

Unravelling taxonomic ambiguity of the Mastacembelidae in the Mekong Delta (Viet Nam) through DNA barcoding and morphological approaches

Thuy-Yen Duong^{1*}, Liem Van Dung Tran¹, Ngoc-Tran Thi Nguyen¹,
Jamsari Amirul Firdaus Jamaluddin² and Mohd Nor Siti Azizah³

¹College of Aquaculture and Fisheries, Can Tho University, Viet Nam; ²School of Biological Sciences, Universiti Sains Malaysia, Malaysia; ³Institute of Marine Biotechnology, Universiti Malaysia Terengganu, Malaysia

Received for publication: 18 June 2020; Accepted for publication: 1 October 2020.

Abstract: Morphological-based species identification can be problematic for a comparative worldwide survey if taxonomic keys are limited and inconsistent, as illustrated in the family Mastacembelidae. This study combined DNA barcoding and morphological methods to test species identification of Mastacembelidae in the Mekong Delta with emphasis on taxonomic ambiguity of the precise identification of the fish locally known as chach bong. Fish specimens were collected from fishermen in different regions of the delta. Five presumed species within two genera were recorded. Samples were morphologically measured for morphometric and meristic traits. Representative samples of each species were sequenced at the cytochrome c oxidase subunit I (COI) gene. The number of dorsal fin spines and general morphological appearance are distinguishable among the five presumed species. However, morphometric measurements overlapped between *Macragnathus semiocellatus* and *Macragnathus siamensis*. K2P distances based on COI sequences among species were high, ranging from 12.4% to 18.7%. All individuals were separated into monophyletic groups of species, clustered into *Mastacembelus* and two *Macragnathus* lineages. Chach bong should be recognized as *Mastacembelus favus* and not *Mastacembelus armatus* as previously classified. No *Mastacembelus armatus* was recorded in the Mekong Delta. GenBank sequences of *Mastacembelus armatus* formed a sister relationship to *Mastacembelus favus* although both have the same range of number of dorsal fin spines and similar reticulated patterns on the body. Misidentification between these two species has been widely recorded in international databases of species taxonomy and DNA barcodes. Nonetheless, their genetic distance (12.4%) is higher than conspecific distances of samples from other regions, indicating the two species can be differentiated by DNA barcoding.

Key words: COI; freshwater fish; phylogenetics; species identification; spiny eel.

Introduction

Species identification is a basic step of research on systematic biology and biodiversity. Current species identification and taxonomic systems of many fishes and taxa have been problematic (Teugels 1996; Kottelat 1998a; Lundberg et al. 2000), which requires the in-

*Corresponding author. E-mail: thuyyen@ctu.edu.vn

©Copyright: the Author(s), 2020 | Licensee PAGEPress, Italy



tegrated approaches of morphology and DNA barcoding (Moritz and Cicero 2004; Teletchea 2009; Kottelat 2013a). Misidentification and misclassification have been reported in diverse fish families such as Cyprinidae (Kottelat 1999), Asian catfish Pangasiidae (Roberts and Vidthayanon 1991; Gustiano *et al.* 2003), etc.

The spiny eel family Mastacembelidae (order Synbranchiformes) has been nominally classified into 141 species belonging to three genera (*Mastacembelus* Scopoli, 1777, *Macrognathus* Lacepède, 1800, and *Sinobdella* Kottelat & Lim, 1994); however, it has undergone several revisions and 88 valid species are currently recognised (Fricke *et al.* 2020). The large number of differences between nominal (141) and valid species (88) implies that species identification and classification in this family are complicated. Traditionally, spiny eel species are identified based on countable traits such as dorsal fin spines, color, spots, bars or reticulated patterns in the body (Rainboth 1996; Tran *et al.* 2013). However, these characteristic keys can vary among life stages or overlap among several species (Tran *et al.* 2013), easily leading to species misidentification.

Among seven species of this family reported in the lower Mekong River (Rainboth 1996), five have been observed in the Mekong Delta, Viet Nam (Truong and Tran 1993; Tran *et al.* 2013). Among them, four species including *Macrognathus circumcinctus* Hora, 1924, *Macrognathus semiocellatus* Roberts, 1986, *Macrognathus siamensis* Günther, 1861, and *Mastacembelus erythrotaenia* Bleeker, 1850 have distinct characteristics which permit ease of identification. The other taxon characterized by brown spots with reticulated pattern on the body is classified by previous authors (Truong and Tran 1993; Tran *et al.* 2013) as *Mastacembelus armatus* Lacepède, 1800 and referred to by the local name of “chach bong” (or chach lau or chach song). However, this species can be misidentified. Rainboth (1996) stated that *Mastacembelus armatus* may be confused with *Mastacembelus favus* Hora, 1924 because both have a reticulated pattern in the body and similar ranges of countable traits. One differentiating characteristic is the arrangement of reticulated pattern which covers the whole body in *Mastacembelus favus* while appearing only on the upper two thirds of the body of *Mastacembelus armatus* (Rainboth 1996). Even then, these characteristics are often inconsistent among individuals, with some presumed *Mastacembelus favus* specimens lacking distinct reticulated pattern covering the whole body. Both species also occupy similar habitats (river, canals, inundated fields, etc.). The type-locality of both species are different, *Mastacembelus favus* was discovered in Phattalung, Southern Thailand while *Mastacembelus armatus* is considered as an Indian subcontinent species (origin in India) (Kottelat 2013b). Over time, the distribution areas of both species overlapped, from South Asia (India, Bangladesh, Pakistan, Sri Lanka) to Southeast Asia (Rainboth 1996; Froese and Pauly 2019). However, because of possible misidentification between the two species, their distribution needs further clarification. We believe misidentification is not only common in Viet Nam but also elsewhere (Rainboth 1996; Kottelat 2016) as their images and descriptions are not consistent within and among global databases such as FishBase, GenBank and BOLD systems.

This study combines morphological and DNA barcoding methods to elucidate species identification and the phylogenetic relationships of spiny eel species belonging to family Mastacembelidae in the Mekong Delta. The morphological and DNA barcode data will assist in filling existing gaps in the geographic distribution and taxonomic understanding of family Mastacembelidae in the Lower Mekong Delta as well as contribute to the global accuracy in identification of spiny eel species in particular between *Mastacembelus favus* and *Mastacembelus armatus*.

