MOLECULAR PHYLOGENY OF CATFISHES (TELEOSTEI: SILURIFORMES) INFERRED FROM MITOCHONDRIAL MARKERS – IMPLICATIONS FOR LOWER MEKONG RIVER BASIN

Quyen Vu Dang Ha[1,2*], Oanh Truong Thi[1], Phuong Thai Thi Lan[1], Thuoc Tran Linh[3], Binh Dang Thuy[1]
[1]Institute for Biotechnology and Environment, Nha Trang University, 02 Nguyen Dinh Chieu, Nha Trang City, **VIETNAM**
[2]Faculty of Biology and Biotechnology, University of Science, Ho Chi Minh City, 227 Nguyen Van Cu, Ho Chi Minh City, **VIETNAM**
[3]Department of Molecular & Environmental Biotechnology, University of Science, 227 Nguyen Van Cu, Ho Chi Minh City, **VIETNAM**
*Corresponding author: quyenvdh@ntu.edu.vn

ABSTRACT

Catfish species in the Order Siluriformes distribute widely in Mekong River Basin, where they are important economic species for aquaculture and fisheries, providing major food sources for local communities. However, phylogenetic relationships of Siluriformes have remained unresolved at high taxonomic level. In this study, phylogenetic trees of 30 catfish species belonging to nine families collected from Lower Mekong Basin (LMB) were constructed based on two mitochondrial gene sequences using Neighbor Joining (NJ), Bayesian inference (BI) and Maximum likelihood (ML) methods. The results obtained from the BI and NJ methods showed more similarity than those obtained from ML method. Monophyly of Siluriformes was strongly supported by all applied phylogenetic methods. Seven of nine studied families performed monophyletic clade, while other families including Pangasiidae (14 species, four genera) and Bagridae (11 species, three genera) displayed paraphyletic. All analyses indicated that: (i) Loricaridae was placed as basal and being a sister clade to remaining families; (ii) Among Silurioidei, Plotosidae showed as a sister clade to the rest of families. Interfamilial relationships of Siluriformes performed divergent between data sets, and methods used, and showing unmatched with traditionally and recently reported catfish phylogenies.

Keywords: Catfish, mitochondrial marker, molecular phylogeny, LMB, Siluriformes.

INTRODUCTION

Catfishes (Order Siluriformes) is a diverse group of ray-finned fishes (Nelson, 2006). Armbruster (2011) listed 4854 nominal species, of which 3407 valid species belong to 36 families following Diogo (2003), Lundberg et al., (2007), Eschmeyer & Fong (2010). They are primarily fresh and brackish water fish with only two secondary marine families Plotosidae and Ariidae (Kailola, 2004).

Catfishes taxonomy have been studied extensively for many years (Diogo, 2003; Nelson, 2006; Eschmeyer & Fong, 2010). In term of molecular evolution and phylogenetic relationships, both mitochondrial and nuclear markers have increasingly applied. Sullivan et al., (2006) conducted deep phylogenetic analysis of Siluriformes covering 110 species of 36 valid families using nuclear recombination activating gene (rag 1 and rag 2). Monophyly of Siluroidei was strongly supported by rag data; Loricarioidea was clustered as the sister group to other catfishes (Diplomystidae and Siluroidei), which contrasts to the hypothesis that Diplomystidae is the sister clade instead. The multi-family clades referred as "Big Africa" and "Big Asia" have been discovered suggest an intra-continental diversification of catfish (Sullivan et al., 2006; Kappas et al., 2016). Kappas et al., (2016) re-examined catfish phylogeny to look for their phylogenetic history, timeline and mode of diversification (using full length metagenome sequence of 62 species represented 20 families). They found somehow similar phylogenetic relationship as Sullivan et al., (2006), except Diplomystoidea placed as a sister group to all the remaining Siluroidei. The basal split of catfish was hypothesized as Pangaean origin (Early Cretaceous). For the specific group of catfish, Peng et al., (2004) obtained completed Cytochrome b gene to verify unresolved glyptosternoid (Siluriformes: Sisoridae) phylogeny in China. Jansen et al., (2006) reported clade monophyly of African anguilliform catfish (Clariidae), and showed anguilliformity seeming to have arisen at least four times from a fusiform *Clarias* like ancestor. Azlina

et al., (2013) showed the phylogenetic relationships of five pangasiid (Pangasiidae) species in Malaysia using Cytochrome c Oxidase subunit I (COI) gene. Recently, Yu & Quilang (2014) applied three mitochondrial genes (COI, 16S rRNA and Cyt b) combined with two nuclear genes (rag1 and rag2) to construct the phylogenetic tree from seven native and four introduced catfishes in the Philippines.

The Mekong River Basin (MRB) represents a global hotspot of aquatic biodiversity (Dudgeon et al., 2006; Allen et al., 2012), second only to the Amazon River in terms of total fish species richness (Poulsen et al., 2004; Baran & Guerin, 2012). Among around 1200 species have been reported, 127 recognized catfish species were listed (MRC, 2005). Although possessing high species diversity, extensive research has only focused on species diversity using morphological characters (Rainboth, 1996; Kottelat, 2001; Tran et al., 2013; Thai, 2016). However, unwell-defined and environmental varied morphological characters make the familial interrelationships and the validity of many nominal species are problematic (Dodson & Lecomte, 2015; Ünlü et al., 2002). Nowadays, single and multiple genomic region DNA barcodes have proved as useful tool for species identification, evaluating taxonomic diagnoses, and examining evolutionary relationships (Thai et al., 2007; Karinthanyakit & Jondeung, 2012). Duong et al. (2016) reported COI mtDNA barcode to distinguish three morphological confusing pangasiid species (*Pangasius elongatus, P. krempfi and P. mekongensis*). Tran et al., (2017) applied mtDNA markers (COI and Cyt b) to differentiate and constructed phylogenetic trees of nine species of family Pangasiidae in Mekong Delta, of which *P. elongatus* was not included.

Generally, Siluriformes has supported by reliable anatomical synapomorphies as monophyletic group, however, relationships among families, genera, and species are not sufficiently resolved. This study aims to investigate the taxonomic and phylogenetic relationships within Siluriformes mainly focus on species distributing in LMB within three countries: Laos, Cambodia, and Vietnam using 16S rRNA and COI mitochondrial genes. Taxonomic characteristics was strongly considered when used up for species identification as well as evolutionary phylogeny.

