Phylogenetic Relationships of Freshwater Fish in Vietnamese Mekong

Vu Dang Ha Quyen[1,2], Thai Thi Lan Phuong[2], Truong Thi Oanh[2], Tran Linh Thuoc[1], Dang Thuy Binh[2]

***Abstract*—**The Mekong River Basin represents a global hotspot of aquatic biodiversity second only to the Amazon River in terms of total fish species richness. Our study focuses on phylogeny of freshwater fish in Vietmamese Mekong. Freshwater fish species were sampling at 7 Provinces along Hau and Tien rivers. Morphologically, 11 species have been identified. Phylogenetic trees were constructed based on 16S and CO1 gene of mitochondrial DNA using Maximum Parimony, Maximum Lilikehood and Baeysian Inference approaches. The 16S phylogram performed similar phylogenetic relationships of CO1 phylogeny, excluding difference in position of *Coilia* species (family Engraulidae) and *Acantopsis* (family Cobotidae). Our results corroborate the monophyletic status at genus level (includes: *Pangasius*, *Ompok*, *Acantopis*, *Trichopodus*, *Glossogobius*, *Coilia*, *Acantopsis* and *Cynoglossus*). Moreover, sequence differences have been found between *G. circumspectus* and *G. olivaceu*in in C01, while they are identical in 16S gene. Current data can be used as data sources for the study of biodiversity and management of fisheries resources in the Mekong Delta in Vietnam.

***Keywords*—**freshwater fish, Mekong River, Phylogenetic relationship

I. Introduction

THE Mekong Delta is a major agricultural and fisheries production zone in Vietnam. Increasing human demand for natural resources, particularly land for agriculture and aquaculture, has significantly reduced the extent of natural and semi-natural habitats in the delta [1]. Additionally, this biodiversity is under pressure from overexploitation, dam building, and environmental pollution.

DNA barcoding is a new technique in ecology and biosystematics of fish to understand genetic variation and clarify species identification where morphology alone is insufficient [2]. It has much to offer to fisheries managers, especially in the provision of tools enabling unequivocal specimen identification and assessment of stock structure [3]. Presently, the barcode analysis is a cost-effective option for species identification in some situations and this will increasingly be the case as reference libraries are assembled and analytical protocols are simplified [4]. The method promises fast and accurate species identifications by focusing analysis on a short standardized segment of the genome [2].

DNA barcodes have been obtained for more than 8000 species of fish and the CO1 sequences deposited in the Barcode of Life Data Systems (BOLD) online workbench and repository [5], [6] built up the barcode for fish in Lake Laut Tawar, Indonesia. Jumawan et al. (2011) applied of the mitochondrial CO1gene (cytochrome c oxidase subunit I) to delineate between two species of exotic suckermouth sailfin catfish *P. terygoplichthys – P. pardalis* and *P. disjunctivus* and their intergrades, which dominate the ichthyofauna of the Marikina River system, Philippines [7]. Zhang (2011) tested the efficacy of DNA barcoding in marine fishes of China [8].

Currently, about 312 freshwater fish (186 species) and brackish water fish [9] are known from the Mekong Delta; of which 252 species were reported only in Vinh Long Province [10]. However, none of those species have been barcoded.

The aim of this study is to build up phylogenetic relationships of freshwater fish Vietnamese Mekong based on morphologic and genetic characters. Freshwater fish Barcode Database is the prime access point for DNA signature sequences together with information on morphological taxonomic characters of freshwater fish. This unique combination of freshwater fish biological information and molecular markers will provide an updated species checklist for Vietnam; these databases are crucial to create input data for natural resources conservation.

II. Materials And Methods

A. Fish sampling

Freshwater fish species were collected at 7 provinces (Can Tho, An Giang, Dong Thap, Vinh Long, Ben Tre, Tra Vinh, Tien Giang) along Hau and Tien rivers based on random sampling methods in the field. Fish specimens are rinsed with freshwater and labeled. The samples were transferred to the laboratory keeping in alcohol or on ice and then store at -40^{0}C before analysis.

For morphological analysis, all specimens were identified based on taxonomic characters such as body and fins colour, the presence or absence of scales on the cheek, number of spine and soft rays belong to dorsal, anal, ventral, pectoral and

Vu Dang Ha Quyen[1,2], is with PhD student in Faculty of University of Science, (VNUHCM), 227 Nguyen Van Cu Str., Dist. 5, Ho Chi Minh City, VIETNAM, on work in Institute for Biotechnology and Environment, Nha Trang University, 02 Nguyen Dinh Chieu Str, Nha Trang City, VIETNAM (corresponding author's phone: +84-914161416; e-mail: quyenvdh.ntu@gmail.com).