Materials and methods

Fish sampling

Adult and sub-adult samples of spiny eel were collected from 2018 to 2019 from fishermen along the Mekong River (Can Tho, CT) and other inland locations (Kien Giang, KG and Long An, LA) in the Mekong Delta, Viet Nam. Based on morphological characteristics, as described in the literature (Rainboth 1996; Kottelat 1998b, 2001; Tran et al. 2013), samples were classified into five taxa (Figure 1). Of these, four were confidently classified into species; *Mastacembelus erythrotaenia*, *Macrognathus circumcinctus*, *Macrognathus semiocellatus* and *Macrognathus siamensis*. The rest of the *Mastacembelus* specimens locally known as chach bong (all specimens morphologically similar, Figure 1A) were tentatively assigned as *Mastacembelus armatus*/*Mastacembelus favus*. These specimens have brown spots displayed fully on the whole body with a reticulated pattern characteristic. No specimens (Figure 1F) with reticulated pattern covering the upper two thirds of the body, diagnostic of *Mastacembelus armatus* (Rainboth 1996) were collected even though sampling had been conducted over two years in different locations in the Mekong Delta. Fin clippings were obtained from each specimen and preserved in 96% alcohol for DNA analysis. The whole specimen was taken photographic images and then kept in ice for immediate morphological analysis. Except for *Mastacembelus erythrotaenia* with only 3 samples, 14 to 32 individuals were analyzed for the other species.

Morphological analysis

The initial morphological analysis focused on examination of external characteristics such as body shape, color and spots. Morphometric traits were measured based on the method described by Cakmak and Alp (2010). Weight of fish and 16 morphometric traits (illustrated in Figure 2, excluding head width and inter-orbital distance) of each specimen were measured using a digital caliper. In addition, the number of dorsal fin spines, the diagnostic meristic character for spiny eels species identification (Rainboth 1996), was counted. For comparative analysis, meristic data of previous studies on *Mastacembelus*

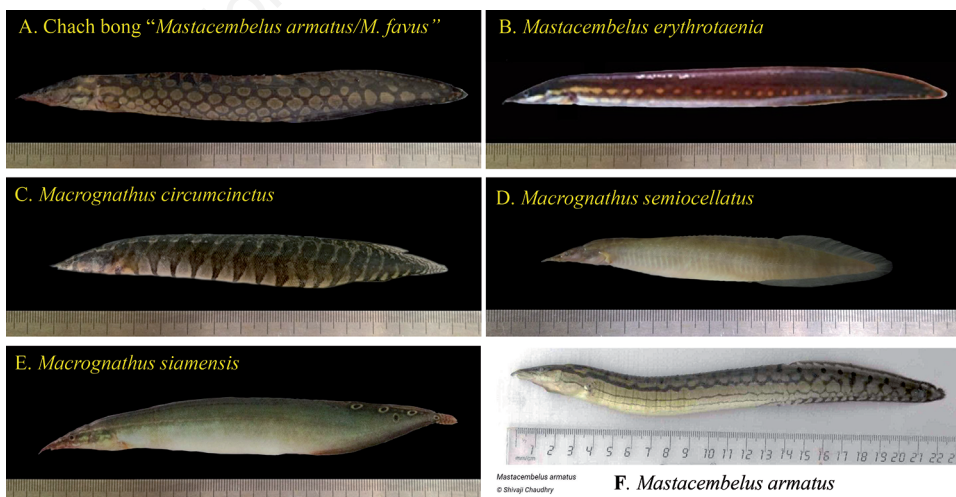


Figure 1. Five spiny eel species (A-E) in the Mekong Delta and *Mastacembelus armatus* (F) in IUCN (picture taken by Shivaji Chaudhry, cited by Fernando et al. 2019).

armatus (Froese and Pauly 2019) and chach bong (Tran *et al.* 2013) were also included (Table 1). After the morphological analysis, 3 to 5 representative samples of each species were preserved as voucher specimens at Can Tho University as reference materials.

DNA analysis

DNA extraction was conducted using ammonium acetate method with the principle of precipitation by salt solution. Cytochrome c oxidase subunit I (COI) gene was amplified using the following universal PCR primer pairs; forward primers Fish F2-t1 and VF2-t1 and reverse primers Fish R2-t1 and VR1d-t1 (Ward *et al.* 2005; Ivanova *et al.* 2007). The components and temperature cycles of PCR reactions (in the machine Swift™ Maxpro) were based on Ward *et al.* (2005). PCR products with a clear band upon electrophoresis as observed on a UV- transilluminator were sent for two-direction sequencing (Apical Scientific Sdn Bhd).

Data analyses

Morphometric measurements were computed as morphometric indices (ratios to standard length or head length, Table 2). Differences in morphometric indices among three

Table 1. Comparing the number of dorsal fin spines in this study with previous reports.

Species	This study	FishBase*	Previous study**
<i>Mastacembelus armatus</i> ***	–	33 – 40	
<i>Mastacembelus armatus/Mastacembelus favus</i>	33 – 34	33 – 37	33 – 36
<i>Mastacembelus erythrotaenia</i>	32	–	32 – 35
<i>Macrognathus circumcinctus</i>	29	26 – 30	26 – 30
<i>Macrognathus semiocellatus</i>	28 – 30	28 – 32	28 – 32
<i>Macrognathus siamensis</i>	17 – 19	13 – 19	13 – 19

Note: *by Froese and Pauly (2019); **by Tran *et al.* (2013); ***this species was not found during the two-year sampling in the present study.

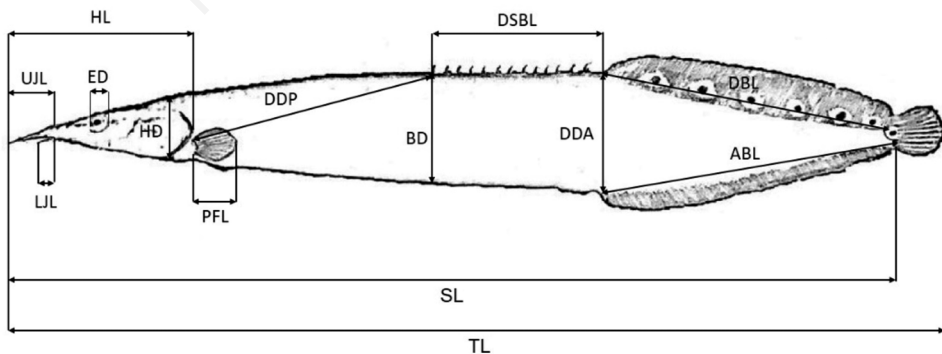


Figure 2. Morphometric measurements of spiny eel species.

