CYTOCHROME OXIDASE I (COI) DIVERGENCE ASSESSMENT OF FAMILY NEMIPTERIDAE FROM MALAYSIAN WATERS

LIEW YOU EN[1], SALWANI ABDULLAH[1], TAN MIN PAU[1], MAZLAN ABD GHAFFAR[2], ALIAS MAN[3] AND TUN NURUL AIMI MAT JAAFAR[1]*

[1]*School of Fisheries and Aquaculture Sciences, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu.*
[2] *School of Marine and Environmental Sciences, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia.*
[3]*Department of Fisheries Malaysia, Ministry of Agriculture and Agro-Based Industry, Federal Goverment Administrative Centre, 62628 Putrajaya, Malaysia*

Corresponding author: tun_aimi@umt.edu.my

Abstract: DNA Barcoding, primarily focusing on cytochrome *c* oxidase subunit I (*COI*) gene has been appraised as an effective tool for species identification. Nonetheless, species identification based on molecular approach is essential for discrimination of look-alike species. In this study, we focused on the marine fishes Family Nemipteridae, one of the commercially important group distributed within the surrounding seas of Malaysia. Some of the samples were collected during the National Demersal Trawl Survey in the Exclusive Economic Zone of East Coast Peninsular Malaysia by the Department of Fishery Malaysia. A 652bp region of *COI* was sequenced for 74 individuals from nine putative species. Additional 34 *COI* sequences from GenBank were also included in this study making the total number of samples analysed to 108 individuals. The average Kimura 2-parameter (K2P) nucleotide divergence was 0.34% among individuals within species and 6.97% within genera. All putative species formed monophyletic clades in both Neighbour-joining (NJ) and Maximum-likelihood (ML) trees. However, there was a potential misidentification in specimen identified as *Nemipterus tambuloides,* as the specimen did not group with their own taxa. It was genetically grouped in *Nemipterus thosaporni* clade. This study supports the effectiveness of *COI* gene in species discrimination of Family Nemipteridae.

Keywords**:** COI, DNA barcoding, Nemipteridae

Introduction

In Malaysia, the Family Nemipteridae is one of the commercially important fish groups. It is known that the fishes from the genus *Nemipterus* are caught throughout the year and are very widely consumed by the Malaysian population (Imtiaz *et al*., 2016). They are considered as the marine, bottom-living fishes (Russell, 1990). The colour of the fish species from Family Nemipteridae is extremely varied. However, they are often found in pinkish or reddish with red, yellow or blue marking (Russell, 1990). There are two similar families occurring in the area of habitat together with the Family Nemipteridae, which are the Lutjanidae and Caesionidae. The taxonomy of the Family Nemipteridae has been extensively described and cleared up on the basis of morphological characters (Russell, 1990; 1993). Nonetheless, a few taxonomic ambiguities still exist. Subtle differences may not be disclosed when depending completely on the external morphological characteristics. This could lead to the problem of misidentification of the species in Family Nemipteridae due to colour changes and body pattern formation after improper and long term storage. When one species is being classified wrongly, severe repercussion of over-exploitation may happen. This will lead to a great impact on the marine system as a whole, particularly the targeted species.

Today, the issues of misidentification have become a frequently happening incident at the fish trading places, particularly at the landing sites as well as the fish markets. Clear-cut identification of the fish is critical for the satisfaction, profit and well-being (Imtiaz *et al*., 2016). The problematic taxonomic issues need to be clarified so as to give broad knowledge to beginner (Seah *et al*., 2017). Different types of methods have been applied in the process of species identification, such as using the morphological identification and biochemical observation (Strauss & Bond, 1990). Today, genetic information, particularly the DNA sequence diversity has been appraised directly or indirectly through protein

analysis to discriminate the closely related species (Blaxter, 2003; Savolainen *et al*., 2005; Ward *et al*., 2005).