Dang Thuy Binh[2], Thai Thi Lan Phuong[2], Trương Thi Oanh[2], work at Institute for Biotechnology and Environment, Nha Trang University, 02 Nguyen Dinh Chieu Str, Nha Trang City, VIETNAM (e-mail: binhdangthuy@gmail.com, lanphuong52sh.ntu@gmail.com; gerberalovely.nt@gmail.com).

Tran Linh Thuoc[1] is with professor in Department of Molecular & Environmental Biotechnology, Molecular Biotechnology Laboratory, University of Science (VNUHCM), 227 Nguyen Van Cu Str., Dist. 5, Ho Chi Minh City, VIETNAM (email: tlthuoc@hcmus.edu.vn)

caudal fin. Fish species was identified following Allen, (1985); Carpenter and Allen, (1989) [11], [12].

B. Molecular analyses

Total DNA was extracted from the muscle of each fish individual using GeneJET Genomic DNA Purification Kit (Thermo scientific) following the manufacturer's instructions. The extractions were used to amplify the 16S of mitochondrial DNA (mtDNA) using the primers 16Sar 5'-CGCCTGTTTATCAAAAACAT-3'and 16Sbr 5'-CCGGTCTGAACTCAGATCACGT- 3') [13], and CO1 mtDNA using the primers HCO 5' – TAAACTTCGGGTGACCAAA – 3', LCO 5' – GGTCAACAAATCATAAAGATA – 3' [14].

PCR reactions were performed with total volume of 50 µl including 20ng DNA template, 5 µl 10X Dream Taq buffer (Fermentat), 0,25 nM each dNTP, 0,2 pM each primer, 1 unit of Taq polymerase (5U/1 µl) and distilled water to the final volume. Biorad thermocyclers (Icycler) were used under the following temperature program: initial denaturation 94^0C for 3 min, followed by 35 cycles of denaturation 94^0C for 30s, annealing for 30s (for 16S, CO1 gene at 48^0C, 42^0C, respectively), and final extension at 72^0C for 5 minutes. PCR products were electrophoresed on 1,5% agarose gel stained with ethidium bromide, and bands were visualized under a UV transilluminator.

1 – 2 µl PCR products were purified using a PCR clean up system kit (Promega), and pre-sequenced using dye – labels dideoxy terminator (Big Dye Terminator v. 3.1, Applied Biosystems) with the same primer as the PCR reaction at the following temperatures: 96^0C for 30s, 50^0C for 30s and 60^0C for 4 min. Products were analyzed using an ABI Prism 3.700 DNA Analyzer (Applied Biosystems).

C. Phylogenetic analysis

Sequence contigs were assembled using Geneious ver 7 (http://www.geneious.com/). The resulting sequences were confirmed by the Basic Logical Alignment Search Tool (BLAST, http://blast.ncbi.nlm.nih.gov/). Sequences were initially aligned by eye using the sequence editor BioEdit 7.0 [15].

16S and CO1 sequences of freshwater fish in this study, together with sequences available from GenBank (Table 1), were used in the phylogenetic analysis. Data were analyzed using three approaches, i.e, Neighbour joining (NJ), maximum parsimony (MP) and Bayesian inference (BI). NJ analyses were conducted from MEGA 6 under 1000 replicate. MP analysis were conducted using PAUP* 4.0 [16]. Bootstrap support values were computed from 1,000 replicates randomized 10 times with tree-bisection-reconnection (TBR) addition sequence to assess the robustness of the findings.

Prior to BI analyses, best-fit models of nucleotide substitution were selected by the Akaike Information Criterion as implemented by and MrModeltest 2.2 [17]. Bayesian analyses were conducted in MrBayes 3.1.2 under the selected best-fit models and parameters. Numbers at the interior branches of the majority-rule consensus tree present posterior probability (PP). Tree display and editing were performed in TreeView 1.6.6 [18].