TL: total length, SL: standard length, BD: body depth, DSBL: dorsal spines base length, DBL: dorsal fin base length, PFL: pectoral fin length, ABL: anal fin base length, DDP: distance between dorsal spines and pectoral fin, DDA: distance between dorsal fin and anal fin, HL: head length, HD: head depth, UJL: upper jaw length, LJL: lower jaw length, ED: eye diameter.

confirmed species (except *Mastacembelus erythrotaenia* due to the small sample size, N=3) and chach bong were tested using ANOVA and Duncan multiple range tests. In addition, species differentiation was visualized and the cross-validated correction of individual assignment to their original species was estimated using discriminant function analysis. Statistical analyses of morphological data were conducted by using Program SPSS version 20 (IBM Corp. Armonk, NY, USA).

The COI (forward and reverse) sequences of each sample were checked for the quality by the program Finch TV (<http://www.geospiza.com>). Blast searches for sequence similarity were performed with GenBank (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and BOLD systems (Ratnasingham and Hebert 2007) databases. GenBank sequences of *Mastacembelus armatus*, *Sinobdella sinensis* [a primitive mastacembeliids and believed to form the sister group to all remaining mastacembelids (Johnson and Patterson 1993; Britz 1996)] and *Chaudhuriia caudata* (outgroup) were included in the analysis (Figure 4). The maximum likelihood (ML) phylogenetic tree reconstruction was carried out with 1000 bootstrap replications using MEGA v7.0 (Kumar et al. 2016). Pairwise genetic distances were estimated using Kimura two-parameter model (Kimura 1980).

Table 2. Morphometric indices (mean \pm SD) of Mastacembelidae species.

Species	<i>Mastacembelus armatus</i> / <i>Mastacembelus favus</i> * (N=32)	<i>Macragnathus circumcinctus</i> (N=14)	<i>Macragnathus semiocellatus</i> (N=31)	<i>Macragnathus siamensis</i> (N=31)	<i>Mastacembelus erythrotaenia</i> (N=3)
Weight (g)	30.6 \pm 14.4	11.4 \pm 4.1	11.4 \pm 2.9	15.6 \pm 7.9	35.4 \pm 17.9
Total length (cm)	22.2 \pm 3.4	14.5 \pm 1.8	15.4 \pm 1.2	15.9 \pm 1.9	37.1 \pm 18.6
<i>Ratio to standard length</i>					
BD	9.63 \pm 1.09 ^a	11.3 \pm 3.2 ^b	10.8 \pm 0.7 ^b	12.7 \pm 1.2 ^c	10.7 \pm 0.8
DSBL	42.4 \pm 7.0 ^c	28.5 \pm 3.7 ^b	30.1 \pm 3.0 ^b	15.6 \pm 2.7 ^a	46.6 \pm 0.8
DBL	41.6 \pm 7.3 ^c	27.3 \pm 4.5 ^{ab}	28.2 \pm 2.4 ^b	24.7 \pm 3.9 ^a	36.2 \pm 0.3
PFL	5.80 \pm 1.20 ^c	3.80 \pm 0.70 ^b	4.06 \pm 0.47 ^a	4.68 \pm 0.72 ^a	4.68 \pm 0.05
ABL	41.9 \pm 7.5 ^c	29.4 \pm 5.8 ^b	31.8 \pm 3.1 ^b	25.1 \pm 3.5 ^a	38.4 \pm 0.5
DDP	6.48 \pm 1.00 ^b	4.79 \pm 1.09 ^a	4.87 \pm 0.88 ^a	18.3 \pm 2.6 ^c	8.44 \pm 0.57
DDA	10.1 \pm 1.7 ^b	8.41 \pm 1.38 ^a	8.39 \pm 1.12 ^a	8.91 \pm 1.57 ^a	10.2 \pm 0.1
HL	19.3 \pm 0.9 ^b	18.6 \pm 1.3 ^a	18.3 \pm 0.9 ^a	20.2 \pm 1.3 ^c	16.6 \pm 0.3
<i>Ratio to head length</i>					
HD	30.7 \pm 6.3 ^b	18.7 \pm 2.3 ^a	18.5 \pm 2.0 ^a	21.8 \pm 3.7 ^b	34.5 \pm 1.3
HW	21.0 \pm 3.5 ^b	10.6 \pm 1.4 ^a	10.9 \pm 1.9 ^a	12.1 \pm 2.5 ^a	15.1 \pm 0.2
UJL	32.8 \pm 7.3 ^b	16.3 \pm 2.2 ^a	17.6 \pm 1.8 ^a	17.4 \pm 1.9 ^a	28.4 \pm 0.3
LJL	20.0 \pm 4.3 ^b	8.26 \pm 2.10 ^a	7.84 \pm 1.52 ^a	7.00 \pm 1.15 ^a	15.3 \pm 0.4
IOD	7.56 \pm 3.06 ^c	3.87 \pm 0.63 ^a	3.73 \pm 0.71 ^a	5.41 \pm 1.00 ^b	6.04 \pm 0.16
ED	7.69 \pm 0.76 ^c	5.27 \pm 0.63 ^a	5.12 \pm 0.62 ^a	6.17 \pm 1.00 ^b	10.9 \pm 0.2

Note: (*) Local name: Chach bong.

BD: body depth, DSBL: dorsal spine base length, DBL: dorsal fin base length, PFL: pectoral fin length, ABL: anal fin base length, DDP: distance between dorsal spines and pectoral fin, DDA: distance between dorsal fin and anal fin, HL: head length, HD: head depth, HW: head width, UJL: upper jaw length, LJL: lower jaw length, IOD: inter-orbital distance, ED: eye diameter.

Means in the same rows sharing the same superscript letters are not significantly different ($P > 0.05$).

Results

Morphological identification and measurements

All specimens of Mastacembelidae displayed distinct color and patterns on their body, characteristic of their respective species (Figure 1). Chach bong “*Mastacembelus armatus*/*Mastacembelus favus*” has many large brown spots (reticulated pattern) on the whole body, congruent to the description of *Mastacembelus favus* by Rainboth (1996). *Macragnathus circumcinctus* has oblique blackish bars on the body. *Macragnathus siamensis* is identified by a series of separate black points along the dorsal fin and caudal fin. *Mastacembelus erythrotaenia* is distinguished by many red stripes and red spots on the body. *Macragnathus semiocellatus* has a row of faint spots along the dorsal fin. These characteristics of the four confirmed species are similar to those observed in the previous studies (Truong and Tran 1993; Tran *et al.* 2013).