Depending on a short, standardised gene region, DNA barcoding is highly approached by the taxonomist for species identification (Hebert *et al*., 2003a; Hebert & Gregory, 2005; Aquilino *et al*., 2011; Mat Jaafar *et al*., 2012). The primary goals of the use of DNA barcoding are to determine the unknown samples to species level, as well as to build up the encounter of new species and facilitate the identification, particularly in the cryptic, microscopic and other organisms with complex or inaccessible morphologies (Hebert *et al*., 2003a; Hebert *et al*., 2004). The ultimate goal of the DNA barcoding evolution is the advancement of extensive barcode libraries for entire species on earth (Mat Jaafar *et al*., 2012). The application of DNA sequence used in the discrimination of species work similarly as the supermarket scanner by scanning the black stripes barcode to classify the products (Seah *et al*., 2017). Cytochrome *c* oxidase subunit I (*COI*) gene from mitochondrial DNA has been widely used as a practical, standardized species-level barcode for most of the animal kingdom. (Hebert *et al*., 2003b; Mat Jaafar *et al*., 2012). The advantages of using the *COI* are that it has a

shorter fragment to be sequenced, rapid and has shown effectiveness in the species discrimination (Mat Jaafar *et al*., 2012).

Materials and Methods
Data Sampling

A total of 74 Nemipteridae specimens were collected from Malaysian waters. The samples were collected from several locations during the National Demersal Trawl Survey in the Exclusive Economic Zone of East Coast Peninsular Malaysia by the Department of Fisheries, Malaysia, from May to July 2016 and some were collected from the West Coast of Peninsular Malaysia (Table 1). The specimens encompassed nine putative species from two genera, namely *Nemipterus* and *Scolopsis* from Family Nemipteridae. All of the fin tissue samples were stored in 96% ethanol and kept in 4°C freezer. All the fish specimens were placed in ice, frozen on site and deposited in Fisheries Science Laboratory, Universiti Malaysia Terengganu. All of the specimens collected were identified based on the external morphology, either with the help of expert local taxonomist or by referring to the fish identification book (Ambak *et al*., 2010; Matsunuma *et al*., 2011; Yoshida *et al*., 2013).

Table 1: Specimens of Family Nemipteridae collected from the Exclusive Economic Zone of East Coast Peninsular Malaysia and West Coast of Malaysia.

Species	n	Collection site
Nemipterus balinensoides	3	Kelantan
Nemipterus bathybius	3	Kelantan
Nemipterus marginatus	7	Pahang
	10	Kelantan
	8	Johor
Nemipterus nematophorus	1	Kelantan
	10	Terengganu
	10	Kelantan
	10	Melaka
Nemipterus nemurus	2	Johor
Nemipterus tambuloides	1	Kelantan
Nemipterus thosaporni	3	Kelantan
Nemipterus virgatus	3	Terengganu
Scolopsis taenioptera	3	Johor

DNA Extraction, PCR Amplification and Sequencing

Total genomic DNA of each specimen was extracted follwing the protocol of salt extraction method (Miller *et al*., 1988). The DNA were extracted from the right pectoral fin of each specimen. After the process of DNA extraction, the DNA stock was stored in 1.5µl microcentrifuge tubes with temperature of -20°C. Concentration and purities of the entire extracted DNA were checked using NanoDrop. The preferred purities for the A260/280 ratio ranged from 1.8 to 2.0.

Every sample had underwent the process of polymerase chain reaction in order to amplify a fragment of 652 base pairs (bp) of *COI*. The primer pair used were 5'-TCAACCAACCACAAAGACATTGGCAC-3' and 5'-TAGACTTCTGGGTGGCCAAAGAATCA-3' (Ward *et al*., 2005). PCR reaction mix of 25 µl consisted of 18.75 µl of ultrapure water, 2.25 µl of 10X PCR buffer, 1.25 µl of MgCl$_2$ (50 mM), 0.25 µl of each primer (0.01 mM), 0.125 µl of each dNTP (0.05 mM), 0.625 U of *Taq* polymerase, and 0.5-2.0 µl of DNA template. The thermal regime consisted of an initial step of 2 minutes at 95°C, followed by 40 cycles of 45 seconds at 94°C, 45 seconds at 50°C and 1 minute at 72°C, then followed in

turn by 10 minutes at 72°C (Imtiaz *et al.*, 2016). DNA amplification products were separated in 1.2% (w/v) agarose gel. The most intense products were sent for sequencing.