TABLE I
GENBANK ACCESSION NUMBERS FOR 16S AND CO1 MTDNA SEQUENCINGS

Family/Subfamily	Taxon	GenBank Accession Numbers	
		CO1	16S
Outgroups			
Petromyzontidae	*Eudontomyzon mariae*		EU404076.1
	Eudontomyzon morii		KM267718.1
	Anguillicoloides crassus	JF805718.1	
Siluriformes			
Pangasiidae	*Pangasius* sp.		
	Pangasius krempfi		
	Pangasius macronema		
	Pangasius krempfi		HM355773.1
	Pangasius macronema	KC627283.1	HM355776.1
	Pangasius sanitwongsei	EU752152.1	HM355780.1
	Pangasius larnaudii	JX997836.1	HM355775.1
	Pangasius pangasius	JF781175.1	GQ411088.1
	Pangasius nasutus	EF609426.1	HM355778.1
	Pangasius conchophilus	JF292426.1	HM355772.1
	Pangasius bocourti		HM355770.1
Siluridae	*Ompok bimaculatus*	JX260923.1	GQ469642.1
	Ompok pabda	FJ229987.1	GQ469568.1
	Ompok bimaculatus		
Perciformes			
Osphronemidae			
	Trichopodus pectoralis	HQ682726.1	AY763712.1
	Trichopodus trichopterus	KC789556.1	AY763713.1
	Trichopodus microlepis		AY763711.1
	Trichopodus microlepis		
	Trichopodus trichopterus		
	Trichopodus pectoralis		
Gobiidae			
	Glossogobius circumspectus	JX536695.1	JX536695.1
	Glossogobius olivaceus	JQ001860.1	JQ001860.1
	Glossogobius circumspectus		
Pleuronectiformes			
Cynoglosiidae	*Cynoglossus cynoglossus*	JX983282.1	AY359669.1
	Cynoglossus lighti	HQ711865.1	DQ112683.1
	Cynoglossus purpureomaculatus	Q738571.1	DQ112680.1
	Cynoglossus interruptus	JF952714.1	JQ939061.1
	Cynoglossus robustus	HM180553.1	HQ003913.1
	Cynoglossus joyneri	KF979127.1	HQ003908.1
	Cynoglossus feldmani		
Clupeiformes			
Engraulidae	*Coilia nasus*	HM180536.1	AM911208.1
	Coilia rebentischii		
Cypriniformes			
Cobitidae	*Acantopsis choirorhynchos*	JN177219.1	AB242161.1
	Acantopsis sp.		

III. RESULTS AND DISCUSSION

A. Phylogenetic relationships of freshwater fish based on the 16S gene of the mtDNA

The 16S dataset consists of more than 600 bp, of which 499 bp were unambiguously aligned. In total, 31 sequences (including 11 from present study, and 25 sequences from Genbank) were used to build phylogenetic tree; *Eudontomyzon mariae* and *Eudontomyzon morii* were used as outgroup. Tree topology from the NJ method (Fig.1) was identical to that of the MP and BI tree.

The phylogram was divided into two main clades with highly bootstrap support (>80%). The **clade I** was subdivided into two separated subgroup: **subGroup1.1** include 25 species belong to 5 Families (Pangasiidae, Siluridae, Cobotidae, Osphronemidae, Gobiidae) of the 3 orders Suliriformes, Cypriniformes and Perciformes.). At the genus level, all species showed monophyly with high BT support. Two species (*Coilia nasus and Coilia rebentischii)* of the Family Engrailidae, order Clupeiformes were clustered in the **subGroup 1.2**. The **Clade II** includes 7 *Cynoglossus* species (Cynoglossidae, Pleuronectiformes).

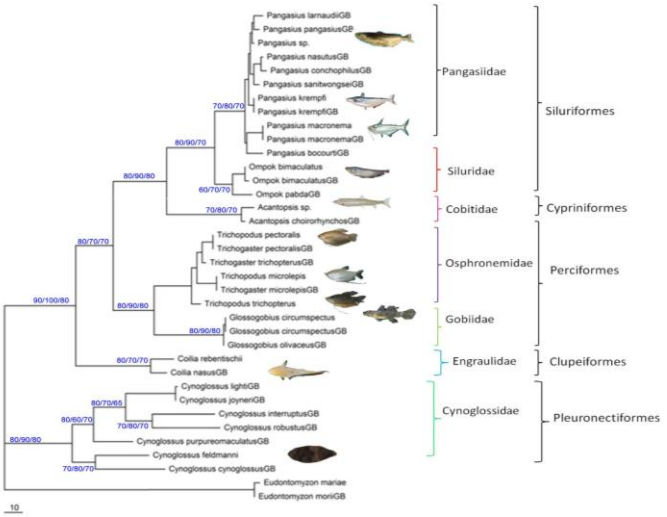

Fig. 1 Phylogenetic tree of freshwater fishes in Vietnamese Mekong based on the 16S mtDNA. Bootstrap value from MP, NJ analysis, posterior probability (BI analysis) along the branch. *Eudontomyzon mariae* and *Eudontomyzon morii* were used as outgroup.