The number of dorsal fin spines varies between the two genera *Mastacembelus* and *Macragnathus* but overlaps between some congeneric species such as *Mastacembelus armatus* - *Mastacembelus favus* and *Macragnathus circumcinctus* - *Macragnathus semiocellatus* (Table 1). Only *Macragnathus siamensis* has species-specific number of dorsal fin spines. The ranges of their dorsal spine numbers matched with previous studies (Tran *et al.* 2013; Froese and Pauly 2019).

Morphometric indices (Table 2) of four species (*Mastacembelus erythrotaenia* was excluded because of the small sample size, N=3) are significantly different ($P < 0.05$). The chach bong, “*Mastacembelus armatus*/*Mastacembelus favus*” group differs in all indices compared to the other species in the ANOVA after Duncan’s multiple range tests. Two species *Macragnathus circumcinctus* and *Macragnathus semiocellatus* are similar in 15/16 indices (except pectoral fin length, PFL/SL). Differentiation among samples and species in multivariate morphometric characters are visualized in the plot of canonical discriminant functions (Figure 3).

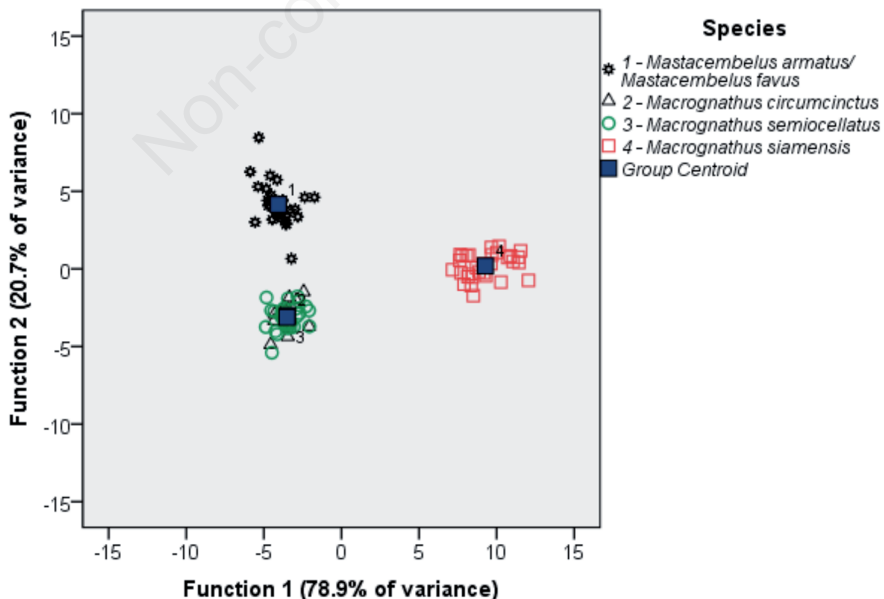


Figure 3. Canonical discriminant analysis based on morphometric indices of four mastacembelid eels (*Mastacembelus armatus*/*Mastacembelus favus* for chach bong).

Overall, there are three clusters; 1. chach bong, “*Mastacembelus armatus/ Mastacembelus favus*”, 2. *Macrognathus circumcinctus* and *Macrognathus semiocellatus* group, and 3. *Macrognathus siamensis*. *Macrognathus circumcinctus* and *Macrognathus semiocellatus* overlap in both Function 1 and Function 2, while *Macrognathus siamensis* is distinct from the other species on both Function 1 and Function 2. Important indices that contribute to Function 1 (differentiating *Macrognathus siamensis*) include dorsal fin base length (DSBL and DBL), distance between dorsal spines and pectoral fin (DDP) and anal fin base length (ABL), while important indices in Function 2 are head depth and width (HD and HW), mouth size (lower jaw and upper jaw length). Discriminant analysis showed that 83.3% of cross-validated grouped cases are correctly classified, in which classification results are 100% for chach bong and *Macrognathus siamensis* and 50.0% and 67.7% for *Macrognathus circumcinctus* and *Macrognathus semiocellatus*, respectively.

Species identification and COI-based phylogeny of mastacembelids

Twenty sequences generated from the study have been submitted to the GenBank databases under accession number from MT508532 to MT508551. Four species including *Mastacembelus erythrotaenia*, *Macrognathus circumcinctus*, *Macrognathus semiocellatus*, and *Macrognathus siamensis* have greater than 99% sequence identity with the same species reported in GenBank (Table 3) and 99 - 100% probability of correct species identification in BOLD systems (data not shown). COI sequences of chach bong which has been classified as *Mastacembelus armatus* in various studies (Truong and Tran 1993; Tran et al. 2013), instead matched closely with *Mastacembelus favus* sequences with several distinct exceptions. It is identical by 99.8 to 100% with *Mastacembelus favus* samples from Southeast Asia (for example, sequence Acc. No KT944605 and KT944606) but has similarity of 93% when compared to *Mastacembelus favus* samples in Thailand (Acc. No MK628409, see Figure 4). A search on the BOLD database aligned chach bong sequences with *Mastacembelus favus* at 100% identity with all five deposited sequences. On the other hand, the chach bong sequences are highly different from those of *Mastacembelus armatus* reported in GenBank. For example, they have 88.6% similarity with *Mastacembelus armatus* sequence number KT944631 (Figure 4). This result indicates that chach bong specimens collected in the Mekong Delta is *Mastacembelus favus*. Therefore, chach bong is herein referred as *Mastacembelus favus*.

Genetic distances based on COI sequences among five spiny eel species (Table 4) vary from 12.4% (*Mastacembelus favus* and GenBank *Mastacembelus armatus*) to 18.7% (*Mas-*

Table 3. Genetic distance (d, %) within species of Mastacembelidae in the Mekong Delta and species identification based on GenBank (GB) and BOLD databases.

Species	N	No of haplotype	d (±SE)	Sequence identity with GB (%)*
<i>Mastacembelus favus</i>	5	2	0.3±0.1	99.9 – 100
<i>Mastacembelus erythrotaenia</i>	3	2	0.4±0.2	99.7 – 100
<i>Macrognathus circumcinctus</i>	4	2	0.7±0.2	99.2 – 99.7
<i>Macrognathus semiocellatus</i>	4	3	0.7±0.2	99.1 – 99.9
<i>Macrognathus siamensis</i>	4	3	0.5±0.2	99.7 – 100

Note: N = sample size. (*) The highest sequence identity of the same species that was aligned with haplotypes in the present study. Following Acc. No KT944606.1, JQ769007.1, KT944565.1, KT944583.1, and KT944652.1 were used for comparison with five species respectively.

tacembelus favius and *Macrognathus siamensis*). Within-genus genetic distances are lower in *Mastacembelus* (from 12.4 to 14.0%) than *Macrognathus* (from 13.1 to 17.9%).

