Journal of Molecular Biomarkers & Diagnosis
ISSN: 2155-9929

Tran et al., J Mol Biomark Diagn 2017, 8:6
DOI: 10.4172/2155-9929.1000363

Research Article Open Access

DNA Barcoding and Phylogenetic Relationships of Nine Catfish Species from Mekong Basin, Vietnam

Ha Thi Thuy Tran[1]*, Thi Nga Tran[1], Hang Nguyen Ai Tran[2] and Huong Thi Nguyen[1]

[1]*Research Institute for Aquaculture No.1, Dinh Bang, Tu Son, Bac Ninh, Vietnam*
[2]*Research Institute for Aquaculture No.2, Nguyen Dinh Chieu, Ho Chi Minh, Vietnam*

Abstract

Fishes that belong to the family *Pangasiidae* are widely recognized to have good potential for aquaculture and are highly valued as flesh food in the markets of Vietnam. However, there is much debate on the identification and phylogeny of the available species of *Pangasiidae* in the Asia. In the present study, nine species of two genera (*Pangasianodon* and *Pangasius*) of *Pangasiidae* were investigated using two partial sequences of the COI and cytochrome b mitochondrial genes to differentiate among them and study their phylogenetic relationships. A total of 42 haplotypes were identified (21 haplotssypes of each gene). The highest interspecies genetic distance was between *Pangasius larnaudii* and *Pangasius bocourti* (0.189) for COI and was between *Pangasius macronema* and *Pangasianodon hypophthalmus* (0.179) for cyt b. Whereas the lowest genetic distance was between *Pangasius macronema* and *Pangasius conchophilus* for both genes (0.65 for COI and 0.92 for cyt b, respectively). The phylogenetic tree analyses of two genes showed two major clusters that are genetically distant from the two genera. The results obtained in this study also show that beside COI gene, the cyt b gene region can be successfully used for differentiating between species and accepted as a standard region for DNA barcoding.

Keywords: DNA barcoding; *Pangasiidae* family; *Pangasianodon; Pangasius*; Coi; Cyt b

Introduction

The Vietnamese Delta forms an integral part of the Lower Mekong Basin (LMB). LMB, originating in the Tibetan Plateau, is supplied with rich alluvial deposits from the Mekong River, which has a high biodiversity of fishes [1]. The water environments of the delta region are numerous, such as large freshwater rivers, irrigation canals, brackish estuaries, mangrove creeks, and mudflats [1]. Thus, fishery resources play an important role in the economy of south Vietnam. Many types of large- and medium-sized fishes, especially those that belong to the family *Pangasiidae*, have a good potential for aquaculture and highly valued as flesh food in the markets of Vietnam, for example, the sutchi catfish (*Pangasianodon hypophthalmus*), basa catfish (*Pangasius bocourti*), *ca bong lao* (*Pangasius krempfi*), and spot pangasius (*Pangasius larnaudii*). In 2016, the production of *Pangasinodon hypophthalmus* was 1.2 million tons, and it was exported to 136 countries across all continents, with an estimated export income of US$ 1.67 billion [2].

The family *Pangasiidae* includes a large number of species and belongs to the order Siluriformes, and it has a relatively wide distribution from Southwest to Southeast Asia. According to Roberts and Vidthayanon (1991), 11 species of the family *Pangasiidae* are found in Thailand, 10 in Indonesia, 3 in Peninsular Malaysia, and 4 endemics to Borneo Island. Most of the species in the family *Pangasiidae* are freshwater fishes. Five species also occur in brackish water: *Pangasius pangasius* [3], *Pangasius krempfi* [4], *Pangasius kunyit* [5], *Pangasius sabahensis* [6] and *Pangasius mekongensis* [6]. In Vietnam, according to some researchers, the family *Pangasiidae* has 13 species that belong to 4 genera: *Pangasianodon*, *Pangasius*, *Pseudolais*, and *Helicophagus* [7,8]. Most of the species are distributed in freshwater environments; only three of them inhabit the brackish waters at estuaries: *Pangasius krempfi, Pangasius mekongensis*, and *Pangasius elongatus.*

Although, many studies on the classification of *Pangasiidae* according to external morphological characteristics have been conducted, there is still much debate on this issue because the species in the family have very similar morphologies, resulting in difficulties in classification in recent years. The website www.fishbase.org provides information on the morphology of fishes, but it is not always accurate. In many cases, the morphological classification is limited by the influence of living conditions or processed products. Therefore, identification becomes difficult or even impossible. Recently, with developments in molecular biology, several markers have been used as effective tools for the identification of fish species in particular and many other animal species in general. Especially, over the last decade, DNA barcoding has emerged as a molecular method for species identification. DNA barcoding is based on the principle of sequencing a short segment of DNA from a uniform region of the mitochondrial genome of the target specimen and comparing these unknown barcodes to an existing barcode database to identify the species [9].

