



Identification of Fish in Semilabeo Genus Using Morphological Taxonomy and Molecular Biology Methods

Tran Thi Thuy Ha^{1,*}, Nguyen Thi Thu Thuy¹, Ngo Sy Van¹

*Research Institute for Aquaculture, No 1. Dinh Bang- Tu Son - Bac Ninh, Viet Nam.

* Corresponding Author: Tel.: 84.241 3842383;
E-mail: thuyha@ria1.org

Received 02 May 2018
Accepted 29 July 2018

Abstract

Fish in Semilabeo genus has been listed in high risk of endangerment in the wild and some species in this genus show similar characteristics. A total of 30 fish in Semilabeo genus collected from Ha Giang, Phu Tho provinces and Research Institutes for Aquaculture No1 (RIA1), Viet Nam was classified using morphological taxonomy and molecular biology methods. The results revealed that the morphological analysis identified 2 species in the genus of Semilabeo, *S. obscurus* (Peters, 1881) and *Semilabeo obscurus* (Lin, 1881). The molecular method using COI sequences revealed highly similar results between Ha Giang and Phu Tho groups but were differ from the group collected in RIA1. Based on BLAST result, COI gene sequences of samples collected in Ha Giang and Phu Tho were homogenous with those of *Semilabeo notabilis* species 99-100%, whereas, RIA1 group showed high similarity with *Semilabeo obscurus*. Besides, genetic distance and Neighbor-Joining tree view using COI indicate the two species: 1) *Semilabeo notabilis* (Peters, 1881) and 2) *Semilabeo obscurus* (Lin, 1881). In this study, the two analysis using morphological taxonomy and molecular biology showed the same findings in term of species identification for fish in Semilabeo genus collected in some areas in Viet Nam.

Keywords: Species identification, COI, *Semilabeo obscurus*, *Semilabeo notabilis*.

Introduction

Fish in the Semilabeo genus family known as the special freshwater fish in the North of Viet Nam are of highly economic value and delicious. However, their production has declined significantly due to overexploitation in recent years. Many fish in this genus has been listed in high risk of endangerment in the wild (VU) (The Red book for animal in Viet Nam, 2007).

Up to now, a number of studies about identification by morphological description has been reported (Wang, 1998; Wu, Shao, & Lai, 1999). In Semilabeo genus, the studies on protein sequence analysis were investigated by Zheng, Yang, and Chen (2012). However, researches on identification by molecular biology methods have still been limited. In Vietnam, studies mainly focused on morphological description, classification and biological characteristics (Yen, 1983; Hao & Van, 1993; Bau, Tuan, Dang, & Thang, 1999).

Recently, methods using morphological taxonomy and molecular biology for identification of aquatic animals have usually been combined in order

to improve accuracy of the analysis and reliability. In these studies, molecular markers and useful softwares with high reliability have been widely applied (Ward, Hanner, & Hebert, 2009). Some genes located on mitochondrial DNA (mtDNA) have been known as DNA barcodes in term of identification such as Cytochrome c oxidase subunit 1 (COI) (Hebert, Cywinska, Ball, & deWaard, 2003), Cytochrome b (Sevilla *et al.*, 2007) and 16s rRNA (Vences, Thomas, Meijden, Chiari, & Vieites, 2005). Among them, COI is the most common use for identification of many animal species (Puckridge, Andreakis, Appleyard, & Ward, 2013).

In this study, samples of animal in Semilabeo genus were collected from Phu Tho and Ha Giang provinces and Research Institute for Aquaculture No1 (RIA1) - Viet Nam which one population being kept. Fish identification was done using two methods: morphological and molecular biology analysis. The results of this study will provide information for establishing a breeding program as well as contribute to conservation of genetic resources of this precious species.

Materials and Methods

Materials

Samples of the fish in *Semilabeo* genus were collected in Chay River belong to Que Lam town, Doan Hung district, Phu Tho province; Lo River belongs to Thanh Thuy town, Vi Xuyen district, Ha Giang province and from one population have been kept at RIA1 (Bac Ninh province). A total of 10 whole fish per population (>100 g/fish) were collected and then kept at -20 °C or stored in 10% formalin for morphological taxonomy. The caudal fins were preserved in 95% ethanol for molecular biology analysis.