The ML phylogenetic tree clustered all six species (five from the study and *Mastacembelus armatus* from GenBank) into three main lineages; Lineage 1 for all other *Mastacembelus* spp., Lineage 2 includes *Macrognathus circumcinctus* and *Macrognathus semiocellatus*, and Lineage 3 consists of *Macrognathus siamensis* (Figure 4). For *Mastacembelus favius* and *Mastacembelus armatus*, each species shows genetic variation in COI among samples from different geographical regions, such as between India and South-east Asia (for *Mastacembelus armatus*) or between Mekong Delta, Viet Nam and Thailand (for *Mastacembelus favius*).

Table 4. Average pairwise K2P genetic distances (%) between mastacembelids based on COI gene.

	1	2	3	4	5	6
1 <i>Mastacembelus favius</i>	1.7					
2 <i>Mastacembelus armatus</i> *	12.4	3.7				
3 <i>Mastacembelus erythrotaenia</i>	13.1	14.0	0.0			
4 <i>Macrognathus circumcinctus</i>	15.5	14.5	15.4	0.4		
5 <i>Macrognathus semiocellatus</i>	16.1	16.0	16.2	13.1	0.6	
6 <i>Macrognathus siamensis</i>	18.7	15.9	17.9	14.3	17.8	0.2

Note: (*) Four COI sequences of *Mastacembelus armatus* from GenBank with accession numbers indicated in Figure 4. Values in the diagonal are genetic distances of within-species in this study and GenBank sequences.

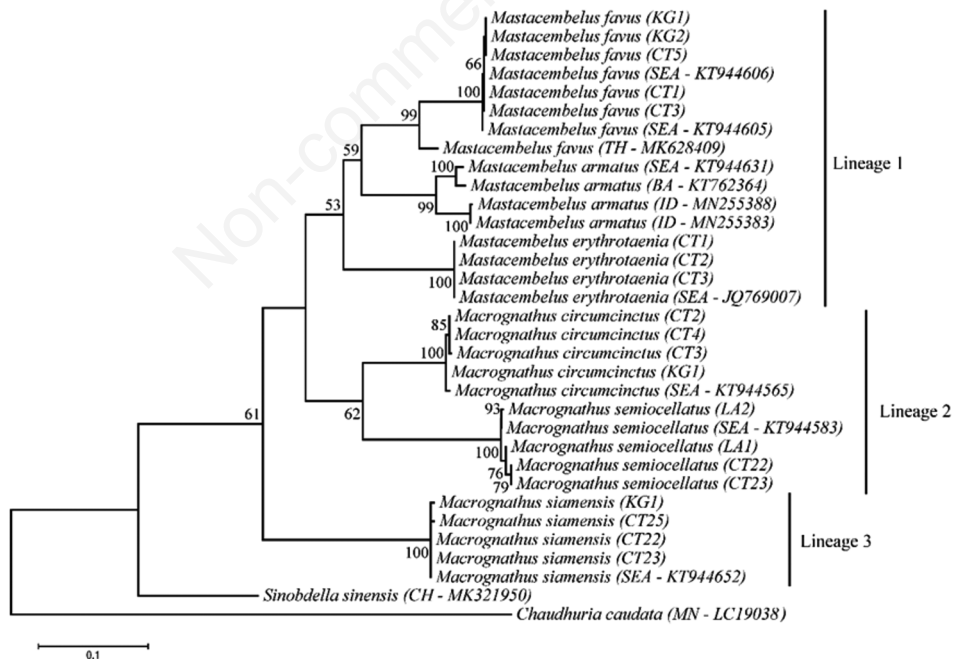


Figure 4. Maximum Likelihood tree of mastacembelids based on mtDNA COI gene. Only bootstrap values greater than 50% are shown. Note: Sequences from Southeast Asia (SEA), Bangladesh (BA), India (ID), Thailand (TH), China (CH) and Myanmar (MN) were obtained from GenBank with accession numbers in parentheses.

Discussion

The combination of morphological and DNA data unambiguously identified four of the five species sampled in family Mastacembelidae in the Mekong Delta. In addition, it adds to the knowledge gap in morphological information of *Mastacembelus favus*, which assists to resolve the taxonomic confusion of the chach bong in the Viet Nam freshwater habitat and clarify the difference between two congeneric species *Mastacembelus armatus* and *Mastacembelus favus*. Our study disputes earlier studies that the chach bong is *Mastacembelus armatus* (Truong and Tran 1993; Tran et al. 2013) but instead it was identified as *Mastacembelus favus*.

Comparative morphology between *Mastacembelus favus* (this study) and *Mastacembelus armatus* (available databases) highlights its close relationship (similar spots in the body and the same range of dorsal fin spines). Due to similarities in external appearance, there is a great deal of taxonomic confusion of the fish taxon, locally known as chach bong in Viet Nam. While several studies have classified it as *Mastacembelus armatus* (Truong and Tran 1993; Tran et al. 2013), images of the two species available in some worldwide databases such as FishBase and BOLD systems can only at best be described as not definitive. In BOLD systems, for example, the species description of *Mastacembelus armatus* and available images are inconsistent and different from the photos described by Rainboth (1996) or Shivaji Chaudhry (Figure 1F) cited by Fernando et al. (2019). In addition, apart for the number of vertebrae, spines and soft rays of the dorsal and anal fins, there is limited information on the morphology of the two species. However, these countable traits are overlapping between the two species (Rainboth 1996), leading to the difficulty in species identification. There also appear to be several errors in COI sequence data of *Mastacembelus armatus* submitted in BOLD systems. Several GenBank sequences of *Mastacembelus armatus* are grouped in the same lineages of *Mastacembelus favus* in the phylogenetic tree of 100 sequences that are closest to a sequence of *Mastacembelus favus* in the present study (Appendix). We opine that these *Mastacembelus armatus* samples (from Cambodia) are misidentification. This confusion could be attributed to inadequate and debatable morphological descriptions in reference materials used in the initial identification. Such discrepancies are also commonly documented in other fish taxa (Will and Rubinoff 2004; Packer et al. 2009; Decru et al. 2016).