MATERIALS AND METHODS

Sample Collection and Morphological Identification

Catfish species were collected at local markets or directly from the fisherman along Bassac and Mekong Rivers, Mekong Delta, Vietnam from 2014 - 2016. Additional samplings were conducted in Cambodia (Tonle Sap and Mekong main stem Kratie and Stung Treng), and Laos (Paske) from 2016 - 2017. The specimens were initially identified based on morphology following Rainboth (1996) and Tran et al. (2013). Muscle tissues (or fin clips) were preserved in 95% ethanol or DNA/RNA Shield™ (Zymo Research Corp., USA). To get the high-quality DNA, preserved fish tissue was kept at – 40^{o}C until analysis.

DNA Extraction, PCR Amplification and Sequencing

Total genomic DNA was isolated from preserved fish tissue (approximately 25 mg of muscle or fin clips) using GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific) following to the manufacturer's instructions. The 16S rDNA region was amplified using the primers 16Sar (5'-GCCT GTTTAACAAAAACAT-3') and 16Sbr (5'-CCGGTCTGAACTCA GATCATGT-3') (Palumbi et al., 1996). The COI mtDNA was amplified with the primers HCO (5'-TAAACTTCGGGTGACCAAA-3') and LCO (5'-GGTCAACAAATCATAAAGATA-3') (Folmer et al., 1994); FISHF1 (5'-TCAACC AACCACAAAGACATTGGCAC-3') and FISHR1 (5'-TAGACTTCTGGGTGGCCAAAGAATCA-3') (Ward et al., 2005).

PCR reactions were performed using a total volume of 50 µl with components as following concentration: 10 µL of Dream Taq buffer 10X, 2 µL dNTP (10 mM), 2 µL each primer (10 mM), 1.25 unit of Dream Taq polymerase (5U/µl), 5 µl DNA template and distilled water to the final volume.

Amplification was implemented using the following PCR profile: a preliminary denaturation at 94 °C for 3 mins, followed by 35 cycles of 94 °C for 45 sec, annealing for 45 sec (for 16S, CO1 gene at 48 °C, 42 ^{o}C, respectively), and then 72 °C for 45 sec. This was followed by a final extension period at 72 °C for 7 min before the samples were cooled to 4 °C. PCR products were run on 1.5% agarose gel for confirmation of equal length against an appropriate size markers.

The PCR products were purified using DNA purification kits (Promega) and pre-sequenced using dye – labels dideoxy terminator (Big Dye Terminator 3.1, Applied Biosystems) with the same primer as the PCR reactions at the following temperatures: 96 ºC for 30 sec, 50 ºC for 30 sec and 60 ºC for 4 min. Sequences of both strands were generated on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) using the amplification primers.

Phylogenetic Analyses

Sequence contigs were assembled using the Geneious Pro 5.5.7 (Kearse et al., 2012). The resulting sequences were confirmed by the Basic Local Alignment Search Tool (BLAST, https:// blast.ncbi.nlm.nih.gov/Blast.cgi). All DNA sequences generated from this study have been submitted to GenBank (**Table 1**). The sequences were aligned and analyzed using BioEdit 7.0.5.3 (Hall, 1999).

Phylogenetic trees were constructed using three different data sets. Xia test was implemented for each date set in DAMBE 6.4.101 (Xia et al. 2003) to check for substitution oversaturation based on the concept of entropy information theory. The first and second data sets were 30 partial 16S rRNA and COI mtDNA sequences from this study and 16 sequences retrieved from GenBank, respectively. According to Goncharow et al., (2004), combined gene analysis enhanced phylogenetic resolution, third data set was combined 16S and COI of current and Genbank available sequences using Geneious Pro 5.5.7 (Kearse et al., 2012). In PAUP*4.0b10 (Swofford, 2002), we used the incongruence-length difference (ILD) test (Farris et al., 1995) with 1000 randomized replicates to estimate any difference in phylogenetic signal among the different molecular sections.

Phylogenetic trees were constructed using 3 approaches, i.e., Neighbour Joining (NJ), Bayesian inference (BI) and Maximum likelihood (ML). All trees inferred from the partial 16S rDNA, CO1 sequences and combined 16S and COI were rooted with *Monopterus albus* as outgroup taxa.

NJ analyses were conducted based on the genetic distance between a pair of sequences, using the Kimura 2-parameter (K2P) model from MEGA 6 (Tamura et al.,2013). The support of clades was tested by the bootstrap method containing 1000 replicate.

Prior to ML and BI analyses, best-fit models of nucleotide substitution were selected by the Akaike Information Criterion as implemented by Modeltest 3.7 (Posada & Crandall, 1998) and MrModeltest 2.2 (Nylander, 2004). BI analyses were conducted in MrBayes 3.2.6 (Huelsenbeck & Ronquist, 2001) under the selected best-fit models and parameters. Four chains were used and the analysis was run for 1 million generations, with the sampling frequency of 100. Each analysis was repeated twice to check for similarity of the likelihood plateau. Additionally, parameter values were evaluated for convergence throughout the run by using the ‘‘sump’’ command in MrBayes and by examining results in Tracer 1.3 (Drummond et al., 2005). Plots from Tracer were used to determine the appropriate number of trees to be discarded in the ‘‘burn in’’ and a final 50% majority-rule consensus tree was constructed from the remaining trees. Numbers at the interior branches of the majority-rule consensus tree represent posterior probability (PP). ML analyses were conducted using PAUP*4.0b10 (Swofford, 2002). Trees were found by 1,000 replicate heuristic searches using the tree-bisection reconnection (TBR) branch-swapping algorithm, with 10 starting trees obtained by stepwise addition. Tree display and editing were performed in TreeView 1.6.6 (Page, 1996).

Table 1: Species list, sampling localities, GenBank accession for all sequences of catfish used in the phylogenetic analysis (*sequences from Genbank). Family classification following Nelson (2006).