Fish Identification Using Morphological Analysis

In addition to fish observation, morphology index (mm), body weight (g) and classification have been done according to the methods described by Rainboth (1996), Nguyen Van Hao *et al.* (1993), Dinh, *et al.* (2013).

Fish Identification Using Molecular Biology

DNA Extraction

Total DNA of 30 samples was extracted from fin clip using Deaasy Tissue Kit of Qiagen (Germany). Quantity and quality of the extracted DNA were examined using 0.8% agarose gel and Nanodrop 2000C (Thermo Scientific).

COI Gene Amplification Using PCR

To amplify COI gene, primers named MAB (MAB Fw và MAB Rw) described by Badhul, Azarudeen, Vignesh, Kumar, & Srinivasan (2012) were used on PCR Perkin-Elmer 2400. The sequences of the primers are: F (5'-3') TCAACCAACCACAAAGACATTGGCAC; R (5'-3'): TAGACTTCTGGGTGGCCAAAGAATCA.

The PCR reaction was amplified in total volume of 25 µl containing 3 µl template DNA (~ 100ng/ µl), 100 mM Tris HCl (pH 8.3), 500 mM KCl (pH 8.3), 2.5 µl MgCl (25mM), 1.0 µl dNTPs (5 mM), 0.5 µl forward and reverse primers (10pm/ µl each) and 1u/ µl Taq Polymerase. Distilled water was added up to total volume of 25 µl. Temperature cycle was: denaturation at 94 °C for 2 min; 35 cycles at 94 °C for 50s, 56 °C for 50s, 72 °C for 1 min, final extension at 72 °C for 10 min and hold at 10 °C.

COI Sequences

Agarose gel of 2% was used to check the quality and the yield of PCR products. Then, the PCR products were purified by Expin™ PCR SV Kit of GeneAll and sequenced by First BASE Laboratories,

Malaysia. The purified products were labelled by bigdye Terminator v3.1 Cycle Sequencing Kit in the mixture reaction of 10 µl: 4.94 µl distilled water, 1.94 µl BigDye buffer 5X (400 mM Tris-HCl pH 9.0 and 10 mM MgCl₂), 0.12 µl BigDye Terminator and 1 µl ExoSAP products. Analysis software of Genomelab system was used to create the sequencing files and reading continuous sequences.

COI Sequences Analysis

The sequences are checked by Finch TV 1.4.0 (<http://www.geospiza.com>), after that the program named ClustalW in BioEdit used to compare and align the sequence.

Phylogenetic Relationships, Genetic Distance and ID Identification

Phylogenetic trees were constructed using neighbor-joining method (NJ) and Kimura 2-Parameter distance model in MEGA 6 (Kimura, 1980). We used bootstrap analysis with 1000 replicates to evaluate support for phylogenetic relationships (Felsenstein, 1985). The genetic distances were analyzed using MEGA 6.0 software (Tamura, Stecher, Peterson, Filipinski, & Kumar, 2013). The identification was done based on the similarity in COI sequences between the samples and the references downloaded from GenBank database with the parameters of coverage and identity. The similarity of COI sequences was compared and analyzed using BLAST software.

Results and Discussion

Morphological Taxonomy Analysis

The results showed that 30 samples were grouped into 2 different species in *Semilabeo* genus. Fish samples collected in Phu Tho and Ha Giang were *Semilabeo notabilis* but fish collected from RIA1 line were *Semilabeo obscurus*. In general, morphological appearances of these fish in the 2 groups was very similar in some points: the body was elongate and flattened sides, the mouth was low and the lower lip was thick and wide with many small specks, the front lip was wide and the back was narrow and in triangles shape. The eyes were on the top of the head and located in the back half. The distance between two eyes was wide. The front bladder was round and the rear was sharp. However, the two species differed in the number of scales around the caudal fin, colors of the line along the body and dorsal fin. In *S. notabilis* species, the number of scales around the caudal fin was 16-18 and colors of the line along the body was very clear. On the other hand, in *S. obscurus* species, the number of scales around the caudal fin was 20-22 and no line along the body appeared and the line behind the dorsal fin was concave. According to some

researchers, environment is the main factor that affects the morphological differences between populations (Norton, Luczkovich, & Motta, 1995; Wainwright, 1996). These findings strongly agree with the results reported by Hao and Van (1993).

