THE GENETIC STRUCTURE OF THE WHITE CYPRINID (*TOR TAMBROIDES)* POPULATIONS BASED ON COI GENE SEQUENCE ANALYZES

Arif Wibowo and Husnah
Research Institute for Inland Fisheries, Mariana, Palembang
Received January 30-2012; Received in revised form May 4-2012; Accepted June 6-2012
email : wibarf@yahoo.com

ABSTRACT

The study analyzed nucleotide sequences from the mitochondrial cytochrome oxidase subunit (COI) gene region (654) to investigate the genetic structure of the white cyprinid (Tor tambroides) among nine populations from the Manna and Semanka Rivers. A total of 36 individuals were collected for this work. Five nucleotides were found to be variable, resulting in 4 haplotypes. Among the nine populations of Melebuy in Semanka River represents the highest level of variability (h = 1.000, π = 0.0015) whereas Kerinjing, Merabung and Kotabumi populations represent the highest level of variability in Manna River (h = 0.667, π = 0.0020). The Batu Aji, Air Sebilo and Kutopadang populations exhibit the lowest level of variability (h = 0.000, π = 0.000). There is an integrated population throughout all sample sites in Manna River. However, the AMOVA analysis provided corroborating evidence for genetic structure obtained from Nei's genetic diversity statistic and the F_{ST} value (0.310), suggested there is genetic divergence among populations of those populations. Of the total genetic diversity, 35% was attributable to inter-population diversity and the remainder (68.92%) to differences within populations. These two approaches produced a picture of genetic structure in Manna and Semanka River. A molecular phylogenetic tree constructed using The Neighbor-joining (NJ) method showed the 4 haplotypes were assigned to two clades associated geographic regions. These results provide basic information for the conservation and sustainable exploitation of this species.

KEYWORDS: Tor tambroides, COI genes, genetic structure, genetic variation, haplotype

INTRODUCTION

The white cyprinid (*Tor tambroides)* is one of the most popular Indonesian freshwater fish food both for human consumption and traditional culture. However, the wild population has undergone a steady decline at the same time, mainly because of overexploitation, pollution, unfriendly fishing practices and construction of hydropower projects. The genetic variation among *T. tambroides* population is largely unknown and there is not information on population differentiation. Information on genetic variation is vital in designing and implementation of adequate management strategies for the species. Furthermore, in order to relieve pressure on wild stock and to implement and to improve industry products through selective breeding, development of domestication program for *T. tambroides* is important, requiring baseline information concerning genetic background information and genetic variations.

Until recently, all major genetic studies on this fish species only based on morphological characteristics (Wibowo *et al.,* 2008), which was simply studied on its genetic variation and there has been no research on population structure before. Therefore, it is necessary to study the genetic diversity of this species, especially using DNA analyzes tools.

As a genetic marker, mitochondrial DNA (mtDNA) has some advantages over the others such as the virtue of its maternal, non recombining mode of inheritance, its rapid pace of evolution and extensive intraspecific polymorphism (Avise *et al*., 1987). MtDNA is widely used in the study of genetic variation in organisms, including some crustaceans. The object of this study is to investigate the genetic structure of the white cyprinids (*T. tambroides)* within six populations from western part of Indonesia, using nucleotide oxidase subunit I (COI) gene region and to provide baseline information for the conservation and sustainable utilization of the wild genetic resources of the white cyprinids.

MATERIAL AND METHODS

Sample and DNA extraction

A total of 38 individuals of *Tor tambroides* were collected from six localities. Details of each locality are shown in Table 1. Samples of muscle tissue and blood were preserved in 96% ethanol until DNA was extracted for genetic analyzes.

A piece of muscle tissue sample was homogenized and incubated in a standard buffer [0.06 mmol/L ethylenediaminetetraacetic acid (EDTA), 0.1

Corresponding author:
Jl.Beringin 308 Mariana-Palembang

mmol/L Tris (pH 8.6) 0.5% sodium dodecyl sulphate (SDS)] overnight at 37^{0}C in the presence of Proteinase-K in tilting apparatus. After buffer incubation, DNA was extracted using genomic DNA extraction kit (GENEAID), concentrated and purified using column purification. The purified DNA was dried and dissolved in TE buffer then stored at -20^{0}C.