DNA barcoding is also used to refine species identification by detecting query specimens with probabilistic algorithms when a set of barcodes of known species is established. Information on the phylogeny of the family *Pangasiidae* is scarce [10], and some species of the family have similar morphological features, such as *Pangasius bocourti* and *Pangasius nasutus*. Rainboth (1996) stated that *Pangasius bocourti* and *Pangasius nasutus* have been misidentified. Gustiano et al. in 2003 used biometric measurements to distinguish between seven species of four genera in four main rivers in Sumatra; however, the classification was solely based on morphology without any molecular evidence. Generally, cytochrome c oxidase I (COI) or cytochrome b (cyt b) sequence is used for DNA barcoding [11,12]. The cyt b gene has

***Corresponding author:** Ha Thi Thuy Tran, Research Institute for Aquaculture No.1, Dinh Bang, Tu Son, Bac Ninh, Vietnam, Tel: +84 0241 3842383; Fax: +84 04 38273070; E-mail: thuyha@ria1.org

Received August 20, 2017; **Accepted** August 24, 2017; **Published** August 26, 2017

Citation: Tran HTT, Tran TN, Tran HNA, Nguyen HT (2017) DNA Barcoding and Phylogenetic Relationships of Nine Catfish Species from Mekong Basin, Vietnam. J Mol Biomark Diagn 8: 363. doi: 10.4172/2155-9929.1000363

been considered one of the most useful genes for phylogenetic studies, and it is probably the best-known mitochondrial gene with respect to the structure and function of its protein product [13]. The cyt b gene contains both slowly and rapidly evolving codon positions as well as more conservative and variable regions or domains overall. It is a powerful indicator for identifying species with DNA analysis techniques [14-16], and it is also used in molecular evolution studies [17]. In this study, we used DNA barcoding for the identification of nine species of *Pangasiidae* from Mekong Delta, Vietnam. The DNA barcoding data obtained in this study will be used for better monitoring, conservation, and management of the family *Pangasiidae*.

Materials and Methods

Sample collection and DNA isolation

A total of 50 individuals (9 species) of *Pangasiidae* were obtained from a local fish market or fishermen from the provinces in Mekong Delta, such as An Giang, Đong Thap, Can Tho, Soc Trang, Tien Giang, and Ben Tre. The fishes were caught in rivers by using lines and nets or in aquaculture facilities (cages or ponds). The specimens were identified according to the descriptions published by Roberts and Vidthayanon and [18], and approximately 1 to 3 g of fin clips or muscle tissue was collected from each specimen and stored in 95% ethanol until further use. A summary of the statistics and sampling localities of some species of the family *Pangasiidae* is presented in Table 1.

The genomic DNA was isolated from the fin clips or muscle tissues by using the standard phenol-chloroform-isoamyl-alcohol method described by Sambrook and Russell [19] with some minor modifications. The DNA quality was checked using 1% agarose gel electrophoresis, and the absorbance at 260 nm was measured using the Ultrospec 2100 Pro UV/visible spectrophotometer to determine the DNA concentration. Fragments were gelpurified and sequenced using an ABI 3730XL DNA Sequencer (Applied Biosystems).

PCR amplification and DNA sequencing

The COI sequences were amplified using primers COI-Fish-F- (5′-CGACTAATCATAAAGATATCGGCAC-3′) and COI-Fish-R- (5′-TTCAGGGTGACCGAAGAATCAGAA-3′) [20], and cyt b DNA was amplified using primers H15149AD-(5′-AAAAACCACCGTTGTTATTCAACTA-3′) and L14735-(5′-GCICCTCARAATGAYATTTGTCCTCA-3′) [21]. The reaction volume contained 100 ng of genomic DNA, 1 U of *Taq* polymerase (TaKaRa, Otsu, Shiga, Japan), 1 × *Taq* PCR buffer, 2 mM $MgCl_2$, 0.2 mM dNTPs, 10 pmol of forward and reverse primers, and double-distilled water to a final volume of 50 µl. The PCR cycle used was as follows: denaturation at 94°C for 3 min, followed by 35 cycles of 94°C for 30 s, 52°C for 30 s (cyt b) or 56°C for 30 s (COI), 72°C for 45 s, and a final extension at 72°C for 10 min. The PCR products were separated using 2% agarose gel electrophoresis in TAE buffer and visualized using a UV illuminator. The fragments were sequenced using the ABI 3730XL DNA Sequencer (Applied Biosystems). Assembly and comparison of the transcripts with *de novo* sequences ensured transcript identity and sequence accuracy.

Genetic diversity analysis and phylogenetic tree construction

The sequence data were edited manually to confirm all base-pair assignments from chromatographs by using BIOEDIT software [22]. All the COI and cyt b sequences of the mtDNA regions were homologous in length and could be easily aligned using Clustal X. For sequence comparisons, pairwise genetic distances were quantified on the basis of the Kimura 2-parameter (K2P) distance model [23] by Hall using MEGA 6.06 [24].