No.	Family	Species	GenBank accession number		Collection site
			16S rRNA	COI	
1	Pangasiidae	*Pangasius krempfi*	MG076881	MG981062	Can Tho, Vietnam
2		*Pangasius macronema*	MG076882	MG981063	Can Tho, Vietnam
3		*Pangasius* sp.	MG076883	MG981064	Can Tho, Vietnam
4		*Pangasius sanitwongsei*	MG076884	MG981065	Strung Treng, Cambodia
5		*Pangasius larnaudii*	MG076885	MG981066	An Giang, Vietnam
6		*Pangasius conchophilus*	MG076886	MG981067	Can Tho, Vietnam
7		*Pangasius borcouti*	MG076887	MG981068	Dong Thap, Vietnam
8		*Pangasius elongatus*	MG076888	MG981069	Strung Treng, Cambodia
9		*Pangasianodon hypophthalmus*	MG076889	MG981070	Can Tho, Vietnam
10		*Pangasianodon gigas**	HM355768	KY118586	Thailand
11		*Helicophagus leptorhynchus*	MG076890	MG981071	Can Tho, Vietnam
12		*Helicophagus waandersii**	KR349217	KP036417	Malaysia
13		*Pseudolais pleurotaenia*	MG076891	MG981072	Pakse, Laos
14		*Pseudolais micronemus**	KR349237	KU692816	Indonesia
15	Clariidae	*Clarias macrocephalus*	MG076892	MG981073	An Giang, Vietnam
16		*Clarias batrachus*	MG076893	MG981074	Kratie, Cambodia
17		*Clarias fuscus*	MG076894	MG981075	Can Tho, Vietnam
18		*Clarias gariepinus**	KT001082	KT001082	China
19	Bagridae	*Mystus mysticetus*	MG076895	MG981076	Dong Thap, Vietnam
20		*Mystus micracanthus*	MG076896	MG981077	Dong Thap, Vietnam
21		*Mystus nigriceps*	MG076897	MG981078	Can Tho, Vietnam
22		*Mystus atrifasciatus**	JQ248054	KF824798	Thailand
23		*Mystus multiradiatus**	HQ257353	JX177677	Thailand
24		*Hemibagrus nemurus*	MG076898	MG981079	Can Tho, Vietnam
25		*Hemibagrus spilopterus*	MG076899	MG981080	Can Tho, Vietnam
26		*Hemibagrus wyckioides*	MG076900	MG981081	Pakse, Laos
27		*Hemibagrus wyckii*	MG076901	MG981082	Can Tho, Vietnam
28		*Hemibagrus peguensis**	KT878169	KJ909352	India
29		*Pseudomystus siamensis*	MG076902	MG981083	An Giang, Vietnam
30	Loricariidae	*Pterygoplichthys pardalis*	MG076903	MG981084	An Giang, Vietnam
31		*Pterygoplichthys disjunctivus**	KJ533242	JF769356	Philippines
32	Siluridae	*Ompok bimaculatus*	MG076904	MG981085	An Giang, Vietnam
33		*Ompok pabda**	GQ469559	FJ229987	India
34		*Ompok malabaricus**	FJ432679	HQ009494	India
35		*Wallago attu*	MG076905	MG981086	Vinh Long, Vietnam
36		*Kryptopterus bicirrhis*	MG076906	MG981087	Strung Treng, Cambodia
37	Schilbeidae	*Laides hexanema*	EU490866	EU490866	India
38		*Clupisoma longianalis*	MG076907	MG981088	Pakse, Laos
39		*Clupisoma garua**	GQ357922	KX455904	India
40	Sisoridae	*Bagarius yarrelli*	MG076908	MG981089	Pakse, Laos
41		*Bagarius bagarius**	KT878056	KJ909351	India
42	Plotosidae	*Plotosus canius*	MG076909	MG981090	Can Tho, Vietnam
43		*Plotosus lineatus**	KJ533241	MG220589	China
44	Ariidae	*Netuma thalassina*	MG076910	MG981091	Can Tho, Vietnam
45		*Arius arius**	KX211965	KX211965	China
46		*Arius dispar**	KJ533187	KJ533143	Philippines

RESULTS

Dataset Characteristics

In total, 60 sequences were generated from 2 gene regions of 30 species belonging to 9 families of Siluriformes fish distributing in Mekong River (no samples from Akysidae were collected as reported from Mekong Delta (Tran et al., 2013), and Amplycipitidae and Heteropneusidae in LMB (Rainboth et al., 2012)). The aligned data contained unambiguous 555 bp and 551 bp of 16S rRNA and COI genes, respectively. All of the datasets passed Xia test (lss < lss.c, in which lss values were significantly lower ($P < 0.05$) than lss.c, indicating that there was no substantial substitution saturation. The concatenated sequences yielded 695 (62.8%) constant characters and 347 (31.4%) parsimony-informative characters out of 1106 characters. Models of evolution and parameters were separately estimated for the 16S rRNA (GTR+I+G), CO1 mtDNA (HKY+I+G), and combined 16S+CO1 (GTR+I+G) datasets **(Table 2)**. ILD test (PAUP 4.0b10) supported combination data sets ($P < 0.05$) (Cunningham, 1997). The appropriate model for each gene and alignment characteristics were summarized in **Table 2**.

Table 2: Basic sequence alignment characteristics and parameter values under the best-fit models selected by the Akaike Information Criterion for different data sets.

Parameter	16S rRNA	CO1 mtDNA	16S + COI
Number of species	47	47	47
Number of aligned sites	555	551	1106
Constant sites	383	315	698
Parsimony uninformative	45	19	64
Parsimony informative	127	217	344
Best fit model	GTR+I+G	HKY+I+G	GTR+I+G
Nucleotide equilibrium frequency μ (A, C, G, T)	0.3487; 0.2309 0.1929; 0.2275	0.3030; 0.3131 0.1060; 0.2780	0.3316; 0.2783 0.1436; 0.2465
Proportion of invariant sites (I)	0.5000	0.5496	0.5763
Gama distribution shape parameter	0.3754	0.9241	0.7060

Phylogenetic Analysis

Catfish Phylogeny

16S data set showed low resolution for phylogeny, especially the taxonomic position of Sisoridae (Supplementary Figure S1), then further excluded from phylogenetic analysis.

Different analysis methods performed divergent topologies with various data sets (**Table 3).** The BI and NJ methods showed more similarity results than results obtained from ML. The greatest similarity between methods and data sets is: i) Siluriformes was resolved monophyletic; ii) 2 lineages were discovered from all topologies: basal Lineage I includes the originated Africa-exotic Asia Loricaridae (suborder Loricarioidei), Lineage II consists the remaining families (suborder Silurioidei) (Betancur-R et al., 2013), iii) in the lineage II, Plotosidae was discovered as sister clade to the rest of families of Silurioidei. The most obvious difference was the arrangement between inter-families among current silurioidei species.

Table 3: Phylogenetic topologies of Lower Mekong Basin catfish families using different phylogenetic analyses for COI and combined 16S+COI data sets. Italic indicated for sister clade families. Family classification was followed the descriptions of Nelson (2006).