DNA Amplification

Universal primer referring Ivanova *et al.,* (2009) was employed. Segment of the COI gene for each individual were amplified using PCR of COI F (5' – TCT ACC AAC CAC AAA GAC ATC GG 3') and COI R (5' – TAC TTC TGG GTG TCC RAA GAA TCA 3').

The PCR reaction mixture of 50 µl containing 2 µl primer COI F, 2 µl primer COI, 25 µl cocktail PCR ready mix and up to 50 ng of genomic DNA 2 µl. The PCR cycling included pre-denaturing for 4 min at 94^{0}C and 30 cycles of 30 sec at 94^{0}C, 30 sec 55^{0}C and 2 min at 72^{0}C, followed by final extension for 5 min at 72^{0}C. Samples were sent to Macrogen Biotechnologies Co, Ltd, Seoul, South Korea for sequencing. For each sample, sequencing was performed in both directions.

Data Analysis

Sequence chromatograms were viewed and edited manually using BIOEDIT Applied Biosystem. Once edited, multiple alignments were performed using Clustal W (Thompson *et al.,* 1997). Haplotype diversity (*%)* (Nei & Tajima, 1981) and nucleotide diversity (*"*) (Nei, 1987) were calculated using MEGA (Version 4.0, Tamura *at al.,* 2007). An analysis of molecular variance (AMOVA; Excoffier *et al.,* 1992) was performed the partition the total phenotypic variance into intra- and inter-population variance using ARLEQUIN 3.01 (Excoffier *et al.,* 2005). The neighbor-joining (NJ) method was used to reconstruct the phylogenetic relationships among haplotypes with MEGA (Version 4.0, Tamura *et al.,* 2007).

Table 1. Population code, locations, sample size and genetic diversity of sample populations

Pop	locality	Sample size	Latitude	Longitude	Haplotype diversity *(h)*	Nucleotide diversity (π)
1	Kerinjing, Bengkulu	3	04^0 07.054'S	103^0 05.529'E	0.6667 ± 0.3143	0.002039 ± 0.002095
2	Air tenam, Bengkulu	11	04^0 15.856'S	103^0 03.771'E	0.4364 ± 0.1333	0.001334 ± 0.001141
3	Batu aji, Bengkulu	4	04^0 15.188'S	103^0 00.033'E	0.000 ± 0.000	0.000 ± 0.000
4	Merabung, Bengkulu	3	04^0 07.226'S	103^0 01.597'E	0.6667 ± 0.3143	0.002039 ± 0.002095
5	Bandar Agung, Bengkulu	4	04^0 20.282'S	102^0 57.306'E	0.5000 ± 0.2652	0.001529 ± 0.001515
6	Air sebilo, Bengkulu	2	04^0 23.800'S	102^0 57.936'E	0.000 ± 0.000	0.000 ± 0.000
7	Kotabumi, Bengkulu	3	04^0 22.541'S	102^0 7.752'E	0.6667 ± 0.3143	0.002039 ± 0.002095
8	Kutopadang, Bengkulu	4	04^0 28.122'S	102^0 55.600'E	0.000 ± 0.000	0.000 ± 0.000
9	Melebuy, Lampung	2	05^0 08.926'S	104^0 14.870'E	1.0000 ± 0.5000	0.001529 ± 0.002162

Results

Sequence variation and haplotype distribution

A total of 654 bp mtDNA COI sequences were amplified successfully from 36 individuals of nine populations resulting in identification of 4 unique haplotypes defined by 5 variable sites, whereas 4 variable sites were parsimony informative and 1 variable site was singleton. The mean total nucleotide composition was T = 28.8%, C = 27.3%, A = 26.5% and G = 17.3%. Among the 4 haplotypes, two (50%) were shared by different population: other haplotypes were unique to the corresponding population. One haplotype (hap1) was shared between seven of the populations and other haplotype (hap2) was shared by another six of the populations (Table. 2).

