population.

3.4. **Haplotype Diversity (Hd)**

Analysis of haplotype diversity (Hd) and nucleotide diversity (π) using the DNA Sequence Polymorphism 5.10 program (Table 7).

Table 7 Haplotype diversity analysis of *R. tornieri*

Locations	**n**	**Hn**	**Hd**	***π***
Kertapati River	7	2	0.28571	0.00045
Kenten Laut River	2	1	0.0000	0.00000

Note: n-number of samples obtained; Hn-number of haplotypes; Hd-haplotype diversity; *π*-nucleotide diversity

Overall, *R. tornieri* populations have a haplotype diversity (Hd) value of 0.28571. [22] stated that the range values of haplotype diversity (Hd) are 0≥0.5≤1. Based on the range value, *R. tornieri* populations have low haplotype diversity.

The low haplotype diversity (Hd) value of *R. tornieri* is proportional to the genetic distance value, which ranges from 0.001579 to 0.014218. The low values of haplotype diversity (Hd) and nucleotide diversity (π) of *R. tornieri* were influenced by the location of origin of the individuals and environmental factors (salinity, water temperature, air temperature and current velocity) which tended to be homogeneous. Theese individuals are in the same river basin with a distance of about 40 KM from each other. In line with [23] high haplotype diversity is influenced by the large population size and environmental heterogeneity within a population.

4. Conclusion

Three haplotypes of 14 *R. tornieri* were observed in Sundaland, based on the CO1 gene. The haplotype diversity of *R. tornieri* was of 0.28571, showing that the genetic variation of *R. tornieri* in Sundaland was included in the low category.

Compliance with ethical standards

Acknowledgments

We thank the Head of the Department who permitted us to work in the Genetic and Biomolecular Laboratory under the Biology Department of Andalas University. We thank the volunteers who assisted with fieldwork: a collection of individual also tissue samples.

Disclosure of conflict of interest

There is no interest in the conflict between the authors of this piece of research work. The authors agreed and assigned in hand to all matter arise to this piece of research work

Reference

[1] Zakaria-Ismail, M. 1994. Zoogeography and biodiversity of the freshwater fishes of Southeast Asia. *Ecology and Conservation of Southeast Asian Marine and Freshwater Environments Including Wetlands*, 41–48.

[2] Lohman, D. J., M. de Bruyn., T. Page., K. V. Rintelen., R. Hall., Peter. K. L. N. H. Shih. G. R. Carvalho., T. V. Rintelen. 2011. Biogeography of the Indo-Australian Archipelago. *Annu. Rev. Ecol. Evol. Syst.* 42, 205- 226.

[3] Sholihah, A., E. Delrieu-Trottin., T. Sukmono., H. Dahruddin. R. Risdawati., R. Elvyra., A. Wibowo., K. Kustiati. F. Busson., S. Sauri., U. Nurharman., E. Dounias., M.S.A. Zein., Y. Fitriana., I.V. Utama., Z.A. Muchlisin., JF. Agnese., R. Hanner., D. Wowor., D. Steinke., P. Keith., L. Ruber dan N. Hubert. 2020. Disentangling the taxonomy of the subfamily Rasborinae (Cypriniformes, Danionidae) in Sundaland using DNA barcodes. *Sci Rep* 10, 2818.

[4] Ng, H.H. dan M. Kottelat. 2013. The identity of the cyprinid fishes *Rasbora dusonensis* and *Rasbora tornieri* (Teleostei: Cyprinidae). *Zootaxa* 3635(1), Hal 62-70.

[5] Voris, H.K. 2000. Maps of pleistocene sea levels in southeast asia: Shorelines, river systems and time durations. *Biogeography*, 27: 1153-1167.

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

[7] Kusuma, W.E., S. Ratmuangkhwang dan Y. Kumazawa. 2016. Molecular phylogeny and historical biogeography of the Indonesian freshwater fish *Rasbora lateristriata* species complex (Actinopterygii: Cyprinidae): Cryptic species and west-to-east divergences. *Molecular Phylogenetics and Evolution* Vol 105, Pages 212-223.

[8] Muchlisin, A.Z., N. Fadli., M.N. dan Siti-Azizah MN. 2012. Genetic variation and taxonomy of Rasbora group (Cyprinidae) from Lake Laut Tawar, Indonesia. *J Ichtyol* 52: 284-290.

[9] Sholihah, A., E. Delrieu- Trottin., T. Sukmono., H. Dahrudin., J. Pouzadoux., M-K. Tilak., Y. Fitriana., J-F. Agnese., F. L. Condamine., D. Wowor., L. Ruber., N. Hubert. 2021. Limited dispersal and in situ diversification drive the evolutionary history of Rasborinae fishes in Sundaland. *Journal of Biogeography* ;48:2153–2173.

[10] Roesma, D.I., D. H. Tjong., M. N. Janra., D. R. Aidil. 2022. Dna barcoding of freshwater fish in Siberut Island, Mentawai Archipelago, Indonesia. *Biodiversitas*, Vol (23) No 4, p 1795-1806.

[11] 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.

[12] Koonin, E. V., M. Y. Galperin. 2003. Sequence- Evolution- Function: Computational Approaches in Comparative Genomics. *Boston: Kluwer Academic*.

[13] Burland, T. G. 2000. DNASTAR's Lasergene Sequence Analysis Software. *Methods Mol Biol* 132: 71-91.

[14] Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol*. 28(10):2731–2739.

[15] Luo, A., A. Zhang., S. Y. Ho., W. Xu., Y. Zhang., W. Shi., S. L. Cameron., C. Zhu. 2011. Potential efficacy of mitochondrial genes for animal DNA barcoding: a case study using eutherian mammals. *BMC Genomics, 12:84.*

[16] Heyden, S. V. D., M. R. Lipinski, C. A. Matthee. 2010. Remarkably low mtDNA control region diversity in an abundant demersal fish. *Mol Phylogenet Evol*, 55(3): 1183-8.

[17] Metcalfe, I. 2005. Asia South-East. *Encyclopedia of Geology ASIA South-East*, 169–*196*.doi:10.1016/B0-12-369396-9/00395-6.

[18] Kartika, G. R. A., P. G. S. Juliyantoro. 2018. DNA Barcoding Reveal The Current Status Unevaluated Species of Rasbora sp (Cyprinidae) From Beratan Lake, Bali. 17th World Lake Conference, Lake Kasumigaura, Ibaraki, Japan.

[19] Voris, H.K. 2000. Maps of pleistocene sea levels in southeast asia: Shorelines, river systems and time durations. *Biogeography*, 27: 1153-1167.

[20] Fitri, Maya. The Settlement Morphology Along Musi River: The Influence of River Characteristics. *Journal of Architecture and Built Environment*, Vol. 45, No. 2, December 2018, 133-140

[21] Syaifullah, Syaifullah., S. Luqyana., D. H. Tjong., I. J. Zakaria., D. I. Roesma. 2022. Haplotype network of three species of *Angguilla* (freshwater eels) in West Sumatera, Indonesia, based on Cytochrome b gene. *AACL Bioflux* Vol 15, 775-787.

[22] Nei, M dan S. Kumar. 2000. Moleculer Evolution and Phylogenetics. Oxford University Press. New York.

[23] Avise, J.1998. Phylogeography. Harvard University Press, Cambridge, Mass.