Both *Mastacembelus armatus* and *Mastacembelus favus* were reported to have similar ranges of natural distribution, from southern to southeast Asia (Rainboth 1996; Froese and Pauly 2019). Our study reveals that the commonly found chach bong is *Mastacembelus favus* while *Mastacembelus armatus* was not recorded in this survey. It is possible that wild populations of the latter species are now rare in the lower Mekong River Basin. However, the red list status of *Mastacembelus armatus* in 2019 was classified as “least concern” species (Fernando et al. 2019). Thus, it is unlikely that they have gone into local extinction in the Mekong Delta. In other places such as India or Bangladesh this species is still abundant (Gupta and Banerjee 2016). A more plausible explanation for the absence of *Mastacembelus armatus* during two-year sampling in the Mekong Delta is that this species does not naturally distribute in the region. Previous studies on fish diversity in the Mekong Delta (Truong and Tran 1993; Tran et al. 2013) mentioned only one species of chach bong (which is identified as *Mastacembelus favus* in the present study). Based on information available from BOLD systems and literature (Figure 5), *Mastacembelus armatus* was mainly reported in South Asia such as India and Bangladesh (BOLD database), while *Mastacembelus favus* was reported in Southeast Asia including Thailand, Cambodia, Vietnam (i.e. Kano et al. 2013; Jamaluddin et al. 2019). In addition, Jamalud-

din *et al.* (manuscript in preparation) found that these two sister species are allopatric in their regional distribution; *Mastacembelus armatus* is distributed across the Indian sub-continent and Myanmar, while *Mastacembelus favus* is restricted to the Southeast Asian biogeographic region, except in Myanmar. Previous reports of overlapping distribution areas of these *Mastacembelus* species (Rainboth 1996; Froese and Pauly 2019) could be based on misidentification between the two species.

In contrast to *Mastacembelus favus*, the other four mastacembelid species (*Mastacembelus erythrotaenia*, *Macrognathus circumcinctus*, *Macrognathus semiocellatus*, and *Macrognathus siamensis*) could be easily identified at sub-adult and adult stages based on the color and spot patterns on their body (Figure 1). Their morphological descriptions are consistent in the literature (Rainboth 1996; Kottelat 1998b, 2001; Tran *et al.* 2013). However, identification based on color and spot patterns of three *Macrognathus* species is more difficult in fingerling or juvenile stages where the body color are similar and spots are unclear among species (Tran *et al.* 2013). In such cases, using the number of dorsal fin spines or morphometric measurements can discriminate *Macrognathus siamensis* from the rest, but could not differentiate between *Macrognathus circumcinctus* and *Macrognathus semiocellatus*. Further studies can explore other internal morphological traits (such as myology and osteology) using micro computed tomography techniques (Faulwetter *et al.* 2013). Here, in the present study, we showed that above problem could be easily resolved through DNA barcoding.

The results on COI sequences showed a high inter-specific divergence (K2P range from 12.4 to 18.7%) in comparison to within-species COI variation (from 0 to 3.7% based on sequences from the present study and GenBank used in Figure 4). Our study

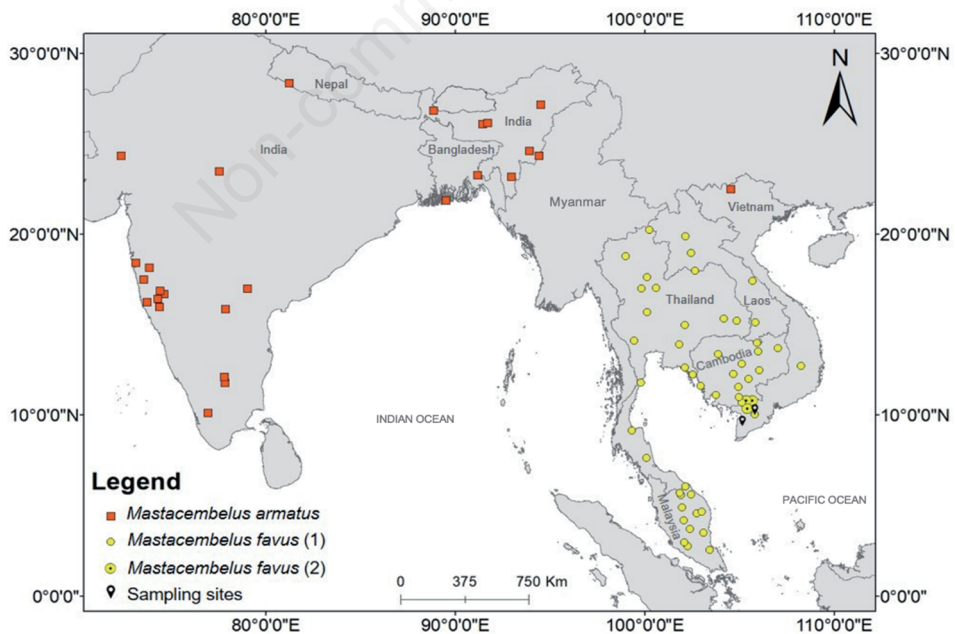


Figure 5. Distribution of *Mastacembelus armatus* based on BOLD database (Ratnasingham and Hebert 2007) and that of *Mastacembelus favus* based on (1) Jamaluddin *et al.* (2019), (2) Nagao project (Kano *et al.* 2013), and from the present study (sampling sites).

noted that *Macrognathus semiocellatus* and *Macrognathus circumcinctus* could not be classified based on morphometric indices (Table 2 and Figure 3) but are highly divergent in the COI sequences (Table 4). Therefore, they could be easily distinguished by DNA barcoding to complement the morphological data such as at larval or juvenile stages or in incomplete samples. Although based on only a few representative species, our study has also been able to capture a general measure of the genetic diversity within this family. The average genetic divergence among species in Mastacembelidae (15.39%) is comparable to results reported from other fish families in Australia 15.46% (Ward et al. 2005) and Canada 15.4% (Hubert et al. 2008). Furthermore, the separation of *Macrognathus* into two different lineages in the phylogenetic tree is in concordance with morphological classification after Travers (1984) and Roberts (1986).

The phylogenetic trees of mastacembelid eel species (Figure 4) and 100 sequences of *Mastacembelus favus* (Appendix) from the BOLD systems highlighted the high levels of conspecific variation in COI in the two species among regions/ countries of fish collection. Jamaluddin et al (2019) found evidence of strong genetic structure of *Mastacembelus favus* in Southeast Asia among geographical regions as resulting from historical dissection of drainages and various routes of species colonisation, based on Cytochrome b. A genetic distance of 6.5% was estimated for a pairwise comparison between a GenBank *Mastacembelus favus* sample collected from Ing River, a tributary of Middle Mekong, in Thailand (Acc. No MK628409.1), which corresponded to Middle Mekong lineage in Jamaluddin et al. (2019). Similarly, *Mastacembelus armatus* samples from India (GenBank Acc. No MN255383 and MN255388) diverged by 5.2% from those of Bangladesh and Southeast Asia (Acc. No. KT762364 and KT762631, respectively). Such levels of conspecific variation is considered atypically high as compared to other fishes. Average K2P distance within species was reported at 2.7% in 190 fish species in Canada (Hubert et al. 2008), and 3.9% in 173 species in Australia (Ward et al. 2005). Regardless of the high intraspecific variation, the two species *Mastacembelus armatus* and *Mastacembelus favus* are highly divergent in the phylogenetic tree (K2P distance: 12.4%). Thus, they can be classified correctly based on DNA barcoding due to the higher inter-specific to intra-specific genetic distances. High levels of COI sequence variation of each species across geographical regions warrant further studies on phylogeography of spiny eel species.