Sample identification based on the COI sequence similarity approach was conducted using two databases: BOLD and GenBank. The highest percent pairwise identity of the consensus sequence from each species searched (BLASTN) in NCBI was compared to the percent specimen similarity scores of the consensus sequence from each species in the BOLD Identification System (BOLD-IDS) [25]. We searched for the consensus sequences in the NCBI database and downloaded one sequence for each species for the phylogenetic tree. For the COI gene: KR080263.1 (*Pangasianodon hypophthalmus*), EF609426.1 (*Pangasius conchophilus*), KY118577.1 (*Pangasius larnaudii*), KT289877.1 (*Pangasius krempfi*), JF292428.1 (*Pangasius bocourti*), KT289890.1 (*Pangasius macronema*), KT289880.1 (*Pangasius mekongensis*), KU568952.1 (*Pangasius sanitwongsei*), and KY118586.1 (*Pangasianodon gigas*). For the cyt b gene: KC009653.1 (*Pangasianodon hypophthalmus*), HM236385.1 (*Pangasius conchophilus*), KC993142.1 (*Pangasius larnaudii*), HM236390.1 (*Pangasius krempfi*), GQ856793.1 (*Pangasius bocourti*), DQ119443.1 (*Pangasius macronema*), KT289880.1 (*Pangasius mekongensis*), KC993147.1 (*Pangasius sanitwongsei*), and

No.	Scientific name	Sampling localities (province)	Number of individual	Number of Haplotype	Haplotype name	Genbank Accession no.
1	*Pangasius krempfi*	Ben Tre, SocTrang	5	3 (2)	H9-H11 (H1-H2)	KY398025-27 (KY451455-56)
2	*Pangasius bocourti*	An Giang, ĐongThap, TienGiang, Can Tho	5	2 (1)	H3-H4 (H3)	KY398019-20 (KY451457)
3	*Pangasius macronema*	An Giang, ĐongThap, TienGiang, Can Tho	10	4 (4)	H7-H8, H13-H14 (H4-H5, H21-H22)	KY398023-24 KY398029-30 (KY451458-59, KY451474-75)
4	*Pangasius conchophilus*	An Giang, ĐongThap, TienGiang, Can Tho	5	2 (3)	H5-H6 (H6-H8)	KY398021-22 (KY451460-62)
5	*Pangasianodon hypophthalmus*	An Giang, ĐongThap, Can Tho	5	2 (3)	H1-H2 (H9-H11)	KY398017-18 (KY451463-65)
6	*Pangasius mekongensis*	SocTrang, Can Gio	5	3 (2)	H15-H17 (H12-H13)	KY398031-33 (KY451466-67)
7	*Pangasianodon gigas*	An Giang, NhaTrang University	5	2 (1)	H20-H21 (H14)	KY398036-37 (KY451468)
8	*Pangasius sanitwongsei*	National Breeding Centre for Southern Freshwater Aquaculture, An Giang	5	2 (4)	H18-H19 (H15-H18)	KY398034-35 (KY451469-72)
9	*Pangasius larnaudii*	An Giang, ĐongThap, Can Tho	5	1 (1)	H12 (H19)	KY398028 (KY451473)
	Total		50	21 (21)	--	--
Values given are for COI with the cyt b values in parentheses						

Table 1: Summary of statistics, sampling localities of some species in Pangasiidae family using COI and Cyt b gene in Mekong Delta, Vietnam.

GQ856795.1 (*Pangasianodon gigas*). The COI and cyt b sequences of *Clarias batrachus* (GenBank: JF292297 and KR007706) were obtained from the NCBI database; *C. batrachus* (family Clariidae) was used as the outgroup and chosen to represent the close relationship between *Clariidae* and *Pangasiidae* [26]. The neighbor-joining (NJ) method was used to infer the phylogenetic relationships among the species by using MEGA version 6.06 [24]. The NJ tree was constructed, and support for monophyly was assessed with 1000 bootstrap pseudo-replicates [27].

Results

Species identification and genetic distance

Two mtDNA regions in all the samples were successfully amplified using PCR; 21 haplotypes of COI and 21 haplotypes of cyt b were investigated from nine species, and all the sequences were submitted to the NCBI database with accession numbers KY398017–KY398037 (COI) and KY451455–KY451455 (cyt b) (Table 1). Table 2 shows the comprehensive barcoding identification results for the COI and cyt b genes by using the GenBank and BOLD databases. The results for the COI gene in both databases revealed definitive identity matches in the range of 98% to 100% for the consensus sequences of six species (*Pangasius macronema, Pangasius conchophilus, Pangasianodon hypophthalmus, Pangasianodon gigas, Pangasius sanitwongsei,* and *Pangasius larnaudii*). GenBank-based identification for all the species yielded an alignment E-value of 0.0. The BOLD-IDS results were consistent with the GenBank results with respect to the identification of these species and yielded 100% identity, except for *Pangasianodon gigas* (100% maximum identity in GenBank, whereas the percent similarity in BOLD was 99.64%). The present study also highlighted that the GenBank databases lack the data record for the cyt b gene of *Pangasius mekongensis* and provided a top hit for a related species, *Pangasius* (92% identity).

