

## DNA BARCODING OF FRESHWATER FISH SPECIES OF MANNA RIVER (BENGKULU) AND SEMANKA RIVER (LAMPUNG)

Arif Wibowo<sup>1</sup>, Achmad Farajalah<sup>2</sup> and Husnah<sup>1</sup>

<sup>1</sup>Research Institute for Inland Fisheries (RIIF), Mariana, Palembang, Indonesia.

<sup>2</sup>Bogor Agricultural University, Dramaga, Bogor, Indonesia

Received February 10-2012; Received in revised form May 07-2013; Accepted May 29-2013

Email: wibarf@yahoo.com.

### ABSTRACT

The identification of fish species is challenging. DNA barcoding provides new perspective in ecology and systematics of fishes. In this study, we assess intraspecific and interspecific genetic divergence among several freshwater fish species from Manna River of Bengkulu Province and Semangka River of Lampung Province, using data from mitochondrial cytochrome c oxidase subunit I (COI) gene. The results show that a total of six freshwater fish species identified in this study exhibited high levels of interspecific variation based on 51 analyzed sequences. Almost all species corresponded to a single species, cohesive array of barcode sequences distinct from other species. Based on Kimura Two Parameters (K2P) within species, mean pairwise sequence distances comparisons range from 0.3 to 6.3%, while interspecific comparisons were much higher range between 1.7 to 29.4%, except species *Tor macrolepis* and *Tor putitora* showed high levels of genetic similarities and no reciprocal monophyly. Using this method will obviously allow the identification eggs, larvae, fillets and fins of freshwater fish.

**KEYWORDS:** Molecular, fish, COI, Manna, Semangka

### INTRODUCTION

Concern has been voiced over the decline of taxonomy and the number of practicing taxonomists (Hopkins & Freckleton, 2002, Gotelli, 2004, Raven, 2004), accurate species identification remains an imperative condition to investigate on biodiversity and conservation. To date, traditional taxonomy relies mostly on diagnostic morphological characters, requiring expert knowledge to identify specimens.

The use of complementary techniques such as DNA sequences may enhance taxonomic and systematic studies. DNA barcoding has proved to be a useful alternative method for rapid global biodiversity assessment, providing an accurate identification system for living organisms (Hebert *et al.*, 2003, Jia-Min & Hickey, 2007, Valentini *et al.*, 2009, Tang *et al.*, 2010). This method of species identification is based on detecting sequence diversity in a single standardized DNA fragment, namely, the mitochondrial Cytochrome c Oxidase Subunit I (COI) (Hebert *et al.*, 2003). Examination of nucleotide sequence diversity of this gene allows the grouping of unknown specimens with a prior defined taxonomic species (Monaghan *et al.*, 2005, Vogler & Monaghan, 2006) based on the assumption that intraspecific genetic divergence is lower than the interspecific one (Hebert *et al.*, 2003, Meyer & Paulay, 2005, Waugh, 2007). This method has provided a high degree of taxonomic

resolution (> 94%) for most of the species examined across several animal groups (Hebert *et al.*, 2003, Clare *et al.*, 2007, Waugh, 2007).

Freshwater fish provides an excellent opportunity to test the efficacy of barcoded-based species delimitation and identification over a broad geographical range (Hubert *et al.*, 2008). Here, we assess intraspecific and interspecific genetic divergence among freshwater fish along Manna River (Bengkulu Province) and Semangka River (Lampung Province), using COI. Our specific goal was to test the degree of accuracy of DNA barcoding to discriminate morphological described species of freshwater fish at different taxonomical levels.

### MATERIALS AND METHODS

#### Sampling

Freshwater fishes were collected in 13 different regions of Manna River and one region of Semangka River. In order to estimate for intraspecific sequence variation and to avoid any misleading results due to restricted sampling in terms of specimens and geographic coverage (Table 1 and Figure 1). Some samples were deposited in Freshwater fish Museum at Research Institute for Inland Fisheries (RIIF) Palembang.