Data Analyses

Multiple alignments were performed for the obtained sequences in MEGA7 (Kumar *et al.*, 2016). In order to ensure that there was no presence of gaps and stop codons in the alignment, the amino acid translation was examined. Kimura's two-parameter model was used to calculate the nucleotide-sequence divergences. Based on

the K2P distances, the intraspecific and interspecific sequences divergences for all species sequences from more than two individuals were obtained (Wang *et al.*, 2011). Additional 34 *COI* sequences from GenBank were included in this study (Table 2). Neighbour-joining and Maximum-likelihood analyses were done using K2P distance model in MEGA7 (Kumar *et al.*, 2016). Phylogenetic trees were constructed with 10000 bootstrap replications to provide a graphic representation of the patterning of divergence between species.

Table 2: Sequences accessed from GenBank with species names and areas of collection sites.

Species	GenBank Accession Numbers	Collection Site
Nemipterus marginatus	KY362833.1	Perlis
	KY362834.1	Perlis
	KY362836.1	Sabah
	KY362837.1	Sabah
Nemipterus nematophorus	KY362838.1	Perlis
	KY362839.1	Perlis
	KY362848.1	Sabah
Nemipterus nemurus	KY362856.1	Sabah
	KY362857.1	Sabah
Nemipterus tambuloides	KY362874.1	Perlis
	KY362877.1	Sabah
	KY362878.1	Sabah
	KY362879.1	Sabah
Scolopsis taenioptera	LC198257.1	Terengganu
	LC198258.1	Terengganu
	LC198259.1	Terengganu
	LC198260.1	Terengganu
	LC198261.1	Terengganu
	LC198262.1	Terengganu
	LC198263.1	Terengganu
	LC198264.1	Terengganu
	LC198265.1	Terengganu
	LC198266.1	Terengganu
	LC198267.1	Terengganu
	LC198268.1	Terengganu
	LC198269.1	Terengganu
	LC198270.1	Terengganu
	LC198271.1	Terengganu
	LC198272.1	Terengganu
	KY362965.1	Terengganu
	KY362966.1	Perlis
	KY362967.1	Perlis
	KY362973.1	Sabah
	KY362974.1	Sabah

Results and Discussion
General Findings

In this study, a total number of nine species and two genera of Family Nemipteridae in Malaysia were analysed. The number of sequences per species varied between three (*N. balinensoides, N. bathybius, N. virgatus*) to 34 (*N. nematophorus*), the most abundant

species caught by trawlers. Therefore, a total of 108 *COI* sequences were acquired for Family Nemipteridae within this region. Among all the analysed sequences, there was no introns, deletion, heterozygous sites or stop codons observed, supporting the view that all the amplified sequences constitute functional mitochondrial *COI* sequences.

COI Divergence Assessment

A total number of 108 sequences (74 sequences from present study and 34 downloaded from GenBank) of nine species and two genera were calculated for the *COI* nucleotide divergences. The average K2P distance within species was 0.34% with far less 0.00% for *N. tambuloides*, *N. thosaporni* and *N. virgatus* (Table 3). On the other hand, the mean congeneric nucleotide distance was 6.97%. The mean K2P distance of the genus *Scolopsis* was recorded to be 0.75% (n=24), which was lower than the genus *Nemipterus*, with a reading of 13.19% (n=84) (Table 4). The congeneric nucleotide K2P distance was approximately 20-fold higher than the conspecific nucleotide K2P distance. By comparing both the intraspecific and congeneric nucleotide K2P distance, it can be concluded that the genetic divergence becomes greater with higher taxonomic level (Mat Jaafar *et al*., 2012). The mean intraspecific K2P distance in the present study was observed to be close to the mean K2P distance reported for marine species (0.30%) (Zhang & Hanner, 2011). The concept of barcoding gap indicates that there is a gap in the distribution of pairwise distance among conspecific individuals and among species, where the branch length among species serves to be much deeper than among conspecific individuals (Meyer & Paulay, 2005). With restricted numbers of specimens gathered, the intraspecific variation was reported to be low. Similarly, Ward *et al*., (2005) reported very low intra-specific genetic divergence for marine teleost species.

COI divergence assessment plays an important role in lending a hand to the taxonomic experts in giving an assurance for a specific species identification. Fish species specimens that show deep intra-species divergence based on the analysis of barcode *COI* gene can be rejected to be classified under the same species by the taxonomic experts as well (Bhattacharya *et al*., 2016). According to a study conducted by Mat Jaafar *et al*. (2012), the effectiveness of *COI* gene was shown in discriminating 36 commercially important marine fish from the Malaysian waters. In the present study, nine species from two genera of Family Nemipteridae in Malaysia were analysed using the *COI* gene.