Pairwise nucleotide and genetic distances (p-distance) using K2P are represented in Table 3 (for COI) and Table 4 (for cyt b). For the COI gene, the highest interspecies genetic distance (0.189) was observed between *Pangasius larnaudii* and *Pangasius bocourti*, and the lowest genetic distance (0.065) was between *Pangasius macronema* and *Pangasius conchophilus*. For the cyt b gene, the lowest genetic distance (0.092) was also between *Pangasius macronema* and *Pangasius conchophilus*; however, the highest interspecies genetic distance (0.179) was observed between *Pangasius macronema and Pangasianodon hypophthalmus*.

Species studied	COI gene				Cyt b gene	
	BOLD-IDS		GenBank (BLASTN)		GenBank (BLASTN)	
	Species identification	% similarity	Species identification	% Max identity	Species identification	% Max identity
Pangasius krempfi	No match	0	*P. krempfi*	100	*P. krempfi*	99
Pangasius bocourti	No match	0	*P. bocourti*	100	*P. bocourti*	100
Pangasius macronema	*P. macronema*	100	*P. macronema*	100	*P. macronema*	100
Pangasius conchophilus	*P. conchophilus*	100	*P. conchophilus*	100	*P. conchophilus*	100
Pangasianodon hypophthalmus	*P. hypophthalmus*	100	*P. hypophthalmus*	100	*P. hypophthalmus*	100
Pangasius mekongensis	No match	0	*P. mekongensis*	100	*P. pangasius*	92
Pangasianodon gigas	*P. gigas*	99.64	*P. gigas*	99	*P. gigas*	99
Pangasius sanitwongsei	*P. sanitwongsei*	100	*P. sanitwongsei*	100	*P. sanitwongsei*	98
Pangasius larnaudii	*P. larnaudii*	100	*P. larnaudii*	100	*P. larnaudii*	100

Table 2: Summary of identification based on each species consensus sequence of two markers using BOLD Identification System (BOLD-IDS) and BLASTN search from GenBank.

	PBO1	PBO2	PCO1	PCO2	PGI1	PGI2	PHY1	PHY2	PKR1	PKR2	PKR3	PLA1	PMA1	PMA2	PMA3	PMA4	PME1	PME2	PME3	PSA1	PSA2
PBO1	0	0.064	0.072	0.074	0.119	0.126	0.116	0.116	0.096	0.096	0.094	0.113	0.088	0.086	0.090	0.088	0.098	0.101	0.098	0.079	0.079
PBO2	26	0	0.143	0.145	0.195	0.192	0.192	0.192	0.170	0.170	0.167	0.189	0.161	0.158	0.163	0.161	0.172	0.175	0.172	0.150	0.150
PCO1	39	64	0	0.002	0.112	0.119	0.119	0.119	0.104	0.104	0.101	0.083	0.069	0.067	0.076	0.074	0.074	0.076	0.074	0.072	0.072
PCO2	40	38	1	0	0.112	0.119	0.119	0.119	0.101	0.101	0.099	0.086	0.067	0.065	0.074	0.072	0.076	0.079	0.076	0.074	0.074
PGI1	61	82	61	60	0	0.006	0.092	0.092	0.132	0.132	0.135	0.137	0.132	0.129	0.134	0.132	0.134	0.132	0.134	0.116	0.116
PGI2	55	54	57	57	1	0	0.099	0.099	0.139	0.139	0.142	0.144	0.139	0.136	0.141	0.139	0.142	0.139	0.142	0.123	0.123
PHY1	60	55	63	63	58	52	0	0.000	0.119	0.119	0.122	0.137	0.126	0.124	0.128	0.126	0.132	0.134	0.132	0.134	0.134
PHY2	58	55	60	60	57	52	3	0	0.119	0.119	0.122	0.137	0.126	0.124	0.128	0.126	0.132	0.134	0.132	0.134	0.134
PKR1	50	47	49	48	65	62	59	59	0	0.000	0.002	0.101	0.106	0.103	0.098	0.096	0.091	0.094	0.091	0.101	0.101
PKR2	50	47	49	48	65	62	59	59	1	0	0.002	0.101	0.106	0.103	0.098	0.096	0.091	0.094	0.091	0.101	0.101
PKR3	47	70	50	48	67	59	58	58	53	1	0	0.103	0.103	0.101	0.095	0.093	0.094	0.096	0.094	0.098	0.098
PLA1	53	50	44	45	70	65	70	67	63	53	52	0	0.103	0.100	0.105	0.103	0.093	0.095	0.093	0.108	0.108
PMA1	52	43	44	42	72	63	66	71	62	63	56	55	0	0.002	0.014	0.012	0.118	0.120	0.118	0.100	0.100
PMA2	51	67	43	41	72	62	65	70	52	62	55	54	1	0	0.012	0.010	0.120	0.123	0.120	0.098	0.098
PMA3	51	70	42	39	72	59	59	64	52	52	57	54	8	7	0	0.002	0.120	0.123	0.120	0.103	0.103
PMA4	51	43	45	44	71	63	66	71	59	59	50	55	6	5	1	0	0.118	0.120	0.118	0.100	0.100
PME1	49	69	43	43	69	65	71	68	50	50	51	50	68	70	64	67	0	0.002	0.000	0.105	0.105
PME2	49	44	43	44	68	64	72	69	51	51	48	51	67	68	61	67	1	0	0.002	0.108	0.108
PME3	47	43	42	43	69	65	68	68	49	49	46	50	65	66	59	65	1	1	0	0.105	0.105
PSA1	43	37	41	40	62	54	66	69	49	49	51	58	58	58	57	58	58	57	53	0	0.000
PSA2	40	37	40	41	63	57	70	73	49	49	47	59	60	59	53	60	60	61	57	1	0