Gene	Maximum likelihood	Bayesian Inferences	Neighbor Joining
COI mtDNA	[Loricariidae]/[{Plotosidae}+{Clariidae+Sisoridae+Ariidae+(*Siluridae*+*Bagridae*)+Schilbeidae+Pangasiidae}]	[Loricariidae]/[{Plotosidae}+{Ariidae+(*Sisoridae*+*Clariidae*+*Siluridae)*+(*Schilbeidae*+*Bagridae*)+Pangasiidae}]	[Loricariidae]/[{Plotosidae}+{((*Clariidae*+*Sisoridae*+*Ariidae*)+(*Siluridae*))+Bagridae+Schilbeidae+Pangasiidae}]
COI+16S mtDNA	[Loricariidae]/[{Plotosidae}+{(*Ariidae*+*Schilbeidae*)+Pangasiidae+Bagridae+Clariidae+(*Sisoridae*+*Siluridae*)}]	[Loricariidae]/[{Plotosidae}+{Ariidae+Schilbeidae+Bagridae+(*Sisoridae*+*Clariidae*+*Siluridae*)+Pangasiidae}]	[Loricariidae]/[{Plotosidae}+{Siluridae+Sisoridae+Ariidae+Clariidae+(*Bagridae*+*Schilbeidae*)+Pangasiidae}]

With COI data set, NJ topology showed that Clariidae, Sisoridae and Ariidae were grouped together, and sister to Siluridae. Bagridae, Schilbeidae and Pangasiidae were all established as strong monophyletic groups. BI was different with Sisoridae, Clariidae, Siluridae, and Schilbeidae, Bagridae were clustered together, while Ariidea and Pangasiidae appeared as separate clades. ML tree showed only Siluridae and Bagridae were grouped together (**Figure 1**).

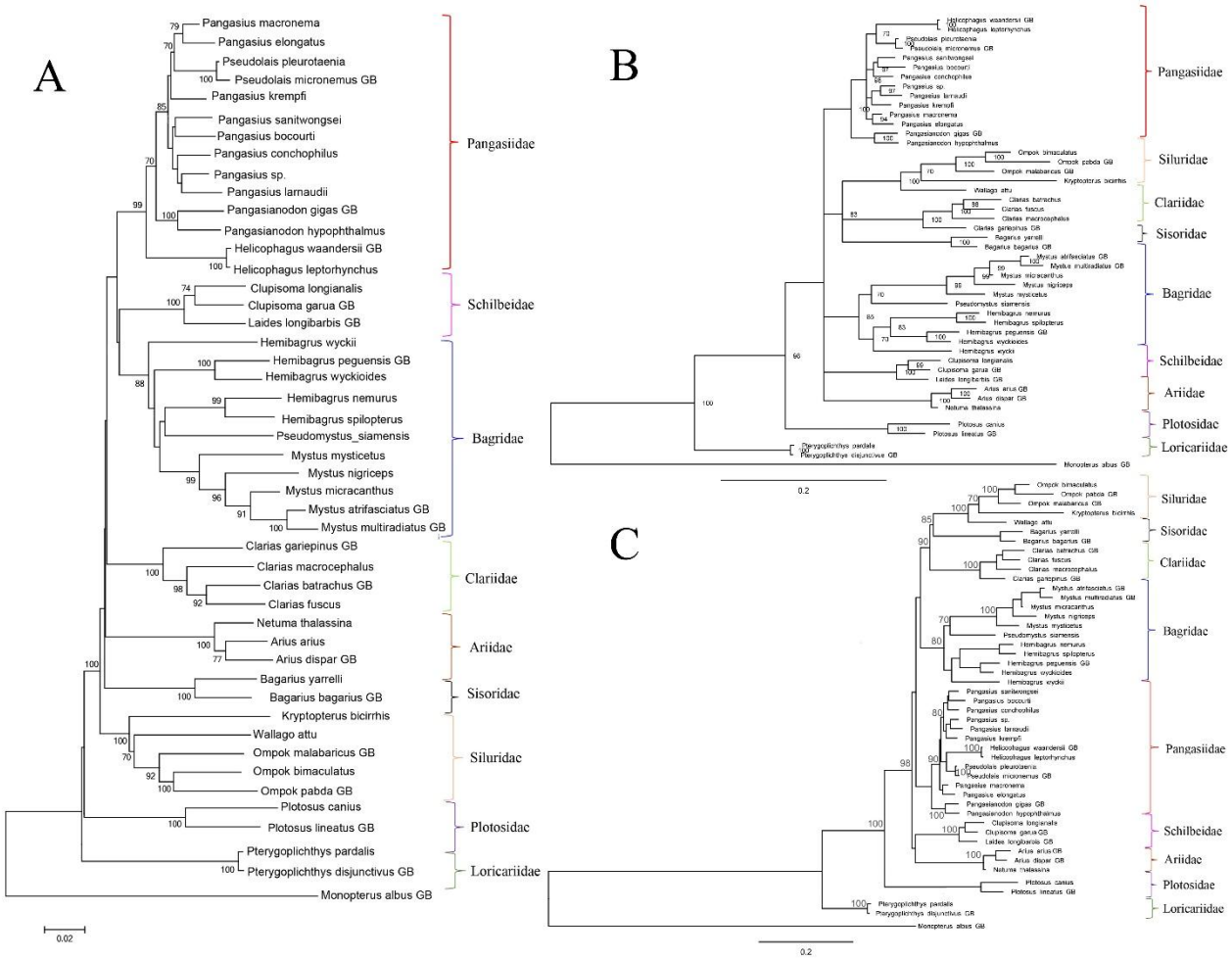


Figure 1: Phylogenetic tree of Siluriformes resulting from the Neighbor Joining (A), Bayesian inference (B) and Maximum Likelihood (C) analysis of COI mtDNA data set. Family classification was followed the descriptions of Nelson (2006).

For the concatenated data, instead of clustered together, Siluridae, Sisoridae, Ariidae and Clariidae were resolved as monophyletic and separated clades, while Bagridae and Schilbeidae (NJ), and Sisoridae, Clariidae, Siluridae (BI) were grouped together. In ML analysis, Schilbeidae was placed closely to Ariidae and Sisoridae was grouped with Siluridae, Pangasiidae, Bagridae, and Clariidae were placed close together in the phylogenetic tree (**Figure 2**).

Monophyly of Catfish Family

At genus level, among nine studied families, seven were well resolved as monophylectic clade. Two rich species families are Pangasiidae (14 species, 4 genera); Bagridae (11 species, 3 genera) displayed uncertain taxonomic positions. Within Pangasiidae, all analyses placed *Pangasianodon* spp. as the sister species to the rest of this clade. While *Helicophagus*, *Pseudolaias* and *Pangasianodon* were recovered with high support as monophyletic group, *Pangasius* spp. appeared to be paraphyletic (grouped with *Helicophagus* and *Pseudolaias* (**Figure 1 & 2**)). Within Bagridae, the phylogenetic position of *Pseudomystus* was not determined, it was grouped with *Hemibagrus* species in the analysis of COI data set (**Figure 1**), while clustered to the genus *Mystus* (analysis ML and BI of 16S and CO1 data set, **Figure 2**).