Molecular Biological Analysis

DNA Extraction and PCR Reaction

The result for checking DNA was presented on agarose gel and the bands were sharp and total DNA was good quality for PCR reaction (Fig not shown).

The PCR reaction was done in order to amplify the COI gen using optimal condition. PCR products showed the band of the targeted gene was clear on the gel (Figure 1). The size of the amplified products ranged from 600-700 bp that was in the range of the expected size reported by Badhul *et al.* (2012).

Analysis of COI Sequences

The analysis results of COI sequences showed the clear differences between sample sequences collected in Phu Tho and Ha Giang and the samples collected from RIA1 (Figure 2). The sequences of the first group were almost similar and the similarities level was very high. The two different sequence types were characterized by the change of nucleotides in the sequences in the same group (purines, A and G, or pyrimidine, T and C) or the change between the groups. This differences may be due to the replacement mutation occurred in the groups that can cause the phenotypic differences in fish collected from RIA1 comparing to those in the other groups.

BLAST results presented a high degree of similarity in the nucleotide sequences of the 30 fish samples in this study with those registered in Genbank NCBI (National Center for Biotechnology

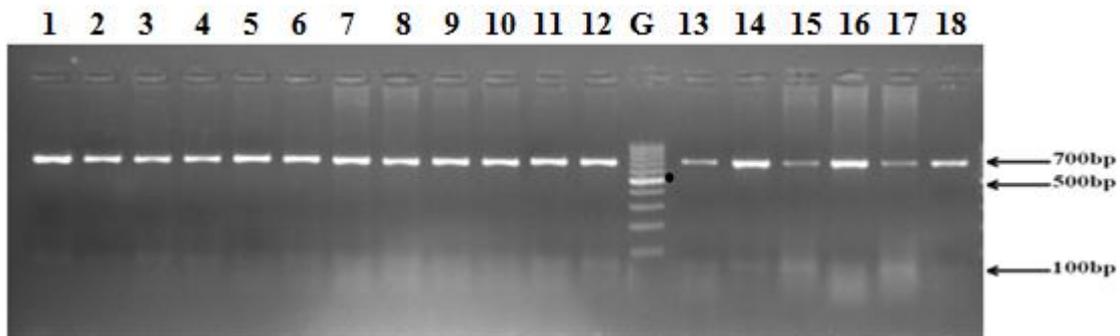


Figure 1. PCR products using MAB primers on 2% agarose gel.

Note: G: Marker (100bp).

Number 1 → 6: PCR product of fish collected in Phu Tho, from PT1 to PT6.

Number 7 → 12: PCR product of fish collected in Ha Giang, from HG1 to HG6.

Number 13 → 18: PCR product of fish collected in RIA1, from VTS1 to VTS6.

Vị trí nu	95	129	152	188	216	244	248	279	312	318	355	415	440	447	450	462	472	492	509	557	593	608	636	643	651	654	660	667	
Mẫu																													
PT1	C	C	C	C	A	G	G	T	C	T	A	T	A	T	C	C	T	C	T	G	C	A	C	T	A	G	T	C	
PT2
PT3
PT4
PT5
PT6
PT7
PT8
PT9
PT10
HG1
HG2
HG3
HG4
HG5
HG6
HG7
HG8
HG9
HG10
VTS1	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	
VTS2	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	
VTS3	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	
VTS4	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	
VTS5	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	
VTS6	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	
VTS7	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	
VTS8	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	
VTS9	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	
VTS10	T	T	T	T	G	A	A	C	T	C	G	C	G	C	T	T	C	T	C	A	T	G	T	C	G	A	C	T	

Figure 2. COI sequence 670 bp at high polymorphic position.