Table 1. Summary of taxonomic classification, number of individuals sequenced (N), number of haplotypes registered for each species (H), sampling localities and georeferences for freshwater fish species collected from the Manna and Semangka Rivers

Order	Family	Species	N	H	Locality	Latitude	Longitude
Cypriniformes	Cyprinidae	<i>Tor tambroides</i> (Bleeker, 1854)	2	2	Kerinjing, Bengkulu	04° 07.054'S	103° 05.529'E
			1	1	Batu rancing, Bengkulu	04° 07.226'S	103° 05.597'E
			11	2	Air tenam, Bengkulu	04° 15.856'S	103° 03.771'E
			2	1	Batu aji, Bengkulu	04° 15.188'S	103° 00.033'E
			1	1	Lubuk tapi, Bengkulu	04° 16.234'S	102° 57.306'E
			2	1	Massat, Bengkulu	04° 16.231'S	102° 57.300'E
			3	2	Merabung, Bengkulu	04° 07.226'S	103° 01.597'E
			4	2	Bandar Agung, Bengkulu	04° 20.282'S	102° 57.306'E
			2	1	Air sebilo, Bengkulu	04° 23.800'S	102° 57.936'E
			3	2	Kotabumi, Bengkulu	04° 22.541'S	102° 7.752'E
			4	1	Kutopadang, Bengkulu	04° 28.122'S	102° 55.600'E
			2	1	Ampera market, Bengkulu	04° 28.100'S	102° 55.570'E
			2	2	Melebuy, Lampung	05° 08.926'S	104° 14.870'E
			1	1	Accession HM. 536923 (Genbank)		
Cypriniformes	Cyprinidae	<i>Tor tor</i> (Hamilton, 1822)	1	1	Accession EU714115.1 (Genbank)		
Cypriniformes	Cyprinidae	<i>Tor malabaricus</i> (Jerdon, 1854)	1	1	Accession HM585024.1 (Genbank)		
Cypriniformes	Cyprinidae	<i>Tor putitora</i> (Hamilton, 1822)	1	1	Accession GQ469826.1 (Genbank)		
Cypriniformes	Cyprinidae	<i>Tor khudree</i> (Sykes, 1839)	1	1	Accession GQ469796.1 (Genbank)		
Cypriniformes	Cyprinidae	<i>Tor macrolepis</i> (Heckel, 1838)	1	1	Accession GQ469832.1 (Genbank)		
Cypriniformes	Cyprinidae	<i>Tor sinensis</i> (Wu, 1977)	1	1	Accession HM536900.1 (Genbank)		
Cypriniformes	Cyprinidae	<i>Tor duorensis</i> (Valenciennes, 1842)	2	1	Accession JN646100.1 (Genbank)		
Anguilliformes	Anguillidae	<i>Anguilla marmorata</i> (Quoy & Gaimard, 1824)	1	1	Bandar Agung, Bengkulu	04° 20.282'S	102° 57.306'E
			1	1	Air tenam, Bengkulu	04° 15.856'S	103° 03.771'E
			2	2	Accession HM345929.1 (Genbank)		
					Accession AP00724 (Genbank)		

Continue.....

Order	Family	Species	N	H	Locality	Latitude	Longitude
Cypriniformes	Cyprinidae	<i>Osteochilus hasseltii</i> (Valenciennes, 1842)	3	1	Ampera market, bengkulu	04° 28.100'S	102° 55.570'E
			1	1	Massat, Bengkulu	04° 16.231'S	102° 57.300'E
			1	1	Accession JF915633.1(Genbank)		
Cypriniformes	Cyprinidae	<i>Rasbora elegans</i> (Volz, 1903)	1	1	Air tenam, Bengkulu	04° 15.856'S	103° 03.771'E
			1	1	Bandar Agung, Bengkulu	04° 20.282'S	102° 57.306'E
			1	1	Melebuy, Lampung	05° 08.926'S	104° 14.870'E
			1	1	Accession HM224227.1(Genbank)		
Cyprinodontiformes	Poeciliidae	<i>Poecilia reticula</i> (Peters, 1859)	2	2	Batu rancing, Bengkulu	04° 07.226'S	103° 05.597'E
			1	1	Kerinjing, Bengkulu	04° 07.054'S	103° 05.529'E
			1	1	Accession GU702152.1(Genbank)		
Cypriniformes	Cyprinidae	<i>Puntius binonatus</i> (Valenciennes, 1842)	1	1	Batu rancing, Bengkulu	04° 07.226'S	103° 05.597'E
			1	1	Accession JF764644.1(Genbank)		