The clustering of mastacembelid species into three lineages in the phylogenetic tree is in concordance with their differences and similarities in morphology (i.e. the number of dorsal fin spines and morphometric measurements). Both types of data show that *Macrognathus siamensis* is distinct from the rest, while species of genus *Macrognathus* (*Macrognathus semiocellatus* and *Macrognathus circumcinctus*) and those of *Mastacembelus* form the other two groups. It also highlights that despite advancement in molecular technique, the morphological approach is equally invaluable particularly where resources are limited.

In conclusion, the present study has expanded the morphological and DNA barcoding data of five spiny eel species in Mekong Delta, which has allowed the resolution of incongruous species identification between *Mastacembelus armatus* and *Mastacembelus favus*. Our results confirm that the chach bong in the Mekong Delta is *Mastacembelus favus* and is abundantly found in this area.

Conflict of Interest/Funding

This study was funded by the Can Tho University Improvement Project VN14-P6, supported by Japanese ODA loan. No potential conflict of interest was reported by the authors.

Contributions

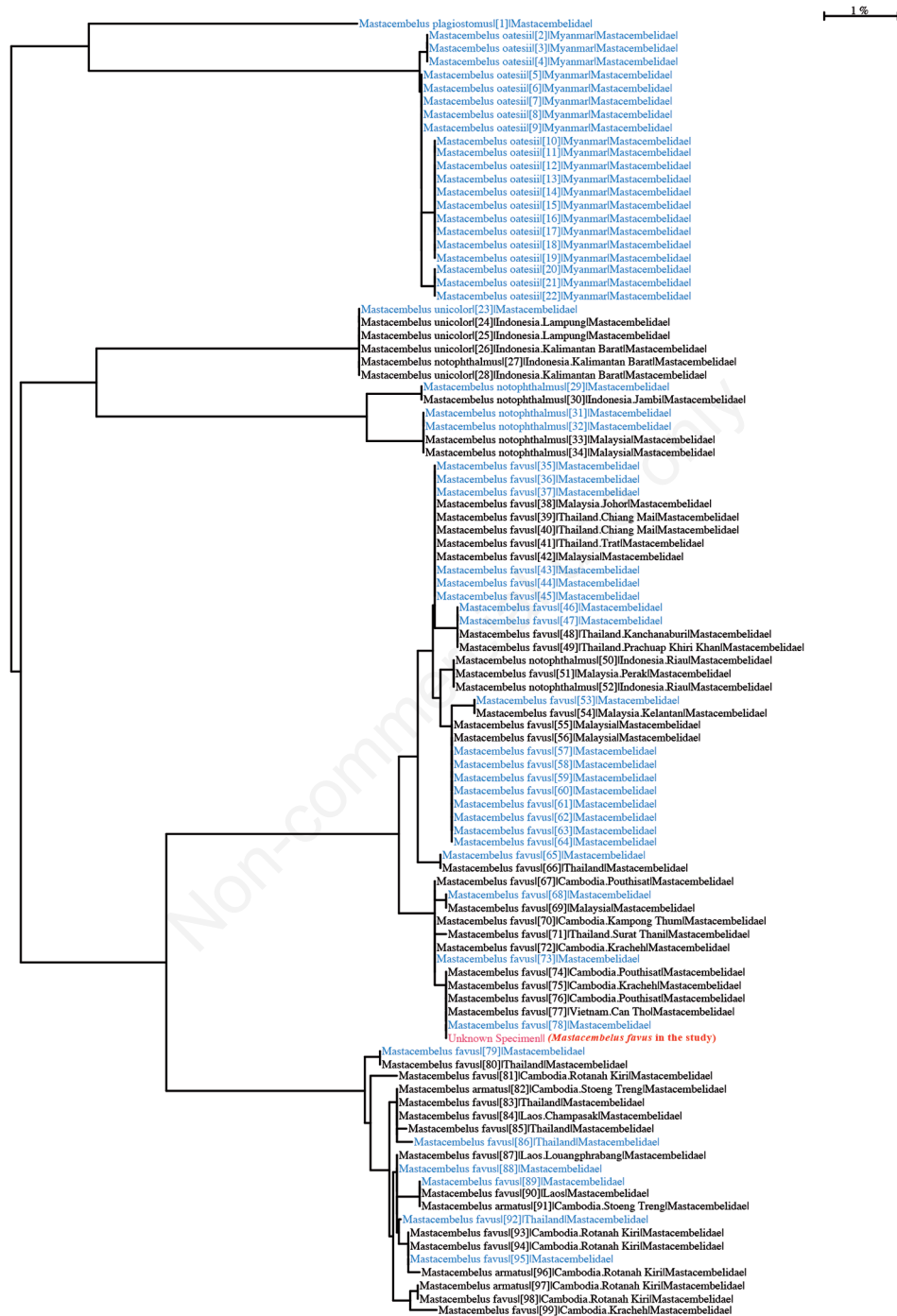
T-YD: Conceptualization, Methodology, Writing- Original draft preparation, Supervision. LVDT: Investigation. N-TTN: Visualization, Investigation. JAFJ: Formal analysis, editing. MNSA: Writing- Reviewing and Editing.