Information) (Table 1). The COI sequences of the samples collected in Ha Giang and Phu Tho provinces were homologous up to 99-100% with those of *Semilabeo notabilis*. Whereas, fish collected from RIA1 had high similarity in the nucleotide sequences with *Semilabeo obscurus*.

Genetic Distances

The analysis of the genetic distance using nucleotide sequence data of 12 samples (04 samples/group) was presented in Table 2. Overall, the genetic distance between three groups ranged from 0.000 to 0.042. This indicated that these fish likely belong to the same genus but was in different species. In addition, low genetic distance (≤ 0.002) between the two groups (Phu Tho and Ha Giang) demonstrates a closely genetic relationship. These results were similar to the classification results of fish in Pangasiidae family based on mtDNA of the Cytochrome b gene reported by Pouyaud, Teugels, Gustiano, & Legendre (2000) with genetic distance value ranged from 0.004 to 0.149. On the other hand, the genetic distance of fish collected from RIA1 was clearly different from those from Phu Tho and Ha Giang (0.040 - 0.042). The differences between individuals of the same species often vary from species to species but usually low ($\sim 1\%$). For

example, this value was 1.2% for India gobies (Viswambaran, et al., 2013), 0.3% for fish Vietnamese gobies (Thao, & Yen, 2015). Totally, it can be clearly that the genetic distance results proved that fish collected in Ha Giang and Phu Tho differed to fish collected from RIA1.

Neighbour-Joining Tree and Genetic Relationship

Neighbor-Joining tree was built using COI sequences of 30 samples belonged to two species in *Semilabeo* genus: *S. obscurus* (Lin, 1981) and *S. notabilis* (Peters, 1881) on the NCBI gene bank with number of GU086581.1 and JX074195.1, respectively. The results showed that 30 samples in this study were clustered into two groups belonging to two species from NCBI. The first cluster (I) includes 20 samples of two populations collected in Phu Tho and Ha Giang, code from PT1 to PT10 and HG1 to HG10 in line with *Semilabeo notabilis* (Peters, 1881) while the second cluster (II) was *Semilabeo obscurus* (Lin, 1981) with 10 samples collected from RIA1, coded from VTS1 to VTS10. Again, this result showed the genetic relationship between the fish from Phu Tho and Ha Giang was closer than

those from RIA1. In this study, methods using morphological and molecular biological analysis showed the same findings with two species:

Table 1. BLAST results from genbank NCBI

No	Code	Similary species	NCBI number	Ident. (%)
1	PT1	<i>Semilabeo notabilis</i>	JX074195.1	99
2	PT2	<i>Semilabeo notabilis</i>	JX074195.1	99
3	PT3	<i>Semilabeo notabilis</i>	JX074195.1	99
4	PT4	<i>Semilabeo notabilis</i>	JX074195.1	99
5	PT5	<i>Semilabeo notabilis</i>	JX074195.1	100
6	PT6	<i>Semilabeo notabilis</i>	JX074195.1	100
7	PT7	<i>Semilabeo notabilis</i>	JX074195.1	99
8	PT8	<i>Semilabeo notabilis</i>	JX074195.1	100
9	PT9	<i>Semilabeo notabilis</i>	JX074195.1	100
10	PT10	<i>Semilabeo notabilis</i>	JX074195.1	99
11	HG1	<i>Semilabeo notabilis</i>	JX074195.1	99
12	HG2	<i>Semilabeo notabilis</i>	JX074195.1	99
13	HG3	<i>Semilabeo notabilis</i>	JX074195.1	99
14	HG4	<i>Semilabeo notabilis</i>	JX074195.1	99
15	HG5	<i>Semilabeo notabilis</i>	JX074195.1	100
16	HG6	<i>Semilabeo notabilis</i>	JX074195.1	100
17	HG7	<i>Semilabeo notabilis</i>	JX074195.1	99
18	HG8	<i>Semilabeo notabilis</i>	JX074195.1	100
19	HG9	<i>Semilabeo notabilis</i>	JX074195.1	100
20	HG10	<i>Semilabeo notabilis</i>	JX074195.1	99
21	VTS1	<i>Semilabeo obscurus</i>	GU086581.1	100
22	VTS2	<i>Semilabeo obscurus</i>	GU086581.1	100
23	VTS3	<i>Semilabeo obscurus</i>	GU086581.1	99
24	VTS4	<i>Semilabeo obscurus</i>	GU086581.1	100
25	VTS5	<i>Semilabeo obscurus</i>	GU086581.1	99
26	VTS6	<i>Semilabeo obscurus</i>	GU086581.1	99
27	VTS7	<i>Semilabeo obscurus</i>	GU086581.1	100
28	VTS8	<i>Semilabeo obscurus</i>	GU086581.1	100
29	VTS9	<i>Semilabeo obscurus</i>	GU086581.1	100
30	VTS10	<i>Semilabeo obscurus</i>	GU086581.1	99