Table. 2. Distribution of the 4 haplotypes in *T. tambroides* populations

Population	Haplotype			
	1	**2**	**3**	**4**
Kerinjing	1	2		
Air tenam	3	8		
Batu aji	4			
Merabung	2	1		
Bandar Agung	3	1		
Air sebilo	2			
Kotabumi	1	2		
Kutopandang		4		
Melebuy			1	1

Population genetic diversity and genetic divergence

The amount of COI gene sequence variation among nine populations is summarized in Table. 1. The Melebuy (Lampung) population exhibits the greatest level of variability (h = 1.0000, π = 0.001529), whereas Batu aji, Air sebilo and Kutopadang populations exhibits the lowest level of variability (h = 0, π = 0 / no variation). AMOVA identified that a high proportion of the total genetic variance was attributable to variations within populations (68.92%), whereas 35.08% and only -4.0% occurred among populations and within populations respectively (Table. 3).

The fixation index (F_{ST}) and genetic distance (Kimura-2, parameter's measure) between pairs of populations are shown in Table. 4. These data suggested that the Kerinjing and Kutabumi population were genetically most similar (F_{ST} = - *0.5),* while the Melebuy and Kutopadang populations were most dissimilar (F_{ST} = *0.724).* However we can conclude that there is a gene flow occurred within Manna River populations.

Table 3. Analysis of molecular variance (AMOVA) for the *T. tambroides* populations

Sources of variation	Degree of freedom	Variance components	Percentage of variation
Among groups (upper reaches, lower reaches of Manna River and Semanka River)	2	-0.01167	-4.00
Among populations within groups	6	0.10241	35.08
Within populations	27	0.20118	68.92
Total	35	0.29191	

F_{ST} = *0.31082*

Table. 4 The value of fixation index (F_{ST}) (below) and genetic distance (above) on sample populations of *T. tambroides*

	Kutopadang	Kutabumi	Bandar agung	Kerinjing	Air Tenam	Air sebilo	Batu aji	Merabung	Melebuy
Kutopadang	-	0.001	0.002	0.001	0.001	0.003	0.000	0.002	0.004
Kutabumi	0.111	-	0.002	0.001	0.001	0.002	0.001	0.002	0.005
Bandar agung	0.667	0.014	-	0.002	0.002	0.001	0.002	0.001	0.006
Kerinjing	0.111	-0.500	0.014	-	0.001	0.002	0.001	0.002	0.005
Air Tenam	0.049	-0.257	0.251	-0.257	-	0.002	0.001	0.002	0.005
Air sebilo	1.000	0.368	-0.263	0.368	0.509	-	0.003	0.001	0.007
Batu aji	1.000	0.579	0.000	0.579	0.584	0.000	-	0.002	0.004
Merabung	0.579	-0.200	-0.388	-0.200	0.103	-0.200	0.111	-	0.006
Melebuy	0.724	0.208	0.344	0.208	0.455	0.500	0.724	0.208	-

Phylogeny Reconstruction

The neighbour-joining tree (Fig. 1) was constructed with the complete data set of 4 haplotypes and nine relative haplotypes from genus Tor around the globe. Two main clades were identified for Indonesian *T. tambroides.* Samples from the Kerinjing, Air tenam, Batu aji, Merabung, Bandar Agung, Air sebilo, Kotabumi and Kutopandang populations formed one cluster (common haplotype of Manna River and a unique haplotype found in middle and upper reaches of Manna River). The other clade was composed of the Melebuy population, which was collected from middle reaches of Semanka River.