**DNA Extraction, Amplification and Sequencing**

A number of 51 specimens belonging to 6 freshwater fish species was preserved in 95% ethanol for subsequent analyses. All DNAs were extracted from soft tissues using the Extraction Kit procedure 'DNeasy Blood & Tissue' (Geneaid). Partial fragments of the mitochondrial gene COI were amplified using universal primers described by Ivanova *et al.* (2007):

Fish-COI-F (5'-ACT TCAAAC TTC CAY AAA GAY aty GG-3) and  
 COI-Fish-R (5'-TAG ACT TCT GGG TGG CCR AAR Aay CA-3 ').

Polymerase chain reaction (PCR) amplifications were made in a 50 µL of reaction volume consisted of 5 µL DNA samples, 16 µL double distilled water, 2 µL

of each primer and 25 µL of PCR ready mixture solution (KAPPA).

PCR cycling parameters consist of an initial denaturation phase at 95°C for 10 min, followed by 35 cycles at 94°C for 1 min, 48°C for 1 min and 72°C for 1.5 min and ended with a final extension at 72°C for 7 min. Finally, all amplicons were automatically sequenced in both directions at MacroGen S.A. Korea ([www.macrogen.com](http://www.macrogen.com)) and these sequences were subject to register in Genbank (NCBI), the accession numbers are JQ665787 – JQ665837.

GenBank is part of the International Nucleotide Sequence Database Collaboration, which comprises the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI.



Figure 1a. Sampling station in Manna River (Husnah *et al.*, 2012)



Figure 1b. Sampling station in Semangka River (Husnah *et al.*, 2012)

**Editing, Alignment and Statistical Analyses**

COI sequences were edited with BioEdit (version 7.0.4.1) (Hall, 1999). Multiple alignments were done using Clustal W (Thompson *et al.*, 1994). Following alignments, COI sequences were translated to amino acids to check for the presence of premature stop codons that indicate the presence of nuclear pseudo genes or sequencing errors. Sequence divergence was estimated using the Kimura two parameters (K2P) model of base substitution (Kimura, 1980). Phylogenetic reconstruction was done using a distance based method, Neighbor-Joining (NJ), carried out in MEGA4 software (Tamura *et al.*, 2007) with the K2P model of substitution. Support for nodes in NJ analyses was assessed using non-parametric bootstrapping with 20,000 pseudo-replicates. We only considered clusters that were supported by at least 95% (Felsenstein, 1985). Several sequences of the freshwater fish species used in Genbank (accession number in figure) to root the tree for comparative purposes.

**RESULTS**

**Sequence Variation**

A total of 51 specimens, of six freshwater fish species were successfully sequenced and analyzed using a 602 bp fragment of the COI gene (Table 1). No insertions/deletions or stop codon were found, all the amplified sequences are part of functional mitochondrial COI sequences. Furthermore, all of the amplified sequence were longer than 600-bp the limit is usually observed in the cell nucleus DNA sequences derived from mtDNA (NUMTs).

The sequences were adenine and thymine (A-T) rich (55%) compared to guanine and cytosine (C-G) content (45%). Freshwater species exhibited high levels of variation among their sequences, 240 sites were variable (39.87%) and 219 of them (36.37%) were parsimonious informative.

Table 2. Mean pairwise sequence distances comparisons based on K2P (in percentage) between the analyzed taxa