Table 2. Genetic distance of 12 samples in three groups

Code	PT1	PT2	PT3	PT4	HG1	HG2	HG3	HG4	VTS1	VTS2	VTS3	VTS4
PT1	0,000											
PT2	0,000	0,000										
PT3	0,002	0,002	0,000									
PT4	0,000	0,000	0,000	0,000								
HG1	0,002	0,002	0,002	0,002	0,000							
HG2	0,002	0,002	0,002	0,002	0,000	0,000						
HG3	0,002	0,002	0,002	0,002	0,000	0,000	0,000					
HG4	0,002	0,002	0,002	0,002	0,000	0,002	0,000	0,000				
VTS1	0,040	0,040	0,040	0,040	0,040	0,040	0,042	0,040	0,000			
VTS2	0,040	0,042	0,042	0,042	0,042	0,042	0,042	0,042	0,000	0,000		
VTS3	0,040	0,040	0,040	0,040	0,040	0,040	0,040	0,040	0,002	0,000	0,000	
VTS4	0,042	0,042	0,042	0,042	0,042	0,042	0,042	0,042	0,002	0,002	0,000	0,000

Semilabeo notabilis (Peters, 1881) and *Semilabeo obscurus* (Lin, 1981).

Conclusions

In conclusion, the results of the present work showed that 20 fish samples collected in Ha Giang, Phu Tho are *Semilabeo notabilis* (Peters, 1881) and the other 10 fish collected from RIA1 are *Semilabeo obscurus* (Lin, 1881) using the morphological taxonomy method. The same findings were found in molecular methods using COI gene sequence analysis. The results can support different studies of these species in the future.