PBO: *Pangasius bocourti*; PCO: *Pangasius conchophilus*; PGI: *Pangasianodon gigas*; PHY: *Pangasianodon hypophthalmus*; PKR: *Pangasius krempfi*; PLA: *Pangasius larnaudii*; PMA: *Pangasius macronema*; PME: *Pangasius mekongensis*; PSA: *Pangasius sanitwongse*; 1: Haplotype 1; 2: Haplotype 2; 3: Haplotype 3; 4: Haplotype 4.

Table 3: Pairwise nucleotide differences (below diagonal) and genetic distances (p-distance) (above diagonal) in COI sequences of nine *Pangasiidae* species.

	PBO1	PCO1	PCO2	PCO3	PGI1	PHY1	PHY2	PHY3	PKR1	PKR2	PLA1	PMA1	PMA2	PMA3	PMA4	PME1	PME2	PSA1	PSA2	PSA3	PSA4
PBO1	0	0.131	0.116	0.120	0.134	0.156	0.156	0.160	0.100	0.103	0.124	0.110	0.110	0.113	0.117	0.107	0.103	0.120	0.120	0.120	0.124
PCO1	44	0	0.011	0.009	0.141	0.140	0.140	0.144	0.110	0.106	0.138	0.106	0.106	0.109	0.113	0.142	0.138	0.099	0.113	0.113	0.109
PCO2	40	4	0	0.003	0.130	0.144	0.144	0.147	0.110	0.106	0.124	0.092	0.092	0.095	0.099	0.127	0.123	0.092	0.106	0.106	0.102
PCO3	41	3	1	0	0.133	0.140	0.144	0.140	0.106	0.103	0.127	0.095	0.095	0.099	0.102	0.131	0.127	0.089	0.102	0.102	0.099
PGI1	42	45	41	46	0	0.106	0.113	0.117	0.142	0.131	0.117	0.123	0.123	0.127	0.131	0.138	0.134	0.134	0.134	0.134	0.131
PHY1	49	50	47	48	34	0	0.006	0.009	0.156	0.152	0.130	0.171	0.171	0.175	0.179	0.141	0.137	0.112	0.126	0.126	0.123
PHY2	51	50	49	50	36	4	0	0.003	0.156	0.152	0.138	0.163	0.163	0.167	0.171	0.148	0.145	0.112	0.126	0.126	0.123
PHY3	51	47	48	47	37	5	5	0	0.160	0.156	0.141	0.167	0.167	0.171	0.175	0.152	0.148	0.116	0.130	0.130	0.126
PKR1	33	38	37	37	44	49	51	50	0	0.009	0.114	0.096	0.096	0.100	0.103	0.128	0.132	0.100	0.114	0.107	0.103
PKR2	33	36	36	35	40	47	49	48	3	0	0.118	0.086	0.086	0.089	0.093	0.132	0.135	0.096	0.103	0.096	0.093
PLA1	40	46	41	43	37	41	45	44	36	37	0	0.139	0.139	0.143	0.139	0.110	0.107	0.100	0.107	0.107	0.103
PMA1	33	34	30	32	43	52	52	53	32	27	44	0	0.000	0.003	0.006	0.131	0.128	0.110	0.110	0.103	0.099
PMA2	33	33	29	31	42	52	52	53	32	27	44	1	0	0.003	0.006	0.131	0.128	0.110	0.110	0.103	0.099
PMA3	38	38	34	36	43	56	56	55	34	29	46	1	1	0	0.003	0.135	0.131	0.113	0.113	0.106	0.103
PMA4	43	39	35	36	43	57	57	54	35	30	45	3	3	2	0	0.139	0.135	0.117	0.117	0.110	0.106
PME1	34	45	40	43	48	45	47	48	41	41	36	45	44	41	43	0	0.003	0.162	0.162	0.162	0.158
PME2	33	44	39	42	48	44	46	47	42	42	35	43	44	40	42	2	0	0.158	0.158	0.158	0.154
PSA1	39	34	32	31	42	37	39	37	33	30	32	35	37	37	40	49	48	0	0.018	0.011	0.009
PSA2	35	36	34	33	40	36	39	38	31	31	30	34	36	36	37	47	46	4	0	0.009	0.011
PSA3	44	41	36	42	42	41	44	41	36	30	34	36	38	39	36	50	49	6	2	0	0.003
PSA4	40	40	35	39	43	40	44	42	35	29	33	34	36	37	40	49	48	5	3	1	0