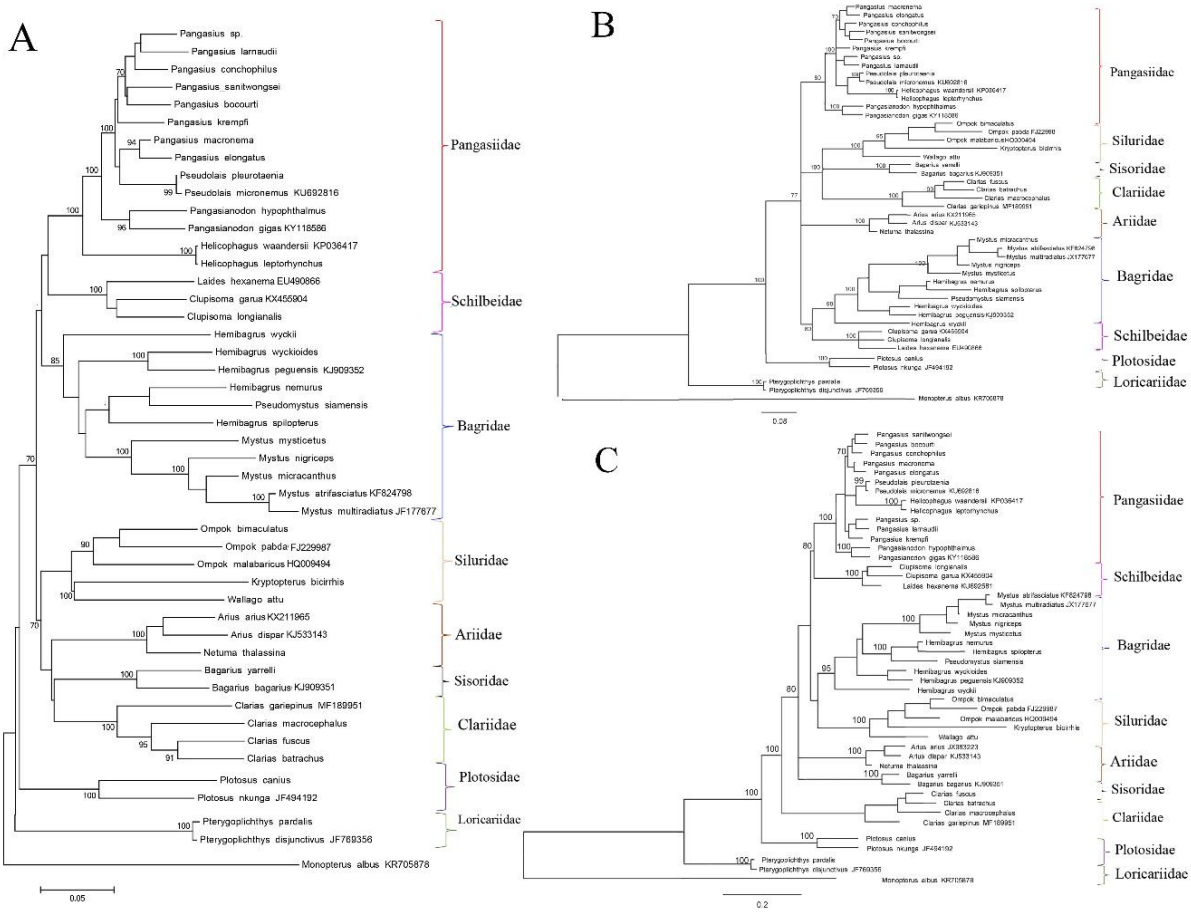


Figure 2: Phylogenetic tree of Siluriformes resulting from the Neighbor Joining (A), Bayesian inference (B) and Maximum Likelihood (C) analysis of 16S rRNA and COI mtDNA datasets. Family classification was followed the descriptions of Nelson (2006).

DISCUSSION

Siluriformes was among the most ecologically and economically important fish in Mekong River Basin (Nelson, 2006; FAO, 2014). From an evolutionary perspective, the order Siluriformes has been used as the model system to investigate the historical biogeography (Kappas et al., 2016). Related continental multi-families have been detected when distinguished gene regions were used (Hardman, 2005; Sullivan et al., 2006; Kappas et al., 2016), such as "Big Asia" clade (including Bagridae, Horabagridae, Akysidae, Amblycipitidae, Sisoridae, Erethistidae), and "Big Africa" (constituting Mochokidae, Malapteruridae, Amphiliidae, Claroteidae, Schilbeidae). Phylogenetic positions of families such as Siluridae, Schilbeidae, Malapteruridae, Bagridae, Mochokidae and Plotosidae remained undefined (Hardman, 2005; Sullivan et al., 2006).

This study found some common ground with other studies, such as the sister relationship of Loricariidae (suborder Loricarioidei) with the remaining families of Siluroidei (Sullivan et al., 2006; Kappas et al., 2016), and the certain family status of Clariidae (Jansen et al., 2006), Bagridae (not include *Rita*) (Jansen et al., 2006), and Pangasiidae (Azlina et al., 2013; Tran et al., 2017). However, the inter-familial relationships were neither consistent among the analyses, nor were there any similarities with other studies.

Following Nelson (2006) and Rainboth et al. (2012), five common suborders have been clarified such that Loricarioidea (Loricariidae), Sisoroidea (Sisoridae), Doradoidea, Siluroidea (Siluridae, Plotosidae, Clariidae), and Bagroidea (Ariidae, Schilbeidae, Pangasiidae and Bagridae). Sullivan et al. (2006) and Kappas et al. (2016) created suborder Clarioidea (Clariidae and Heteropneustidae), and Arioidea

(Ariidae and Anchariidae), and placed Siluridea, and Pangasiidae on the list of unresolved families. An important finding is that "Big Asia" and "Big Africa" clades include multi families related to continental distribution. Bagridea and Sisoridea were grouped together in "Big Asia", while Schilbeidae clustered in "Big Africa" (**Figure 3**). Betancur-R et al. (2013) suggested sister relationships of Loricarioidei and Diplosistoidei (not included in current research); among Siluroidei, Sisoridea was sister clade to Bagridae/Schilbeidae, Pangasiidae clustered to Anchariidea/Ariidae and Siluridea was grouped to Plotosidae.