References

- Badhul, H.M.A., Azarudeen, M.D., Vignesh, R., Kumar, A.T.T. & Srinivasan, M. (2012). Identification and intra species delineation of ornamental silver pompano *Trachinotus blochii* (Lacépède, 1801) with DNA barcodes. *International Research Journal of Biochemistry and Bioinformatics*, 3, 26-36. <http://www.interestjournals.org/IRJBB>.
- Bau, P., Tuan, N.D., Dang, B.D., & Thang, N.C. (1999). Some biology characteristics of *Semilabeo obscurus* Lin, 1881. *Proceedings published by Research Institute for Aquaculture, No1*. TU.43.6.
- Dinh, T.K., Shibukawa, K., Phuong, N.T., Hung, H.P., Loi, T.X., Hieu, M.V., & Utsugi, K. (2013). Description of fish identification in Mekong Delta river in Viet Nam. Can Tho university publisher.
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39, 783–791. <https://dx.doi.org/10.1111/j.1558-5646.1985.tb00420.x>
- Hao, N.V., & Van, N.S. (1993). Freshwater fish in Viet Nam. T42.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., & deWaard J.R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270, 313–321. <https://www.ncbi.nlm.nih.gov/pubmed/12614582>
- Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111–120. <https://www.ncbi.nlm.nih.gov/pubmed/7463489>
- Norton, S.F., Luczkovich, J.J., & Motta, P.J. (1995). The role of ecomorphological studies in the comparative biology of fishes. *Environment. Biology. Fish.*, 44, 287–304. <https://link.springer.com/article/10.1007/BF00005921>
- Pouyaud, L., Teugels, G.G., Gustiano, R., & Legendre, M. (2000). Contribution to the phylogeny of pangasiid catfishes based on allozymes and mitochondrial DNA. *Journal of Fish Biology*, 56, 1509–1538. <https://dx.doi.org/10.1111/j.1095-8649.2000.tb02161.x>
- Puckridge, M., Andreakis, N., Appleyard, S.A., & Ward, R.D. (2013). Cryptic diversity in flathead fishes (Scorpaeniformes: Platycephalidae) across the Indo-West Pacific uncovered by DNA barcoding. *Molecular Ecology Resources*, 13, 32–42. <https://dx.doi.org/10.1111/1755-0998.12022> <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1095-8649.2000.tb02161.x>
- Rainboth, W.J. (1996). Fishes of the Cambodian Mekong. FAO, 1996. <http://www.fao.org/docrep/010/v8731e/v8731e00.htm>
- Sevilla, R.G., Diez, A., NorÉN, M., Mouchel, O., JÉRÔME, M., Verrez-Bagnis, V., Van Pelt, H., Favre-Krey, L., Krey, G., Consortium, T.F., & Bautista, J.M. (2007). Primers and polymerase chain reaction conditions for DNA barcoding teleost fish based on the mitochondrial cytochrome b and nuclear rhodopsin genes. *Molecular Ecology Notes*, 7, 730–734. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1471-8286.2007.01863.x>
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, 30, 2725–2729. <https://dx.doi.org/10.1093/molbev/mst197>
- Thao, N.P., & Yen, D.T. (2015). Comparison of morphology and DNA barcoding for two species *Butis butis* and *Butis humeralis*. *Scientific journal of Can Tho University, Vietnam*, 40, 23–30. <https://sj.ctu.edu.vn/ql/docgia/download/baibao-17521/04-CN-NGUYEN%20The%20Red%20book%20for%20animal%20in%20Viet%20Nam,2007,pp26>
- Vences, M., Thomas, M., van der Meijden, A., Chiari, Y., Vieites, D. (2005). Comparative performance of the

- 16S rRNA gene in DNA barcoding of amphibians. *Frontiers in Zoology*, 2, 5. <https://dx.doi.org/10.1186/1742-9994-2-5>.
- Viswambharan, D., Pavan-Kumar, A., Singh, P.D., Jaiswar, A.K., Chakraborty, S.K., Nair, J.R., & Lakra, W.S. (2013). DNA barcoding of gobiid fishes (Perciformes, Gobioidae). *Mitochondrial DNA*, 1–5. <https://dx.doi.org/10.3109/19401736.2013.834438>.
- Ward, R.D., Hanner, R., & Hebert, P.D.N. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology*, 74, 329–356. <https://dx.doi.org/10.1111/j.1095-8649.2008.02080.x>
- Wainwright, P.C. (1996). Ecological explanation through functional morphology: the feeding biology of sunfishes. *Ecology*, 77, 1336–1343. <https://onlinelibrary.wiley.com/doi/pdf/10.2307/2265531>
- Wang, T. (1998). China red data book of endangered animals. Beijing: Science press, 1998. ISBN: 7–03–006401. <https://www.nhbs.com/series/china-red-data-book-of-endangered-animals>
- Ward, R.D., Zemplak, T.S., Innes, B.H., Last, P.R., & Hebert, B.D.N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 360, 1847–1857. <https://dx.doi.org/10.1098/rstb.2005.1716>
- Wu, H.L., Shao, K.T., & Lai, C.F. (eds.) (1999). Latin-Chinese dictionary of fishes names. The Sueichan Press, Taiwan. http://www.hceis.com/home/book_view.aspx?id=14175&ISBN=9787567015
- Yen, M.D. (1983). The economic freshwater fish in Viet Nam. T 263. Identification for freshwater fish in the North VietNam (1978).
- Zheng, T., Yang, D., & Chen, R. (2012). Sequences analysis of protein interphotoreceptor retinoid-binding protein at locus AFK65716, *Dong Wu Xue Bao journal*.