PBO: *Pangasius bocourti*; PCO: *Pangasius conchophilus*; PGI: *Pangasianodon gigas*; PHY: *Pangasianodon hypophthalmus*; PKR: *Pangasius krempfi*; PLA: *Pangasius larnaudii*; PMA: *Pangasius macronema*; PME: *Pangasius mekongensis*; PSA: *Pangasius sanitwongse*; 1: Haplotype 1; 2: Haplotype 2; 3: Haplotype 3; 4: Haplotype 4.

Table 4: Pairwise nucleotide differences (below diagonal) and genetic distances (p-distance) (above diagonal) in Cyt b sequences of nine *Pangasiidae* species.

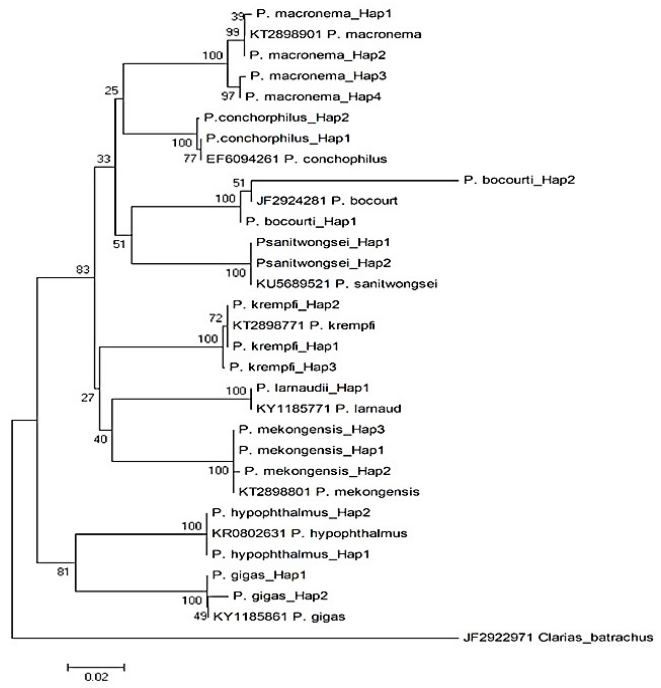

Figure 1: Neighbor-joining tree constructed on the basis of COI gene sequences showing phylogentic relationship among nine *Pangasiidae* species. A *Clarias batrachus* used as out-group species.

Phylogenetic relationships among the *Pangasiidae* species: The phylogenetic trees (NJ (Figures 1 and 2), MP, and ML) for the two genes revealed almost identical phylogenetic relationships among the species. Two major clusters were revealed: the first cluster formed by *Pangasianodon hypophthalmus* and *Pangasianodon gigas*, and the second cluster further divided into two subclades in all the constructed trees in which *Pangasius macronema, Pangasius conchophilus, Pangasius bocourti,* and *Pangasius sanitwongsei* were placed in the first subclade and *Pangasius krempfi, Pangasius larnaudii,* and *Pangasius mekongensis* formed the second subclade. All the trees were produced using phylogeny reconstruction analysis, with 1000 bootstrap replicates.