B. Phylogenetic relationships of freshwater fish based on the CO1 gene of the mtDNA

The CO1 dataset consists of more than 600 bp, of which 555 bp were unambiguously aligned. Tree topology was similar from three different methods.

Phylogram (Fig.2) was divided into two main clades. In **clade 1**, s**ubroup1.1** include 22 species; showing difference in position of *Coilia* species (family Engraulidae). These species were sister clade to *Pangansius* and *Ompok*, while in 16S tree, They were separated cluster. **Subgroup 1.2** contained two species of *Acantopsis* (family Cobotidae), that are sister clade to P*angansisus* and *Ompok* in 16S tree,

The **clade II** performed similar phylogenetic relationships of 16S phylogeny as possessing 7 species of *Cynoclossus*.

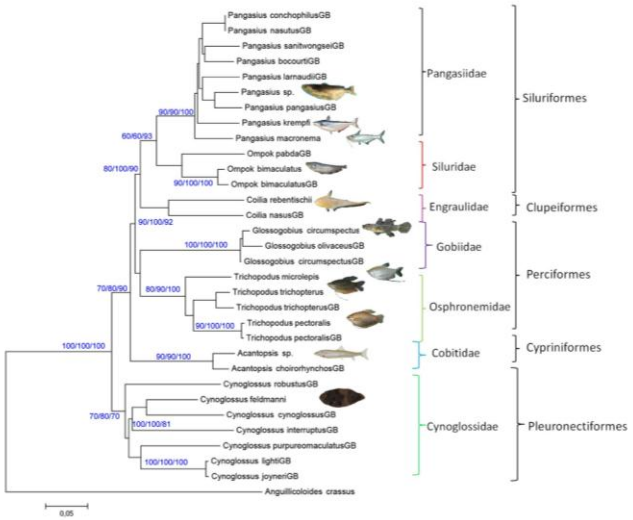

Fig. 2 Phylogenetic tree of freshwater fishes in Vietnamese Mekong based on the COI gene of the mtDNA. MP bootstrap value, posterior probability (BI analysis) and bootstrap value from NJ analysis along the branch). *Anguillicoloides crassus* was used as outgroup

The current research constructed phylogeny of economic and ecological important freshwater fish in Mekong Vietnam. Tree topology from 16S and CO1 mtDNA strongly support monophyly at genus level, while also confirm family and order of fish species (Fig 1 and 2).

In 16S phylogeny, (**Clade 1, subgroup 1.1**), *Pangasius macronema* showed close relationship with *P. bocourti*GB; and *Pangasius* sp. has sister clade to *P. larnaudii*GB and *P. pangasius*GB. This result is consistent with phylogenetic research donned by Pouyaud et al. (1998) [19]. Present phylogeny of *Pangasius* species was also identical to the phylogram analyzed in Thailand [20], in which *Pangasius krempfi* and *Pangasius* sp. are sister groups of *P. macronema*, *P. nasutus*, *P. conchophilus*, *P. larnaudii* and *P. polyuranodon*. Karinthanyakit and Jondeuung (2012) [20] also detected closed relationship between *Pangasius krempfi* and *P. nasutus*, *P. conchophilus*; while *P. sanitwongsei* close to *P. larnaudii*GB.

In the **subgroup 1.1 (clade 2**), the two species of the genus *Coilia* (*Coilia nasus*GB and *Coilia rebentischii*) appeared as a monophyletic group. The monophyly of the family Engraulidae had already been proposed by several authors [21], [22], [8]. However, they show limited population genetic structure of *C. nasus*. This can be explained by their ecotypes have high genetic diversity on mitochondrial markers [22].

Monophyly of *Trichopodus, Acantpsi, and Cynoglossus* (**clade 2)** were reported by Ruber et al. (2006) [23], Mayden (2009) [24], and Yokogawa (2008) [25]. These relationships are strongly supported by current study. Species of genus *Cynoglossus* were characterized by flat body, demersal fish; commonly occur in estuarine environment, some strict to freshwater. Morphological characters, as well as living mode is concrete evidence for the separation of these taxa.