Our study supported family monophyly of Siluriformes, however, the arrangement of inter-families did not follow the suborder classification as reported by Nelson (2006) and Rainboth et al. (2012). Plotosidae was strongly supported as basal clade from other remaining of Siluroidei, Loricarioidea was sister clade to Siluroidea. Additionally, we did not find Asia/Africa distribution pattern as Bagridae and Sisoridae were not arranged together in all analyzes as they clustered in Big Asia clade, Schilbeidae (Big Africa) was mostly placed closed to Bagridae, and/or Pangasiidae (Sullivan et al., 2006). Due to topology divergent within different methods used, none of the family/suborder arrangements was concordant with those reported by Betancur-R et al. (2013).

In total, 13 pangasiid species were reported distributed in LMB that belong to 4 genera: *Pangasianodon*, *Pangasius*, *Pseudolais*, and *Helicophagu*s. In this study, 14 Pangasiidae (one unidentified *Pangasius* sp.) species presented by four above taxonomic diagnostic genera (Ferraris et al., 2007) was well monophyletic resolved. *Pangasianodon* was discovered as sister taxon to the rest of pangasiid in most analyses (except COI data set, NJ analysis), while *Pangasius* species were clustered in the same clade to *Helicophagus* and *Pseudolais* species (**Figure 1 and 2**). As the genus level, *Pangasius* (Pangasiidae) was previously reported as paraphyletic (Pouyaud et al., 2004; Azlina et al., 2013), 2 subgroups dividing (Tran et al., 2017), and/or monophyletic (Kappas et al., 2016; Sullivan et al., 2006; Karinthanyakit & Jondeung, 2012). Gustiano et al. (2003) resolve problem with *Pangasius kunyit* species complex, and described 2 new species (*P. sabahensis* from Malaysia, and *P. mekongensis* from Mekong Delta, Vietnam). Tran et al. (2017) has reported limited morphologic characters to distinguish *Pangaius* species. However, Karinthanyakit & Jondeung (2012) applied mitochondrial DNA markers (cyt b, 12S rRNA, tRNA-Val and 16S rRNA) detecting four supported genus - clades, of which *Pangasianodon* was strongly supported as the most basal taxon within pangasiids, whereas *Pseudolais* and *Helicophagus* were recovered as a sister group of *Pangasius.* In contrast, Pouyaud et al. (2004), and recent study shared the same finding of uncertainty in the arrangement of *Pangasius* species.

Bagridae is among rich species genus of Siluriformes of which 210 have been valid (19 extant genera) out of 314 nominal species (Armbruster, 2011; Ferraris et al., 2007). Family status was strongly supported if Africa *Rita* species was excluded. Bagridae was involved in "Big Asia" clade along with Horabagridae, *Ailia* and *Laides*, Akysidae, Amblycipitidae, Sisoridae, and Erethistidae (Sullivan et al., 2006; Kappas et al., 2016). Peng et al. (2004) was also recovered bagrid catfishes form a monophyletic clade. Among 4 genera, *Mystus* and *Pseudobagrus* were monophyletic groups, while the remaining genus *Pelteobagrus* and *Leiocassis* are complicated. Dodson & Lecomte (2015) revealed the phylogeny of *Hemibagrus*, and reported several species group which distributed in Southeast Asia. With three genera examined, *Hemibagru*s showed the same pattern as paraphyletic group when placed close to *Pseudomystus*, and then sister taxon to monophyletic *Mystus* genus.

Our mtDNA data show unequivocal support for monophyly of the individual families, and the suborder Loricarioidei and Siluroidei as well. Plostosidae was placed as basal clade of Siluroidei, while traditional suborders as well as recent supported clades (Big Asia, Big Africa) were not recovered herein. It seems that the synapomorphic taxonomic characters supported diagnostic families of Mekong fish, however, additional taxon and markers, as well as divergent time estimate need to be conducted to better understand the diversity and evolutionary radiation of fish in the cradle biodiversity of the Mekong River.