Discussion

Morphological studies of specimens raise questions regarding observed features versus described features. In a few cases, key morphological characteristics are difficult to discern [28]. The DNA

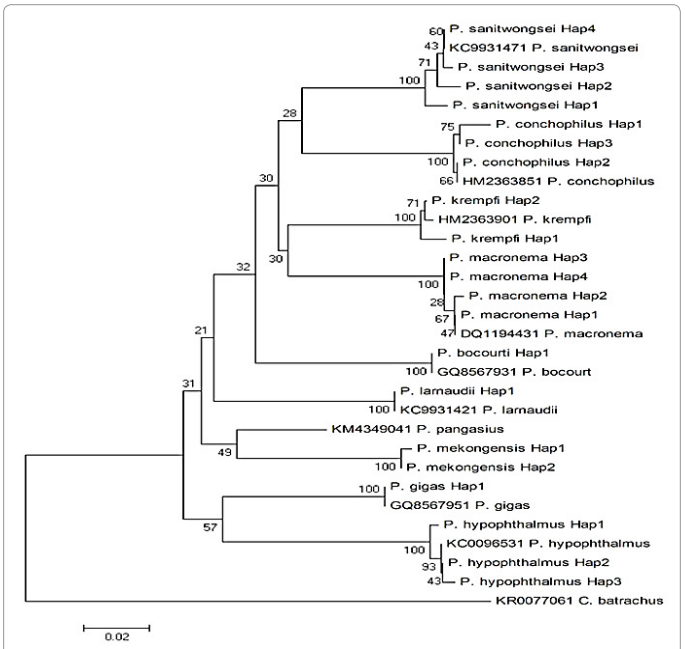

Figure 2: Neighbor-joining tree constructed on the basis of cyt b gene sequences showing phylogentic relationship among nine *Pangasiidae* species. A *Clarias batrachus* used as out-group species.

barcoding approach has resolved some identification issues and elucidated the actual species composition in certain regions [29]. In this study, we sequenced the COI region and cyt b gene of mtDNA to create a set of barcode sequences and identify nine catfish species belonging to two genera (*Pangasianodon* and *Pangasius*) in Vietnam. We extensively compared our results to the BOLD and GenBank databases. We found that out of the nine-species studied, only six matched the reference sequences in both databases. The remaining catfish have not yet been barcoded in the BOLD database. However, when we used BLAST to search for the cyt b gene in the NCBI databases, eight species (except *Pangasius mekongensis*) matched the reference sequences. Our *Pangasius mekongensis* COI sequences used as queries in GenBank yielded only one result: *Pangasius mekongensis* (100% identity), followed by *Pangasius pangasius* (94% to 95% identity). The mislabeling is not unexpected since both *Pangasius pangasius* and *Pangasius mekongensis* are genetically homologous [30] and morphologically similar. The adults of both species have yellow caudal fins [6, 31]. BOLD-IDS validates the identification search only if the species in the reference database has at least three barcoded specimens and identifies the query sequences if they match the reference sequence within a conspecific distance of less than 1% [25]. Therefore, correct species labeling, morphological taxonomy, and voucher documentation should be prioritized in cases where reassessment of spurious data is necessary [20].

In this study, we also aimed to understand the potential of COI and cyt b genes as DNA barcoding tools for identifying almost all the species of the two genera (*Pangasianodon* and *Pangasius*) in Vietnam. A total of 42 haplotypes were identified, and 21 haplotypes per gene were studied. For the nine-studied species, the interspecies distances were greater than 0.02 for both genes and ranged from 0.065 to 0.189 for COI and 0.092 to 0.179 for cyt b). No intraspecies and/or interspecies distance overlaps were detected, and a distinct barcoding gap was found between intraspecies and interspecies distances in each species. These results indicate that the COI and cyt b gene sequences can be effectively used to identify the nine species of the two genera with DNA barcoding. The observed transition vs. transversion ratios in pangasids are also comparable to those in many teleosts [32,20]. Generally, a larger excess of transitions related to transversion is observed in teleost mtDNA [20].

The phylogenetic trees for the nine species belonging to *Pangasiidae* in Vietnam on the basis of the COI and cyt b genes are shown in Figures 1 and 2. The phylogenetic relationships based on the COI sequence are concordant with the morphological and osteological comparisons made by [1,8]. Our results are consistent with those of Azlina et al. in 2013, who studied the phylogenetic relationships of *Pangasiidae* in Malaysia.

Conclusion

In conclusion, DNA barcoding is emerging as an invaluable tool for species identification. The gene sequences of COI and cyt b have been submitted directly to GenBank. The results obtained in this study show that the cyt b gene region can be successfully used for differentiating between species and accepted as a standard region for DNA barcoding. The phylogenetic trees represented two major clusters that are genetically distant from the two genera.