For *Glossogobus* species, 16S sequences of *G. circumspectus* (from present study and Genbank) and *G. olivaceus*GB were identical, while the CO1 sequence

differences are 2.3 and 3.1, respectively. Further study need to be conducted to confirm taxonomic position of these species.

As 16S and CO1 phylogeny displayed almost similar pattern. The difference is only the position of *Coilia* and *Acantopsis*. They both appear as separate clade due to different genes applied. More species of this genus need to be added, and genomic DNA markers should be considered for future analysis.

Overall, the taxonomy and systematic of freshwater fishes in Vietnamese Mekong are fragmented and uncompleted resolved. The current study has provided a robust attempt to reconstruct the phylogeny of freshwater fish taxa in Vietnam using the molecular approach. In addition, the species richness of the fauna should be properly documented for appropriate management and conservation. Meanwhile, more molecular studies should be undertaken to examine larger datasets in order to obtain a more comprehensive understanding of their systematic relationships.

IV. Conclusion

The current research constructed phylogeny of economic and ecological important freshwater fish in Mekong Vietnam. Tree topology from 16S and CO1 mtDNA displayed almost similar pattern. The difference is only the position of *Coilia* and *Acantopsis*. The present analyses corroborated the monophyletic status of 8 genus that includes: *Pangasius, Ompok, Acantopis, Trichopodus, Glossogobius, Coilia, Acantopsis and Cynoglossus*. These data can be used as data sources for the study of biodiversity and management of fisheries resources in the Mekong Delta in Vietnam.

Acknowledgment

This study was partly supported by a project “Conservation Genetics for Improved Biodiversity and Resource Management in a Changing Mekong Delta”. We also thank Prof. Kent Carpenter (Old Dominion University) for assistance in identity of fish species.

References

[1] I. C. Cambell, “Biodiversity of the Mekong Delta,” in *The Mekong Delta System*, 2012, pp. 293–313.

[2] P. D. N. Hebert, A. Cywinska, S. L. Ball, and J. R. DeWaard, “Biological identifications through DNA barcodes.,” *Proc. Biol. Sci.*, vol. 270, no. 1512, pp. 313–321, Feb. 2003. http://dx.doi.org/10.1098/rspb.2002.2218

[3] B. R. Ward, “Declivity in steelhead (Oncorhynchus mykiss) recruitment at the Keogh River over the past decade,” *Can. J. Fish. Aquat. Sci.*, vol. 57, no. 2, pp. 298–306, 2000 http://dx.doi.org/10.1139/f99-243.

[4] M. Hajibabaei, J. R. DeWaard, N. V Ivanova, S. Ratnasingham, R. T. Dooh, S. L. Kirk, P. M. Mackie, and P. D. N. Hebert, “Critical factors for assembling a high volume of DNA barcodes.,” *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 360, no. 1462, pp. 1959–1967, Oct. 2005 http://dx.doi.org/10.1098/rstb.2005.1727.

[5] S. RATNASINGHAM and P. D. N. HEBERT, “bold: The Barcode of Life Data System (http://www.barcodinglife.org),” *Mol. Ecol. Notes*, vol. 7, no. 3, pp. 355–364, May 2007 http://dx.doi.org/10.1111/j.1471-8286.2007.01678.x.

[6] Z. A. Muchlisin, Z. Thomy, N. Fadli, M. A. Sarong, and M. N. Siti-Azizah, “Barcoding of freshwater fishes from Lake Laut Tawar, Indonesia,” *ACTA Ichthyol. Piscat.*, vol. 43, no. 1, pp. 21–29, 2013 http://dx.doi.org/10.3750/AIP2013.43.1.04.

[7] J. C. Jumawan, B. M. Vallejo, A. A. Herrera, C. C. Buerano, and I. K. C. Fontanilla, “DNA barcodes of the suckermouth sailfin catfish Pterygoplichthys (Siluriformes: Loricariidae) in the Marikina River system, Philippines: Molecular perspective of an invasive alien fish species,” *Philipp. Sci. Lett.*, vol. 4, no. 2, 2011.

[8] J. Zhang, “Species identification of marine fishes in china with DNA barcoding.,” *Evid. Based. Complement. Alternat. Med.*, vol. 2011, p. 978253, Jan. 2011 http://dx.doi.org/10.1155/2011/978253.

[9] Tran Khac Dinh, S. Koichi, Nguyen Thanh Phương, Ha Khac Hung, Tran Xuan Loi, Mai Van Hieu, and U. Kenzo, “Identified and classification fish in Vietnamese Mekong Delta based on morphology,” *CanTho University Publiser*, 2013.