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 <i>Chitala lopis</i>	***													
2 <i>Tor tor</i>	24.10	***												
3 <i>Tor malabaricus</i>	25.00	3.10	***											
4 <i>Tor putitora</i>	24.10	1.70	2.40	***										
5 <i>Tor macrolepis</i>	24.10	1.70	2.40	0.00	***									
6 <i>Tor khudree</i>	25.70	3.40	2.40	2.40	2.40	***								
7 <i>Tor sinensis</i>	25.00	3.40	2.70	3.10	3.10	3.10	***							
8 <i>Tor duorenenensis</i>	24.60	4.40	3.00	3.90	3.90	4.10	3.70	***						
9 <i>Tor tambroides</i>	24.40	3.20	2.80	2.90	2.90	3.20	3.20	3.80	***					
0 <i>Anguilla marmorata</i>	23.70	21.80	22.70	22.20	22.20	23.80	23.60	22.60	22.20	***				
1 <i>Osteochilus hasseltii</i>	26.30	17.70	16.90	17.60	17.60	16.80	16.90	18.20	16.10	25.80	***			
2 <i>Poecilia reticulata</i>	29.40	24.60	25.30	25.60	25.60	25.40	26.80	26.00	25.90	22.70	28.40	***		
3 <i>Puntius binonatus</i>	21.90	16.40	17.30	16.00	16.00	17.00	16.10	16.70	15.90	21.80	18.90	27.60	***	
4 <i>Rasbora elegans</i>	27.60	20.60	19.40	19.90	19.90	18.90	19.30	19.70	18.80	23.00	20.80	26.70	20.30	***

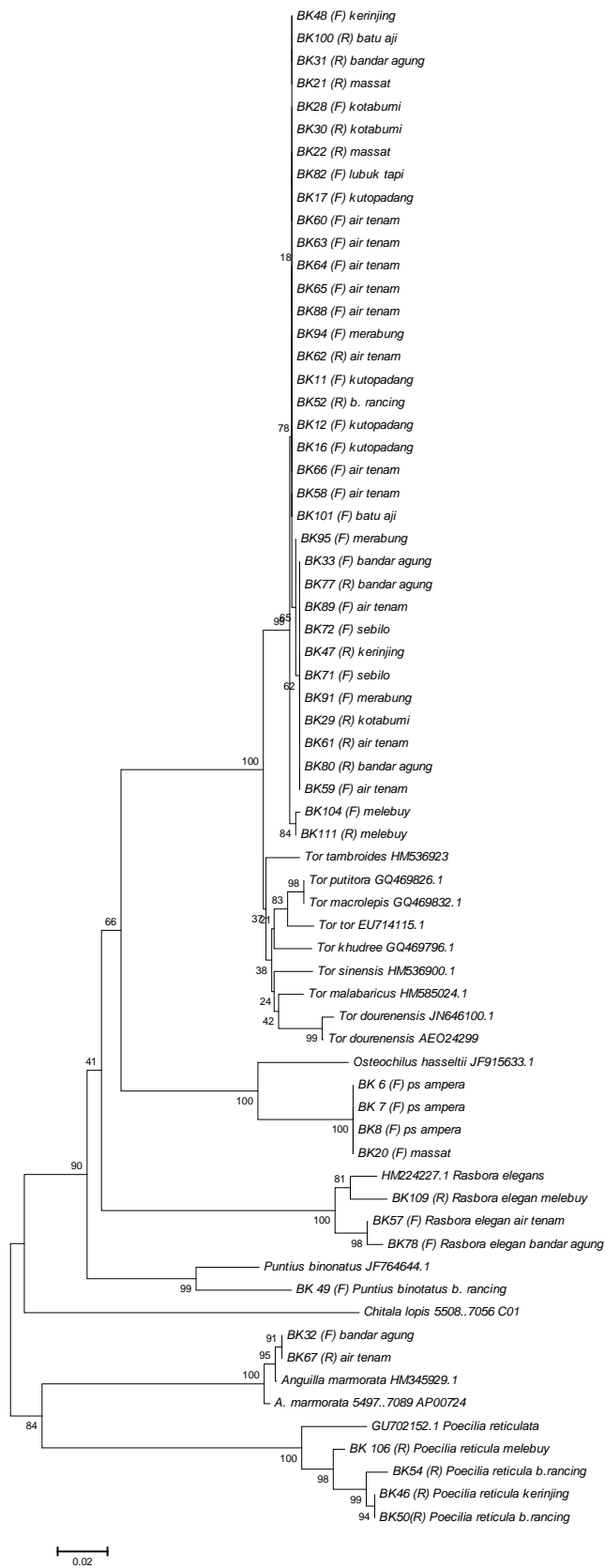


Figure 2. Neighbor-joining tree analysis of the COI sequences (K2P) for 14 freshwater fish species from the Manna and Semanka River from 20,000 iterations are shown.