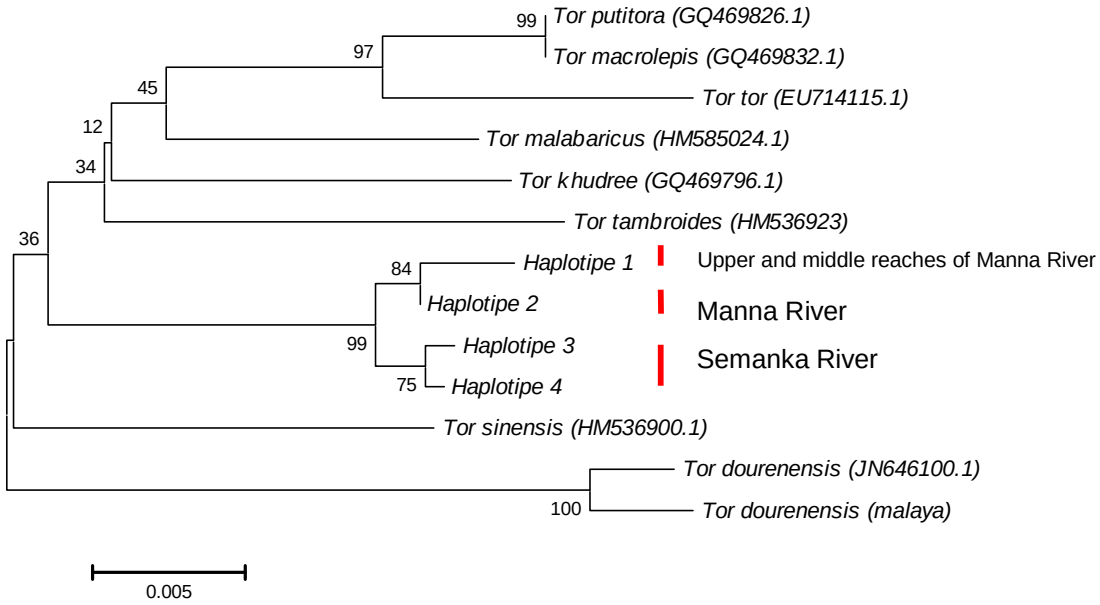

Figure 1. Neighbor-joining tree of COI haplotypes of *T. tambroides* and its relatives

DISCUSSION

This study used nucleotide sequences from the COI gene region to investigate the genetic structure of *T. tambroides* among nine populations with a total of 36 individuals. The percentage of A + T base composition (55.3%) was much higher than C + G, which coincides with vertebrate protein-coding genes (Miller *et al*., 2005; Ping *et al*., 2007).

The present study sequencing an approximately 654 bp fragment of this gene region revealed 4 haplotypes based on polymorphisms at 5 nucleotide sites. Compared to others vertebrate study this present research resulted in less variable sites and haplotypes due to the typically short distance river in eastern Sumatra in Indonesia and less informative character of the COI gene. Cook *et al.,* (2002) investigated the genetic structure of *Macrobrachium australiense* throughout the rivers of western Queensland in Australia. Their study revealed that sequences of a 607 bp fragment of COI gene revealed 17 haplotypes in 28 individuals from six localities, based on polymorphisms at 36 nucleotide sites. Bruyn *et al*., (2004) studied the phylogeography of *M. rosenbergii* from the Lake Carpentaris region, Australia. They amplified the 602 bp COI gene for all samples taken 378 individuals.

Pairwise F_{ST} analysis and Nei's genetic distance analysis indicated that the genetic different of *T. tambroides* between the Kerinjing and Kutabumi population is the smallest which is not corresponds to the geographic distance. The smaller in Kerinjing and Kutabumi population (3 individuals respectively) sample size might be part of the reason for the unrelated between genetic different and geographic distance. In this study, Melebuy in Semanka River presents the highest level of variability (h = 1.000, π = 0.0015) whereas Kerinjing, Merabung and Kotabumi represent as the highest level of variability in Manna River (h = 0.667, π = 0.0020). The Batu aji, Air sebilo and Kutopadang populations exhibit the lowest level of variability (h = 0.000, π = 0.000).