References

- Britz R. 1996. Ontogeny of the ethmoidal region and hyopalatine arch in *Macrognathus pancalus* (Percomorpha, Mastacembeloidei), with Critical Remarks on Mastacembeloid Inter- and Intra-relationships. *American Museum Novitates*. 3181(2):1–18.
- Cakmak E, Alp A. 2010. Morphological differences among the mesopotamian spiny eel, *Mastacembelus mastacembelus* (Banks & Solander 1794), populations. *Turkish Journal of Fisheries and Aquatic Sciences*. 10:87–92.
- Decru E, Moelants T, De Gelas K, Vreven E, Verheyen E, Snoeks J. 2016. Taxonomic challenges in freshwater fishes: A mismatch between morphology and DNA barcoding in fish of the north-eastern part of the Congo basin. *Molecular Ecology Resources*. 16:342–352.
- Faulwetter S, Vasileiadou A, Kouratoras M, Thanos Dailianis, Arvanitidis C. 2013. Micro-computed tomography: Introducing new dimensions to taxonomy. *Zookeys*. 263:1–45.
- Fernado M, Kotagama O, de Alwis Goonatilake S. 2019. *Mastacembelus armatus*. The IUCN Red List of Threatened Species 2019: e.T166586A60592409. <https://dx.doi.org/10.2305/IUCN.UK.2010-4.RLTS.T166586A6241626.en>.
- Fricke R, Eschmeyer WN, Fong JD. 2020. Eschmeyer's Catalog of Fishes: Species by Family/Sub-family. Electronic version accessed 4 May 2020.
- Froese R, Pauly D. 2019. FishBase, World Wide Web electronic publication. www.fishbase.org.
- Gupta S, Banerjee S. 2016. Food, feeding habit and reproductive biology of tire-track spiny eel (*Mastacembelus armatus*): A Review. *Journal of Aquaculture Research & Development*. 7(5):1000429.
- 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). *Journal of Natural History*. 37(3):357–376.
- Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, Burrige M, Watkinson D, Dumont P, Curry A, Bentzen P, et al. 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS ONE*. 3:e2490.
- Ivanova N V., Zmlak TS, Hanner RH, Hebert PDN. 2007. Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes*. 7(4):544–548.
- Jamaluddin JAF, So N, Tam BM, Ahmad A, Grudpan C, Page LM, Khaironizam MZ, Mohd Nor SA. 2019. Genetic variation, demographic history and phylogeography of tire track eel, *Mastacembelus favus* (Synbranchiformes: Mastacembelidae) in Southeast Asia. *Hydrobiologia*. 838(1):163–182.
- Johnson GD, Patterson C. 1993. Percomorph phylogeny: a survey of acanthomorphs and a new proposal. *Bulletin of Marine Science*. 52(1):554–626.
- Kano Y, Adnan MS, Grudpan C, Grudpan J, Magtoon W, Musikasinthorn P, Natori Y, Ottomanski S, Praxaysonbath B, Phongsa K, et al. 2013. An online database on freshwater fish diversity and distribution in Mainland Southeast Asia. *Ichthyological Research*. 60:293–295.
- Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*. 16(2):111–120.
- Kottelat M. 1998a. Systematics, species concepts and the conservation of freshwater fish diversity in Europe. *Italian Journal of Zoology*. 65(1):65–72.
- Kottelat M. 1998b. Fishes of the Nam Theun and Xe Bangfai basins, Laos, with diagnoses of twenty-two new species (Teleostei: Cyprinidae, Balitoridae, Cobitidae, Coidae and Odontobutidae). *Ichthyological Exploration of Freshwaters*. 9(1):1–128.
- Kottelat M. 1999. Nomenclature of the genera Barbodes, Cyclocheilichthys, Rasbora and Chonerhinos (Teleostei: Cyprinidae and Tetraodontidae), with comments on the definition of the first reviser. *Raffles Bulletin of Zoology*. 47(2):591–600.

- Kottelat M. 2001. Fishes of Laos. Colombo: WHT Publications Ltd. 198 pp.
- Kottelat M. 2013a. The valid generic names for the fish species usually placed in Cyclocheilichthys (Pisces: Cyprinidae). *Zootaxa*. 3640(4):479–482.
- Kottelat M. 2013b. The fishes of the inland waters of Southeast Asia: A catalogue and core bibliography of the fishes known to occur in freshwaters, mangroves and estuaries. *Raffles Bulletin of Zoology*. Supplement 27:1–663.
- Kottelat M. 2016. The fishes of the Nam Theun and Xe Bangfai drainages, Laos. *Hydroecologie Appliquee*. 19:271–320.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*. 33(7):1870–1874.
- Lundberg JG, Kottelat M, Smith GR, Stiassny MLJ, Gill AC. 2000. So many fishes, so little time: An overview of recent ichthyological discovery in continental waters. *Annals of the Missouri Botanical Garden*. 87(1):26–62.
- Moritz C, Cicero C. 2004. DNA barcoding: Promise and pitfalls. *PLoS Biology*. 2(10):e354.
- Packer L, Gibbs J, Sheffield C, Hanner R. 2009. DNA barcoding and the mediocrity of morphology. *Molecular Ecology Resources*. 9:42–50.
- Rainboth WJ. 1996. Fishes of the Cambodian Mekong. FAO species identification field guide for fishery purposes. Rome: Food and Agriculture Organization of the United Nations. 310pp.
- Ratnasingham S, Hebert PDN. 2007. BOLD: The Barcode of Life Data System. *Molecular Ecology Notes*. 7:355–364.
- Roberts TR. 1986. Systematic review of the Mastacembelidae or spiny eels of Burma and Thailand, with description of two new species of *Macrognaathus*. *Japanese Journal of Ichthyology*. 33(2):95–109.
- Roberts TR, Vidthayanon C. 1991. Systematic revision of the Asian catfish family Pangasiidae, with biological observations and descriptions of 3 new species. *Proceedings of the Academy of Natural Sciences of Philadelphia*. 143:97–144.
- Teletchea F. 2009. Molecular identification methods of fish species: Reassessment and possible applications. *Reviews in Fish Biology and Fisheries*. 19(3):265–293.
- Teugels GG. 1996. Taxonomy, phylogeny and biogeography of catfishes (Ostariophysi, Siluroidei): an overview. *Aquatic Living Resources*. 9(SupplementS1):9–34.
- Tran DD, Shibukawa K, Nguyen TP, Ha PH, Tran XL, Mai VH, Utsugi K. 2013. Fishes of the Mekong Delta, Viet Nam. Viet Nam: Can Tho University Publishing House. 174 pp.
- Travers RA. 1984. A review of the Mastacembeloidei, a suborder of synbranchiform teleost fishes. Part 2: phylogenetic analysis. *Bulletin of the British Museum (Natural History)*. 47(2):83–150.
- Truong TK, Tran TTH. 1993. Classification of freshwater fish in the Mekong Delta, Viet Nam. Viet Nam: Can Tho University. 361 pp.
- Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 360(1462):1847–1857.
- Will KW, Rubinoff D. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics*. 20(1):47–55.

Appendix



Appendix. COI-species tree for *Mastacembelus favus* generated by BOLDsystems. *Mastacembelus favus* sequence in the present study is marked red (as unknown sequence) on the tree with BOLD sequences in black. GenBank sequences are marked in blue.