## Phylogeny Reconstruction

Based on Kimura Two Parameters (K2P) within species, mean pairwise sequence distances comparisons range from 0.3 to 6.3%, while interspecific comparisons were much higher ranged between 1.7 to 29.4% (Table 2), except species *Tor macrolepis* and *Tor putitora* showed high levels of genetic similarities and no reciprocal monophyly (Fig. 2).

## DISCUSSION

In the seminar paper describing DNA barcoding, Hebert *et al.* (2003) suggested that DNA-based identification founded on the mitochondrial cytochrome c oxidase Subunit I would serve as the core of a global bioidentification system for animal life. This line of reasoning holds that when fully developed, COI identification system will provide a reliable, cost-effective and accessible solution to the current problem of species identification, generating also new insights into the diversification of life and the rules of molecular evolution. Hebert *et al.* (2003) proposed two principal elements in the DNA barcoding initiative. First, the ability to assign the unknown sample to a known species, and second, the ability to detect previously unsampled species as distinct. Assigning an unknown to a known element is promising especially for well-known, comprehensively sampled groups studied by genetic and morphological taxonomy, thus, DNA barcoding holds promise for identification in taxonomically well-understood groups (Meyer & Paulay, 2005). Species assignment is very important in the context of fisheries research, fisheries control, and consumer protection.

The results obtained in this study support the effectiveness of DNA barcoding for identification of intraspecific and interspecific genetic divergence among freshwater fish species, where almost all species analyzed corresponded to a single species, cohesive array of barcode sequences that differ from other species. According to Hebert *et al.* (2003), the general success of COI in recognizing relationships among taxa in these cases is important because it signals that character convergence and horizontal gene transfer (i.e. via retroviruses) have not disrupted the recovery of expected taxon affinities. Moreover, it establishes that the information content of COI is sufficient to enable the placement of organisms in the deepest taxonomic ranks.

The mean genetic distance between conspecifics are generally much smaller than the average genetic distance among individuals of different species, even if only considering sister species. This is due to the

fact that diversification in the species is driven by mutations at a higher rate of speciation within lineages. Therefore, the branch length between species tends to be deeper than similar among individuals that lead to disparities in the distribution of distances between pairs of individuals among species of similar and better known as the barcoding gap (Meyer & Paulay, 2005). Thus, the DNA barcoding method effectively discriminates the two species within the genus *Tor* with 1.7 – 4.4% of mean pairwise sequence distances (see Table 2 and Fig. 2).

Confusion in taxonomic assignments as a result of inter-specific hybridization (Verspoor & Hammar, 1991) does not seem to be a major issue since only two of 67 sequences (*Tor macrolepis* and *Tor putitora*,) appeared in the 'wrong' congeneric species. This may represent either introgressive hybridization, or incorrect identification of the original specimen. In this respect, it is interesting to note that the two species studied do not present strong morphological differences. Thus, a deeper understanding of the morphological taxonomy with the development of the molecular taxonomy in this genus is required for future application of barcoding.

This study has shown the first data on DNA barcoding in freshwater fish from Manna and Semangka River. We have established the effectiveness of DNA barcode for freshwater fish species identification in west Sumatra River, making it available to a much broader range of scientists, with possibilities to extend to freshwater fish in the Indonesian rivers. The degree of taxonomic resolution in the studied species is comparable to the results in other freshwater fishes (Ward *et al.*, 2005, Ardura *et al.*, 2010). Using this method will obviously allow the identification eggs, larvae, fillets and fins of freshwater fish.

Despite its methodological shortcomings and limitations, DNA barcoding studies have reinvigorated the development of systematic studies and taxonomic inventories around the world. The increased development of molecular systematic and taxonomy provide exciting opportunities to enrich our understanding of widespread taxa and many new tools are very useful for the practice of forensic and genetic conservation in freshwater fish. From a systematic perspective, COI barcodes provide a new and fast approach for screening the actual number of species that are characterized by a set of specific diagnostic character. Identification of some cases of polyphyletic or paraphyletic lineage COI species further supports the view that the iterative process of DNA barcoding followed by taxonomic analysis using other character will be a productive way to catalog biodiversity (Barber & Boyle, 2006).