[10] Nguyen Thanh Tung, Nguyen Nguyen Du, and Lam Ngoc Chau, “Composition of aquatic species in the rivers of Vinh Long Province,” Department of Agriculture and Rural Development, 165, 2005.

[11] G. R. Allen, “Snappers of the World.An annotated and illustrated catalogue of Lutjanid species known to date,” *FAO Fish. Synopsis*, vol. 6, no. 125, 1985.

[12] K. E. Carpenter and G. R. Allen, “Emperor fishes and large-eye breams of the world (family lethrinidae). An annotated and illustrated catalogue of lethrinid species known to date,” *FAO Fish. Synopsis*, vol. 9, no. 125, 1989.

[13] S. Palumbi, S. Romano, W. O. Mcmillan, and G. Grabowski, “THE SIMPLE FOOL ’ S GUIDE TO PCR,” vol. 96822, no. 808, pp. 1–45, 2002.

[14] O. Folmer, M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek, “DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates.,” *Mol. Mar. Biol. Biotechnol.*, vol. 3, no. 5, pp. 294–299, Oct. 1994.

[15] T. A. Hall, “BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT,” *Nucleic Acids Symp. Ser.*, vol. 41, pp. 95–98, 1999.

[16] D. Swofford, “PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.” Sinauer Associates, Sunderland, Massachussets, 2002.

[17] J. A. A. Nylander, “MrModeltest,” *Evol. Biol. Centre, Uppsala Univ.*, 2004.

[18] P. R. D., “TreeView: an application to display phylogenetic trees on personal computers,” *Bioinformatics/computer Appl. Biosci.*, vol. 12, no. 4, pp. 357–358, 1996.

[19] P. Relationships, A. Pangasiid, C. Species, N. E. W. I. On, E. I. R. Zoogeography, L. Pouyaud, R. Gustiano, M. Legendre, C. A. Project, I. Penelitian, P. Air, J. Ragunan, P. Minggu, M. Cedex, I. Penelitian, P. Tawar, J. Ragunan, and P. Minggu, “Phylogenetic relationships among pangasiid catfish species (Siluriformes, Pangasiidae) and new insights on their zoogeography,” no. 1, 1998.

[20] W. Karinthanyakit and a. Jondeung, “Molecular phylogenetic relationships of pangasiid and schilbid catfishes in Thailand,” *J. Fish Biol.*, vol. 80, no. 7, pp. 2549–2570, 2012 http://dx.doi.org/10.1111/j.1095-8649.2012.03303.x.

[21] D. Liu, H. Y. Guo, W. Q. Tang, and J. Q. Yang, “Comparative evolution of S7 intron 1 and ribosomal internal transcribed spacer in coilia nasus (Clupeiformes: Engraulidae),” *Int. J. Mol. Sci.*, vol. 13, no. 3, pp. 3085–3100, 2012 http://dx.doi.org/10.3390/ijms13033085.

[22] Q. Yang, Z. Han, D. Sun, S. Xie, L. Lin, and T. Gao, “Genetics and phylogeny of genus Coilia in China based on AFLP markers,” *Chinese J. Oceanol. Limnol.*, vol. 28, no. 4, pp. 795–801, 2010 http://dx.doi.org/10.1007/s00343-010-9093-3.

[23] L. Rüber, R. Britz, and R. Zardoya, “Molecular Phylogenetics and Evolutionary Diversification of Labyrinth Fishes (Perciformes: Anabantoidei),” *Syst. Biol.* , vol. 55 , no. 3 , pp. 374–397, Jun. 2006 http://dx.doi.org/10.1080/10635150500541664.

[24] R. L. Mayden, W. J. Chen, H. L. Bart, M. H. Doosey, A. M. Simons, K. L. Tang, R. M. Wood, M. K. Agnew, L. Yang, M. V. Hirt, M. D. Clements, K. Saitoh, T. Sado, M. Miya, and M. Nishida, “Reconstructing the phylogenetic relationships of the earth’s most diverse clade of freshwater fishes-order Cypriniformes (Actinopterygii: Ostariophysi): A case study using multiple nuclear loci and the mitochondrial genome,” *Mol. Phylogenet. Evol.*, vol. 51, no. 3, pp. 500–514, 2009. http://dx.doi.org/10.1016/j.ympev.2008.12.015

[25] K. Yokogawa, “yokogawa_2008.pdf.” Fish Genetic and Breeding Science, pp. 113–121, 2008.