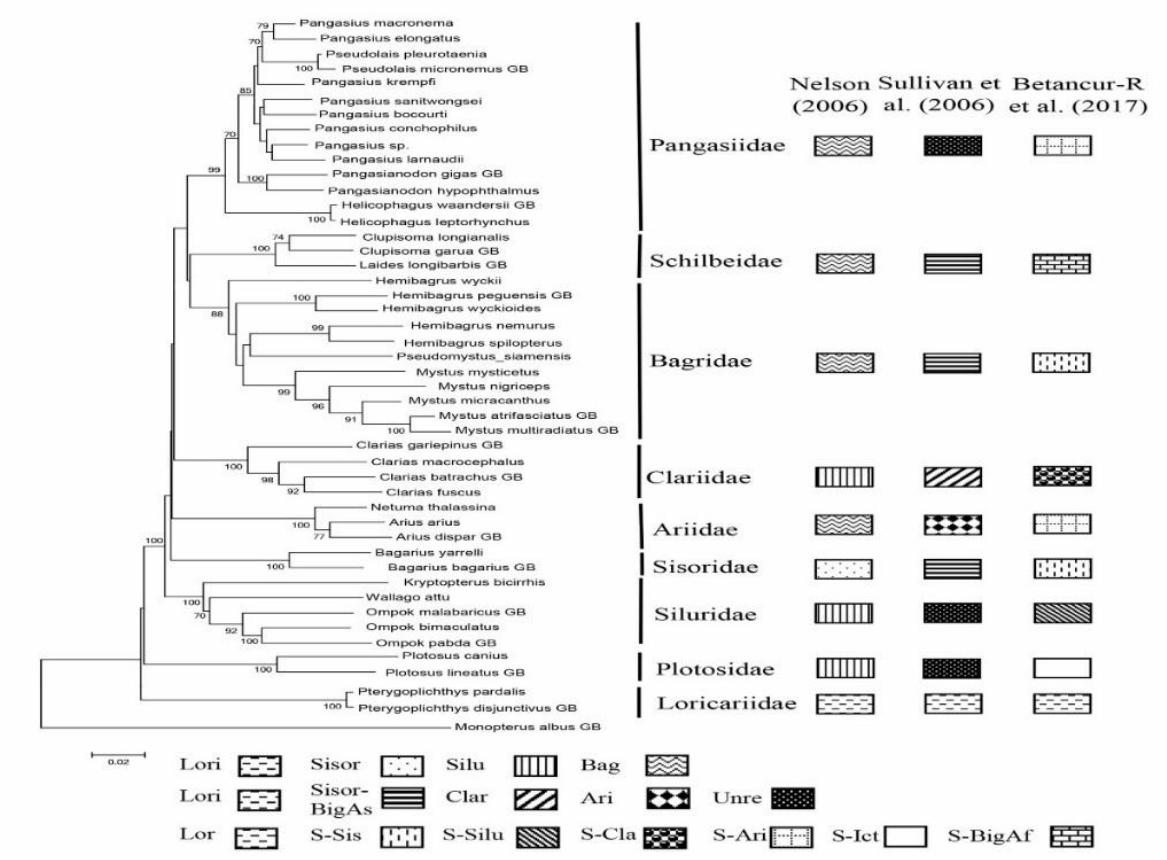
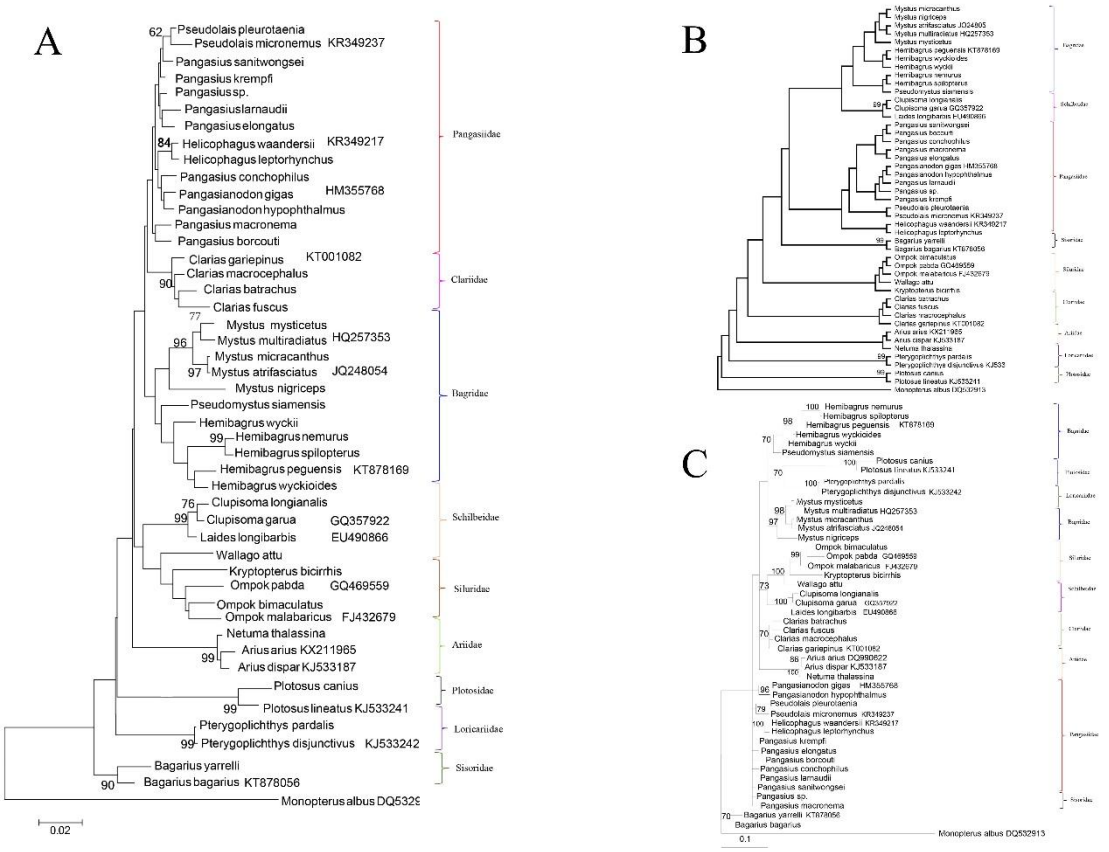


Figure 3: Comparative phylogenetic topologies from current study (MJ tree from combined 16S and COI datasets, bootstrap value with 1000 replicates), and recently reported phylogenies based on morphologic characters (Nelson, 2006), Rag sequencing data (Sullivant et al., 2006), and molecular phylogeny (Betancur-R et al., 2017).

SUPPLEMENTARY



Supplementary Figure S1: Phylogenetic tree of Siluriformes resulting from the Neighbor Joining (A), Bayesian inference (B) and Maximum Likelihood (C) analysis of 16S rRNA data set. Family classification was followed the descriptions of Nelson (2006).

ACKNOWLEDGEMENTS

We thank Mekong Delta communes who help us collect fish samples from market. We are very grateful to Prof. Kent Carpenter from ODU for assistance in fish identification. This project was funded by the PEER 2-7 Project

REFERENCES

1. Allen, D. J., Smith, K. G., & Darwall, W. R. (2012). *The status and distribution of freswater biodiversity in Indo - Burma*. The IUCN Red List of Threatened SpeciesTM, Cambridge, UK and Gland.
2. Armbruster, J. W. (2011). Global Catfish Biodiversity. *American Fisheries Society Symposium* 77:15–37.
3. Azlina, Z. M., Daud, S. K., Siraj, S., Aliabadian, M., & Moghaddam, F. (2013). Molecular phylogeny of Malaysian Pangasiid based on cytochrome C oxidase I (COI) DNA sequences. *Iranian Journal of Animal Biosystematics* 9(January): 17–28.
4. Baran, E., & Guerin, E. (2012). River sediments and coastal fish production: What about the MeKong?. Report for the Project "A Climate Resilient Mekong: Maintaining the Flows that Nourish Life" led by the Natural Heritage Institute. World Fish Center, Phnom Penh, Cambodia.
5. Betancur-R, R., Li, C., Munroe, T. A., Ballesteros, J. A., & Orti, G. (2013). Addressing gene tree discordance and non-stationarity to resolve a multi-locus phylogeny of the flatfishes (Teleostei: Pleuronectiformes). *Systematic Biology* 62(5): 763–785. doi:10.1093/sysbio/syt039