## CONCLUSION

Almost all species of freshwater fish from the river Manna and Semanka showed a similar pattern of genetic diversity at the COI, each into a single cluster of closely related mtDNA sequences differing from other species. This survey supports the view that the COI barcoding is a useful tool for species identification.

## ACKNOWLEDGMENT

We thank to Apriadi, Sigit Bergii and Marson for their technical assistance during samples collection. This study was partially funded by Research Institute for Inland Fisheries, Indonesia through Government Income and Expenditure grant 2011.

## REFERENCES

- Ardura A, A.R. Linde, J.C. Moreira & E. Garcia-Vazquez. 2010. DNA barcoding for conservation and management of Amazonian commercial fish. *Biological Conservation* 143: 1438-1443.
- Barber P & S.L Boyce. 2006. Estimating diversity of Indo-Pacific coral reef stomatopods through DNA barcoding of stomatopod larvae. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, 273:p. 2053–2061.
- Clare E.L., B.K. Lim, M.D. Engstrom, J.L. Eger & P.D.N. Hebert. 2007. DNA barcoding of neotropical bats: species identification and discovery within Guyana. *Molecular Ecology Notes* 7: 184-190.
- Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.
- Gotelli N.J. 2004. A taxonomic wish-list for community ecology. *Philosophical Transaction of the Royal Society B*. 359: 585-597.
- Hall T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41: 95–98.
- Hebert P.D.N., A. Cywinska., S.L. Ball & J.R. de Waard. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B* 270: 313-321.
- Hopkins G.W & R.P. Freckleton. 2002. Declines in the numbers of amateur and professional taxonomists: implications for conservation. *Animal Conservation* 5: 245-249.
- Hubert N., R. Hanner., E. Holm., N.E. Mandrak & E. Taylor. 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One* 3: e2490.
- Ivanova N.V., T.S. Zemlak., R.H. Hanner & P.D.N. Hebert. 2007. Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes* 7: 544–548.
- Jia-Min X & D.A Hickey. 2007. Assessing the effect of varying sequence length on DNA barcoding of fungi. *Molecular Ecology Notes* 7: 365-373.
- Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16:111-120.
- Lakra, W.S & M.S. Verma. 2008. *DNA Barcoding of Indian Fishes*. Unpublished. Submitted (11-NOV-2008) International Centre for DNA Barcoding, National Bureau of Fish Genetic Resources, Canal Ring Road, P. O. Dilkusha, Lucknow, Uttar Pradesh 226002, India.
- Meyer C.P & G. Paulay. 2005. DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology*, 3: p.2229–2238.
- Monaghan M.T., M. Balke., T. Ryan-Gregory & A.P. Vogler. 2005. DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. *Philosophical Transactions of the Royal Society B* 360: 1925-1933.
- Raven P.H. 2004. Taxonomy: where are we now? *Philosophical Transactions of the Royal Society B* 359: 729-730.
- Tamura K., J. Dudley, M. Nei & S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 10.1093/molbev/msm092.
- Tang R.W.K., C. Yau & W.C. Ng. 2010. Identification of stomatopod larvae (Crustacea: Stomatopoda) from HongKong waters using DNA barcodes. *Molecular Ecology Resources* 10: 439-448.
- Thompson J.D., D.G. Higgins & T.J. Gibson. 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673-4680.



- Valentini A., F. Pompanon & P. Taberlet. 2009. DNA barcoding for ecologist. *Trends in Ecology and Evolution* 24: 110-117.
- Verspoor, E. & J. Hammar. 1991 Introgressive hybridization in fishes: the biochemical evidence. *J. Fish Biol.* 39 (Suppl. A), 309–334.
- Vogler A.P & M.T. Monaghan. 2006. Recent advances in DNA taxonomy. *Journal of Zoological Systematics Evolutionary Research* 45: 1-10.
- Ward R.D., T.S Zemplak, B.H Innes, P.R. Last & P.D.N. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society, Series B*, 360: p.1847–1857.
- Waugh J. 2007. DNA barcoding in animal species: progress, potential and pitfalls. *BioEssays* 29: 188-197.