Table 3: Intraspecific nucleotide K2P distances for nine species of Family Nemipteridae.

Species	No. of sequences (n)	Mean K2P distance (%)	Min (%)	Max (%)
Nemipterus balinensoides	3	0.14	0.00	0.20
Nemipterus bathybius	3	0.14	0.00	0.20
Nemipterus marginatus	29	0.48	0.00	0.80
Nemipterus nematophorus	34	1.38	0.00	2.60
Nemipterus nemurus	4	0.20	0.00	0.40
Nemipterus tambuloides	4	0.00	0.00	0.00
Nemipterus thosaporni	4	0.00	0.00	0.00
Nemipterus virgatus	3	0.00	0.00	0.00
Scolopsis taenioptera	24	0.75	0.00	1.90

Table 4: Congeneric nucleotide K2P distances for two genera in Family Nemipteridae.

Genus	No of sequences (n)	Mean K2P distance (%)
Nemipterus	84	13.19
Scolopsis	24	0.75

Phylogenetic Analysis

A total number of 108 sequences which encompassed nine species from Family Nemipteridae in Malaysia were used to construct the Neighbour-joining (NJ) tree and Maximum-likelihood (ML) tree. All of the specimens generated in this study were clustered into their own taxa, except for one of the potential misidentification species (Figure 1 and Figure 2). It was noticed that the misidentified species *N. tambuloides* grouped together with *N. thosaporni* with a high bootstrap value of 100%. This indicated the node was well-supported. On the other hand, the NJ and ML trees both exhibited two major clades dividing the *N. nematophorus*. Samples collected from Kelantan, Melaka and Sabah were grouped together in cluster 1, whereas the specimens from Terengganu and Perlis were grouped in cluster 2. Notwithstanding the smaller sample sizes, the trend was homogeneous in *S. taenioptera*.

It is known that the recognition of a specific species and phylogenetic relationship based on both the conventional methods and molecular methods are predominantly consistent (Ward *et al*., 2005). The present research shows that nine selected species from Family Nemipteridae in Malaysia were found genetically distinct from one another, where the *COI* sequences displayed simplicity and unambiguity. However, the data associated with *N. nematophorus* and *S. taenioptera* showed a reading of 2.60% and 1.90% respectively for

maximum intraspecific divergence. This revealed that the individuals formed two clusters based on the phylogenetic analysis in NJ tree. For *N. nematophorus*, the specimens collected from Kelantan, Melaka and Sabah were grouped in the same cluster, while the other cluster was made up of specimens from Terengganu and Perlis. Despite that, no clear geographical structure was perceived. According to a study conducted by Farivar *et al.*, (2017) based on the Japanese threadfin bream, *N. japonicus*, it was found that same species from the same locality may be divided into two significant clades due to the large evolutionary effective population size and high gene flow. It is noted that nucleotide distance should be lower than 3% for individuals in the same species (Hebert *et al.*, 2003b). With a maximum divergence of 2.60%, which was near to 3%, there is a possibility of cryptic species found in *N. nematophorus* within this region.

Scolopsis taenioptera, one of the species which is widely distributed in the Andaman Sea and West Pacific (Matsunuma *et al.*, 2011) has also been found to form two clusters based on the phylogenetic analysis. With a maximum divergence of 1.90%, it does not support the point for the appearance of cryptic species of *S. taenioptera* within this region. On the other hand, based on a study conducted by Kakioka *et al.*, (2017) in the region of western Pacific Ocean, there is a genetic break

in *S. taenioptera*. The study also revealed that *S. taenioptera* collected from Terengganu, Malaysia and Rayong populations invariably exhibit clear signs of demographic expansion, where both of the locations are found to be on Sunda Shelf.