Although unevenly sample size among sampling sites, the existence of shared haplotypes and very small genetic difference in grouping analysis of upper reaches, lower reaches of Manna River and Semanka River) (-4%), its contributed substantial evidence that there is an integrated population throughout all sample sites in Manna River. However, the AMOVA analysis provided corroborating evidence for genetic structure obtained from Nei's genetic diversity statistic. Of the total genetic diversity, 35% was attributable to inter-population diversity and the remainder (68.92%) to differences within populations. The F_{ST} value (0.310) suggested there is genetic divergence among populations. These two approaches analysis provided a picture of genetic structure in Manna and Semanka River.

The mtDNA data resolved two divergent monophyletic clades, with the first representing *T. tambroides* in Manna River and the second corresponding to the Semanka River. The present

study showed a significant correlation between genetic differentiation and geographical distance. The passive dispersal of fry would be contributed to the close relationship of the middle and upper reaches of Manna River. Although the results of this study may be affected by relatively small sample size and collection areas, they can be used as foundation for further nuclear DNA-based studies and baseline information for protection and management of this species.

The maintenance of genetic polymorphism in natural populations can reflect the process of adaptation to environmental heterogeneity. The reduction of genetic variation of a species could have serious consequences, such as reducing survival, growth and reproduction and suppressing the ability of individuals in the population to adapt to a changing environment. In breeding hatchery-reared individuals may escape or be released into the wild, which would reduce the genetic variability of wild stock. The genetic resources of *T. tambroides* are likely to decline with ongoing human activity. Therefore, it is necessary to design and implement adequate management strategies for the species to continue to exploit natural sources of *T. tambroides.*

REFERENCES

Avise J.C., J. Arnold, R.M. Ball, E. Bermingham, T. Lamb, J.E Neigel, C.A Reeb & N.C Saunders. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *J. Annu Rev Ecol Res.* 18: 489-522.

Bruyn M.D., J.C Wilson & P.B Mather. 2004. Reconciling geography and genealogy: Phylogeography of giant freshwater prawns from the Lake Carpentaria region. *J. Mol. Ecol,* 13 (11): 3515-3526.

Cook B.D., S.E Buan & J.M Hughes. 2002. Genetic structure and dispersal of *Macrobrachtium austrsaliense* (Decapoda: Palaemonidae) in West Queensland, Australia. *J. Freshwater Biol.* 47: 2098-2112.

Excoffier L., G. Laval & S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population data analysis. *J. Evolutionary Bioinformatics Online.* 1: 47-50.

Excoffier L., P.E Smouse & Quattro. 1992. Analysis of molecular variance inferred from metric distance among haplotypes: Application to human mitochondrial DNA restriction data. *Genetics.* 131 (2): 479-491.

Ivanova N.V., T.S. Zemlak., R.H. Hanner & P.D.N. Hebert. 2009. Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes.* 7: 544–548.

Miller A.D., N.P Murphy, C.P Burridge & C.M Austin. 2005. Complete mitochondrial DNA sequences of the decapod crustaceans *Psedocarcinus gigas (*Menippedae) and *Macrobrachium rosenbergii* (Palaemonidae). *J. Mar Biotechnol.* 7(4): 339-349.

Nei M & F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases (J). *Genetics*. 97 (1): 145-163.

Nei M. 1987. *Molecular Evolutionary Genetics*. New York. Colombia University Press. 457 pp.

Ping, Y., Z. Hao., C. Li-qiao., Y. Jin-yun., Y. Na., G. Zhi-min & S. Da-xiang. 2007. Genetic structure of the oriental river prawn (*Macrobrachium nipponense)* from Yangtze and Lancang River, inferred from COI gene sequence. *Zoological Research*. 28 (2): 113-118.

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.

Thompson, J.D., T.J. Gibson., F. Plewniak., F. Jeanmougin & D.J Higgins. 1997. The clustal X windows interface: Flexible strategies for multiple sequences alignment aided by quality analysis tool. *Nucleic Acid Res.* 25 (24): 4876-4882.

Wibowo A., Subgja & S. Makmur. 2008. Morphology analysis of *Tor tambra* on three different environments (in Bahasa Indonesia). *Proceeding of national seminar on fisheries*. Fisheries School. Jakarta. P 1-5.