References

1. Tran DD, Shibukawa K, Nguyen PT, Ha HP, Tran LX, et al. (2013) Fishes of the Mekong Delta, Vietnam.
2. https://tongcucthuysan.gov.vn/en-us/News/-Tin-v%E1%BA%AFn/doc-tin/006583/2016-12-15/tong-ket-san-xuat-tieu-thu-ca-tra-nam-2016-va-ban-giai-phap-phat-trien-ben-vung%20ed
3. Hamilton F (1822) An account of the fishes found in the river Ganges and its branches. Printed for A. Constable and company.
4. Hogan Z, Baird IG, Radtke R, Vander Zanden MJ (2007) Long distance migration and marine habitation in the tropical Asian catfish, Pangasius krempfi. J Fish Biol 71: 818-832.
5. Pouyaud L, Teugels GG, Legendre M (1999) Description of a new pangasiid catfish from South-East Asia (Siluriformes). Cybium 23: 247-258.
6. Gustiano R, Teugels GG, Pouyaud L (2003) Revision of the *Pangasius kunyit* catfish complex, with description of two new species from South-East Asia (*Siluriformes; Pangasiidae*). J Nat Hist 37: 357-376.
7. Khoa TT, Huong TTT (1993) Identification of freshwater fish in the Mekong Delta. Can Tho University.
8. Mai DY (1992) Identification of southern freshwater fish species. Publisher of Scientific and Technical, UK.
9. Hebert PD, Cywinska A, Ball SL (2003) Biological identifications through DNA barcodes. Proceedings of the Proc R Soc Lond B Biol Sci 270: 313-321.
10. So N, Van Houdt JK, Volckaert FA (2006) Genetic diversity and population history of the migratory catfishes *Pangasianodon hypophthalmus* and *Pangasius bocourti* in the Cambodian Mekong River. Fish Sci 72: 469-476.
11. Wong LL, Peatman E, Lu J, Kucuktas H, He S, et al. (2011) DNA barcoding of catfish: Species authentication and phylogenetic assessment. PLoS One 6: e17812.
12. Yacoub HA, Fathi MM, Mahmoud WM (2013) DNA barcode through cytochrome b gene information of mtDNA in native chicken strains. Mitochondrial DNA 24: 528-537.
13. Degli Esposti M, De Vries S, Crimi M, Ghelli A, Patarnello T, et al. (1993) Mitochondrial cytochrome b: Evolution and structure of the protein. Biochimica et Biophysica Acta (BBA)-Bioenergetics 1143: 243-271.
14. Hsieh HM, Chiang HL, Tsai LC, Lai SY, Huang NE, et al. (2001) Cytochrome b gene for species identification of the conservation animals. Forensic Sci Int 122: 7-18.
15. Kappel K, Haase I, Käppel C, Sotelo CG, Schröder U (2017) Species identification in mixed tuna samples with next-generation sequencing targeting two short cytochrome b gene fragments. Food Chem 234: 212-219.
16. Zehner R, Zimmermann S, Mebs D (1998) RFLP and sequence analysis of the cytochrome b gene of selected animals and man: Methodology and forensic application. Int J Legal Med 111: 323-327.
17. Armani A, Castigliego L, Tinacci L, Gianfaldoni D, Guidi A (2011) Molecular characterization of icefish, (*Salangidae* family), using direct sequencing of mitochondrial cytochrome b gene. Food Control 22: 888-895.
18. Rainboth WJ (1996) Fishes of the cambodian mekong. Food and Agriculture Organisation, Rome, Italy.
19. Sambrook J, Russell DW (2001) Molecular cloning: A laboratory manual. third. Cold pring Harbor Laboratory Press, New York, USA.
20. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD (2005) DNA barcoding Australia's fish species. Philos Trans R Soc Lond B Biol Sci 360: 1847-1857.
21. Kocher TD, Lee WJ, Sobolewska H, Penman D, McAndrew B (1998) A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus)*. Genetics 148: 1225-1232.
22. Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: Nucleic acids symposium series Jan 41: 95-98.
23. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120.
24. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30: 2725-2729.
25. Ratnasingham S, Hebert PD (2007) BOLD: The Barcode of Life Data System. Mol Ecol Resour 7: 355-364.
26. http://tolweb.org/Siluriformes/15065
27. Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783-791.
28. Bhattacharjee MJ, Laskar BA, Dhar B, Ghosh SK (2012) Identification and re-evaluation of freshwater catfishes through DNA barcoding. PloS one 7: e49950.

29. Zhang YH, Qin G, Zhang HX, Wang X, Lin Q (2017) DNA barcoding reflects the diversity and variety of brooding traits of fish species in the family *Syngnathidae* along China's coast. Fish Res 185: 137-144.

30. Quyen VD, Phuong TT, Oanh TT, Thuoc TL, Binh DT (2015) Phylogenetic relationships of freshwater fish in Vietnamese Mekong. International Conference on Biological, Environment and Food Engineering (BEFE-2015), Singapore.

31. Roberts TR, Vidthayanon C (1991) Systematic revision of the Asian catfish family *Pangasiidae*, with biological observations and descriptions of three new species. Proc Acad Nat Sci Philadelphia, USA. pp. 97-143.

32. Barman AS, Singh M, Singh RK, Sarkar T, Lal KK (2014) RETRACTED: Molecular identification and phylogeny of Channa species from Indo-Myanmar biodiversity hotspots using mitochondrial COI gene sequences. Biochem Syst Ecol 5.