Potential Misidentification of Species

Based on the phylogenetic analysis, only one sequence was remarked as potential misidentification. Based on morphology examination, the species was identified as *N.tambuloides*. However, it was genetically identified as *N. thosaporni* as it did not group with its own taxa. Both of the species *N. tambuloides* and *N. thosaporni* can be commonly found on the sand or mud bottoms (Matsunuma *et al.*, 2011; Froese & Pauly, 2017). Hence, these two species are highly associated. Besides, based on their morphological characteristics, both of the species appeared to be pink with silver in colour. The pectoral and pelvic fins appeared to be long for these two *Nemipterus* species. Generally, *N. tambuloides* can be distinguished from *N. thosaporni* by having five well-defined yellow stripes along the body. In this study, the misidentification of species may arise ascribed to the altered conditions in the specimen due to trawl pressure which resulted in faint colour and deformities (Seah *et al.*, 2017).
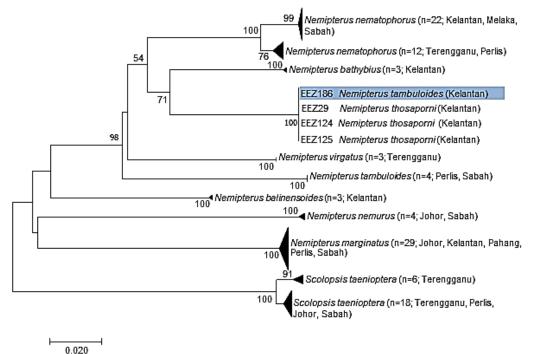

Figure 1: Neighbour-joining tree (K2P distance) of nine species from Family Nemipteridae. Only bootstrap value that are greater than 50 are shown. Blue box indicates the potential misidentification of species.
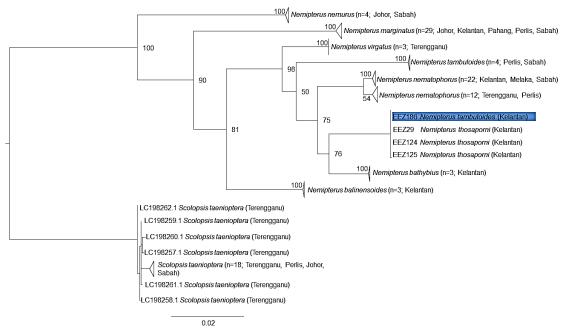
Figure 2: Maximum-likelihood tree (K2P distance) of nine species from Family Nemipteridae. Only bootstrap value that are greater than 50 are shown. Blue box indicates the potential misidentification of species.

Conclusion

Understanding the taxonomy of different aquatic organisms is much needed through the maximization of global efforts. The present study has successfully shown the effectiveness of *COI* divergences in discriminating fishes of Family Nemipteridae in Malaysia. It was proved beyond doubt that the mean intraspecific nucleotide K2P distance obtained was lower than the mean congeneric nucleotide K2P distance. Besides, phylogenetic analysis based on Neighbour-joining (NJ) tree and Maximum-likelihood (ML) tree clearly classified nine species into their respective clades, supporting the use of *COI* in classifying specimen into species level. The data procured in this study may assist in the management of the selected species from Family Nemipteridae in Malaysia waters as well as facilitate biodiversity conservations. Further research could be carried out by using higher evolutionary rate gene or increasing the sample size to strengthen the results on species identification.

Acknowledgement

References

Ambak, M. A., Mat Isa, M., Zakaria, M. Z. & Ghaffar, M. A. (2010). *Fishes of Malaysia*. Universiti Malaysia Terengganu. 334p.

Aquilino, S. V. L., Tango, J. M., Fontanilla, I. K., Pagulayan, R. C., Basiao, Z. U., Ong, P. S. & Quilang, J. P. (2011). DNA barcoding of the ichthyofauna of Taal Lake, Philippines. *Molecular Ecology Resources*, 11, 612-619.

Bhattacharya, M., Sharma, A. R., Patra, B. C., Sharma, G., Seo, E., Nam, J., Chakraborty, C. & Lee, S. (2016). DNA barcoding to fishes: current status and future directions. *Mitochondrial DNA*, 27, 2744-2752.

Blaxter, M. (2003). Molecular systematics - counting angels with DNA. *Nature*, 421, 122124. doi:10.1038/421122a

Farivar, S., Jalil-Piran, Z., Zarei, F. & Hosseinzadeh, S. H. (2017). Intraspecific phylogeography of the Japanese threadfin bream, *Nemipterus japonicas* (Perciformed: Nemipteridae), from the Persian Gulf and Indo-West Pacific: a preliminary study based on mitochondrial DNA sequence. *Iranian Journal of Fisheries Sciences*, 16, 587-604.

Froese, R. & Pauly, D. Editors. (2017) FishBase. World Wide Web electronic publication. www.fishbase.org,version (06/2017).