6. Cunningham, C. W. (1997). Can three incongruence tests predict when data should be combined? *Molecular Biology and Evolution* 14(7): 733–740. https://doi.org/10.1093/oxfordjournals.molbev.a025813
7. Diogo, R. (2003). Higher-level Phylogeny of Siluriformes.An Overview. In A. Gloria, B. G. r Kapoo, M. Chardon, & R. Diogo (Eds.), *Catfishes*. Science Publishers.
8. Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22(5): 1185–1192. doi:10.1093/molbev/msi103
9. Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z., Naiman, R. J., Knowler, D. J., & Le, C. (2006). Freshwater biodiversity : importance, threats, status and conservation challenges. *Biological Reviews* 81(2006): 163–182. doi:10.1017/S1464793105006950
10. Duong, T. Y., Nguyen, K., Bui, S. N., Nguyen, V. T., Loan, N. B., & Tran, D. D. (2016). DNA barcodes and morphology of *P. krempfi*, *P. mekongensis*, and *P. elongatus*. *Vietnam Journal of Biotechnology* 14(1): 29 - 37. doi: 10.15625/1811-4989/14/1/9289
11. Eschmeyer, W. N., & Fong, J. D. (2010). *Species by family/subfamily in the Catalog of Fishes*. California Academy of Sciences. Retrieved from http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp
12. Farris, J. S., Källersjö, M., Kluge, A. G., & Bult, C. (1995). Constructing a Significance Test for Incongruence. *Systematic Biology* 44(4): 570–572. doi:10.1093/sysbio/44.4.570
13. Ferraris, M., Carl, J., Miya, M., Azuma, Y., & Nishida, M. (2007). *Checklist of catfishes, recent and fossil (Osteichthyes: Siluriformes), and catalogue of siluriform primary types (Zootaxa 1418)* (1st ed.). Magnolia Press.
14. Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology 3*(5): 294–299. Retrieved from http://europepmc.org/abstract/MED/7881515
15. Gustiano, R., Teugels, G. G., & 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. doi:10.1080/713834687
16. Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.
17. Hardman, M. (2005). The phylogenetic relationships among non-diplomystid catfishes as inferred from mitochondrial cytochrome b sequences; the search for the ictalurid sister taxon (Otophysi: Siluriformes). *Molecular Phylogenetics and Evolution* 37(3): 700–720. doi:10.1016/j.ympev.2005.04.029
18. Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8): 754–755. doi:10.1093/bioinformatics/17.8.754
19. Jansen, G., Devaere, S., Weekers, P. H. H., & Adriaens, D. (2006). Phylogenetic relationships and divergence time estimate of African anguilliform catfish (Siluriformes: Clariidae) inferred from ribosomal gene and spacer sequences. *Molecular Phylogenetics and Evolution* 38(1): 65–78. doi:10.1016/j.ympev.2005.09.011
20. Kailola, P. (2004). A Phylogenetic Exploration of the Catfish Family Ariidae (Otophysi: Siluriformes) [online]. *The Beagle: Records of the Museums and Art Galleries of the Northern Territory* 20: 87–166.
21. Kappas, I., Vittas, S., Pantzartzi, C. N., & Drosopoulou, E. (2016). A Time-Calibrated Mitogenome Phylogeny of Catfish (Teleostei : Siluriformes). Plos One 11(12): 1–16. doi:10.1371/journal.pone.0166988
22. Karinthanyakit, W., & Jondeung, A. (2012). Molecular phylogenetic relationships of pangasiid and schilbid catfishes in Thailand. *Journal of Fish Biology* 80(7): 2549–2570. doi:10.1111/j.1095-8649.2012.03303.x
23. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)* 28(12): 1647–1649. doi:10.1093/bioinformatics/bts199
24. Lundberg, J. ., Sullivan, J. ., Rodiles-hernández, R., & Dean, A. (2007). Discovery of African

roots for the Mesoamerican Chiapas catfish, *Lacantunia enigmatica*, requires an ancient intercontinental passage. Proceedings of the Academy of Natural Sciences of Philadelphia 156(1): 39–53.
25. Nelson, J. S. (2006). *Fishes of the World*. John Wiley & Sons, Inc.doi: 978-0-471-25031-9
26. Nylander, J. A. A. (2004). MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
27. Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12(4): 357–358.
28. Palumbi, S., Romano, S., Mcmillan, W. O., & Grabowski, G. (2002). The Simple Fool's Guide To PCR, 96822(808), 1–45.
29. Peng, Z., He, S., & Zhang, Y. (2004). Phylogenetic relationships of glyptosternoid fishes (Siluriformes: Sisoridae) inferred from mitochondrial cytochrome b gene sequences. Molecular Phylogenetics and Evolution 31: 979–987. doi:10.1016/j.ympev.2003.10.023.
30. Poulsen, A. F., Hortle, K. G., Booth, S. J., Visser, T. A. M., Viravong, S., Bouakhamvongsa, K., … Tran, B. Q. (2004). *Distribution and ecology of some important riverine fish species of the Mekong River Basin*. MRC Technical Paper No. 10, ISSN: 1683-1489.
31. Pouyaud, L., Gustiano, R., & Teugels, G. (2004). Contribution to the phylogeny of the Pangsiidae based on mitochondrial 12 rDNA. *Indonesia Journal of Agriculture Science* 5(2): 45-62.
32. Rainboth, W., Vidthayanon, C., & Mai, D. Y. (2012). *Fishes of the Greater Mekong Ecosystem with Species List and Photographic Atlas*. *Miscellaneous publication*. Museum of Zoology, University of Michigan.
33. Rainboth, W. J. (1996). *Fishes of the Cambodian Mekong*. FAO Species Identification Field Guide for Fishery Purposes. FAO, Rome.
34. Sullivan, J. ., Lundberg, J. ., & Hardman, M. (2006). A phylogenetic analysis of the major groups of cat W shes (Teleostei : Siluriformes) using rag1 and rag2 nuclear gene sequences. *Molecular Phylogenetics and Evolution* 41: 636–662. doi:10.1016/j.ympev.2006.05.044
35. Swofford, D. L. (2002). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10. Sinauer Associates. Doi: 10.1111/j.0014-3820.2002.tb00191.x
36. Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, 30(12), 2725–2729. doi:10.1093/molbev/mst197.
37. Thai, N. T. (2016). *Study on fishes biodiversity in Mekong Delta, Vietnam and its variation due to Climate change and Socio-economic development*. PhD Thesis in Institute of Tropical Biology, Vietnam Academy of Science and Technology, Vietnam.
38. Tran, D. D., Nguyen, T. P., Ha, P. H., Tran, X. L., Mai, V. H., Kenzo, U., & Koichi, S. (2013). *Fishes of the Mekong Delta, Vietnam*. Can Tho University Publishing House.
39. Tran, H. T. T., Tran, T. N., Tran, H. N. A., & Nguyen, H. T. (2017). DNA Barcoding and Phylogenetic Relationships of Nine Catfish Species from Mekong Basin, Vietnam. *Journal of Molecular Biomarkers & Diagnosis* 8(6). doi:10.4172/2155-9929.1000363
40. Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360(1462): 1847-1857.
41. Yu, S. C. S., & Quilang, J. P. (2014). Molecular Phylogeny of Catfishes (Teleostei : Siluriformes) in the Philippines Using the Mitochondrial Genes COI , Cyt b , 16S rRNA , and the Nuclear Genes Rag1 and Rag2. Philippine Journal of Science 143(2): 187–198.