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003a). Biological identifications through DNA barcodes. *Proceedings of The Royal Society of London B*, 270, 313-322.

Hebert, P. D. N., Ratnasingham, S. & deWaard, R. (2003b). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *The Royal Society*, P4.

Hebert, P. D. N. & Gregory, T. R. (2005). The promise of DNA Barcoding for taxonomy. *Systematic Biology*, 54, 852-859.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. (2004). Ten species on one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 14812-14817.

Imtiaz, A., Yen, D. T., Mohd Nor, S. A. & Md. Naim, D. (2016). Molecular identification of commercially important species of *Nemipterus* (Perciformes: Nemipteridae) in surrounding seas of Malaysia. *Biodiversitas*, 17, 571-577.

Kakioka, R., Muto, N., Takeshima, H., Gaje, A. C., Cruz, R. S., Alama, U. B., Guzman, A. M. T., Traifalgar, R. F. M., Babaran, R. P., Muda, O., Arshaad, W. M., Arnupaphoon, S., Phuttharaksa, K., Nguyen, Q. V., Pham, T. T., Motomura, H., Muro, F. & Ishikawa, S. (2017). Cryptic genetic divergence in *Scolopsis taenioptera* (Perciformes: Nemipteridae) in the western Pacific Ocean. *Ichthyological Research*, 1-9.

Kumar, S., Stecher, G. & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Society for Molecular Biology and Evolution*, p11.

Mat Jaafar, T. N. A., Taylor, M. I., Mohd Nor, S. A., de Bruyn, M. & Carvalho, G. R. (2012). DNA barcoding reveals cryptic diversity within commercially exploited Indo-Malay Carangidae (Teleosteii: Perciformes). *PLoS ONE*, 7.

Matsunuma, M., Motomura, H., Matsuura, K., Shazili, N. A. M. & Ambak, M. A (eds.). (2011*). Fishes of Terengganu, East Coast of Malay Peninsula, Malaysia.* National Museum of Nature and Science, Universiti Malaysia Terengganu and Kagoshima University Museum, 251p.

Meyer, C. P. & Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology*, 3, 2229-2238.

Miller, S. A., Dykes, D. D. & Polesky. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, 16, 1215.

Russell, B. C. (1990). FAO species catalogue. Vol. 12. Nemipterid fishes of the world (threadfin breams, whiptail breams, monocle breams, dwarf monocle breams, and coral breams). Family Nemipteridae. An annotated and illustrated catalogue of nemipterid species known to date. *FAO Fish. Synop.*, (125)12. Rome, FAO, 149p.

Russell, B. C. (1993). A review of the threadfin breams of the genus *Nemipterus* (Nemipteridae) from Japan and Taiwan, with description of a new species. *Japanese Journal of Ichthyology*, 39, 295-310.

Savolainen, V., Cowan, R. S., Vogler, A. P., Roderick, G. K. & Lane, R. (2005). Towards writing the encyclopaedia of life: an introduction to DNA Barcoding. *Philosophical Transactions of the Royal Society B*, 360, 1805-1811.

Seah, Y. G., Ariffin, A. F. & Mat Jaafar, T. N. A. (2017). Levels of *COI* divergence in Family Leiognathidae using sequences available in GenBank and BOLD Systems: a review on the accuracy of public databases. *AACL Bioflux*, 10, 391-401.

Strauss, R. E. & Bond, C. E. (1990). Taxonomic methods: morphology. In: Schreck, C. B., Moyle, P. B. *Methods for fish biology.* American Fisheries Society, Maryland, pp109-140.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transaction of the Royal Society*, 360, 1847-1857.

Wang, J. F., Jiang, L. Y. & Qiao, G. X. (2011). Use of mitochondrial *COI* sequence to identify species of the subtribe Aphidina (Hemiptera, Aphididae). *ZooKeys*, 122, 1-17.

Yoshida, T., Motomura, H., Musikasinthorn, P. & Matsuura, K. (eds.). 2013. *Fishes of Nothern Gulf of Thailand*. National Museum of Nature and Science, Tsukuba, Research Institute for Humanity and Nature, Kyoto, and Kagoshima University Museum, Kagoshima. 239p.

Zhang, J & Hanner, R. (2011). DNA barcoding is a useful tool for the identification of marine fishes from Japan. *Biochemical Systematics and Ecology*, 39, 31-42.