Borneo Journal of Resource Science and Technology (2015) 5(2): 1-15

Population Genetics of *Tor douronensis* in Sarawak – A Revisit

MUHAMMAD AMIRUL ARIB MD ADZHAR* & RUHANA HASSAN

Department of Aquatic Science, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Malaysia

ABSTRACT

Tor douronensis, known locally as Semah, is one of the valuable resources of Sarawak due to its high nutritional value and demand as game fish. Several molecular studies related to *T. douronensis* have been carried out and deposited in GenBank including the data collected from Ba Kelalan, Layar and Limbang. Although there are other studies on *Tor* spp., the data was not available in GenBank, thus there are not included in the analysis. One fieldtrip to Bakun Dam was carried out on June 2014 with initial aim to assess overall fish diversity. During this sampling, 11 individuals of *T. douronensis* were caught, which later subjected to molecular work to investigate the genetic structure and evolutionary relationship among four of *T. douronensis* in Sarawak using partial Cytochrome c oxidase I (COI) mtDNA gene. A fragment of 465 bp of CO1 gene of *T. douronensis* was successfully amplified. Based on the phylogenetic trees generated, three clades could be observed namely Central, Southern, and Northern populations; 1st clade (haplogroup I) from Bakun, 2nd clade (haplogroup II) from Layar and 3rd clade (haplogroup III) from Ba Kelalan and Ulu Limbang. Overall, there were 13 haplotypes and none was shared among populations, suggesting low level of inter-population gene flow has been observed. The small number of migrants per generation ($Nm<1.0$) among the population indicated that the small populations were isolated possibly due to large geographical areas. All population had undergone expansion with a large negative value and significant test of Fu's F in Bakun population suggested recent expansion. In addition, result also suggested that all populations did not deviate from evolutionary neutrality.

Keywords: COI gene, evolutionary neutrality, population expansion, population subdivision, *Tor douronensis*

INTRODUCTION

Tor douronensis Valenciennes (1842) is a member of the mahseer group from the genus *Tor* Gray in the family Cyprinidae. It is one of the most important freshwater fishes in Malaysia (Mohsin & Ambak 1983; Roberts 1989; Litis *et al.,* 1997; Ng, 2004) inhabiting the upper streams and headwaters of most major river systems (Kottelat *et al.*, 1993; Rainboth, 1996). In Sarawak, *T. douronensis* is locally known as Semah. This species has high economic value to the local people.

About a decade ago, the price of *T. douronensis* was valued RM45/kg in the open market in Kapit, Sarawak (Ingram *et al.*, 2005). During field visit to Bakun Dam in April 2013, local people claimed that 'Semah' could fetch a price of between RM60/kg to RM100/kg in Bakun area. Thus in agreement with Ingram *et al.* (2005), this species has great potential for freshwater aquaculture industry.

Populations of *Tor* spp. are declining due to degrading environmental conditions by deforestation, logging and development of hydropower dam that may have disturbed their natural habitat (Ng, 2004). Uncontrolled fish harvest (overfishing) due to its high price has also contributed to the reduction of their population size (Ng, 2004). Their distributions in Malaysian Borneo are now limited to the upper streams and protected areas of Sarawak and Sabah (Litis *et al.*, 1997; Nyanti *et al.*, 1999; Ng, 2004).

Although currently not listed by the IUCN as a protected or endangered species, the drastic decline in natural populations of *T. douronensis* has increased awareness among relevant authorities (*e.g.*, Fisheries Department,

*Corresponding author: *citet11@gmail.com*

Malaysia and policy makers) of the importance of the conservation and proper management of this species (Esa *et al.*, 2008). Due to its economic importance, high commercial, recreational and conservational value, *T. douronensis* needs a proper management plan to ensure its sustainability. It is crucial to understand their taxanomic status, population distribution, genetic variability, levels of gene flow and population subdivisions and for understanding factors contributing to fitness of *T. douronensis.* This would ensure a more effective management plan could be developed with the aim for long-term maintenance of genetic diversity of cultured stocks, as well as to minimize potential adverse effects on the genetic integrity of the wild populations through proper stock enhancement practices.

Esa *et al.* (2008) had carried out genetic studies on *T. douronensis* involving Layar, Ba Kelalan and Ulu Limbang populations. However, he suggested that larger sample sizes per population and samples from other areas are required to reveal the actual genetic variation at the inter-population and intra-populations levels. In this current study, a re-analysis of the previous study was conducted with the addition of the Bakun Dam population.

Thus, the aim of this study is to conduct a population genetic study among four populations from Sarawak using partial DNA sequence of the Cytochrome c Oxidase I (COI) mtDNA gene analysis.

MATERIALS AND METHODS

Collection of Samples

Eleven *T. douronenis* samples were obtained from Bakun Dam (N 02° 45′ 23″, E 114° 03′ 47″) (Figure 1). Other data were obtained from Genbank with accession number, EF192444, EF192445 (Ulu Limbang samples), EF192453, EF192454, EF192455, EF192456 (Layar samples) and EF192446, EF192447 (Ba Kelalan samples.

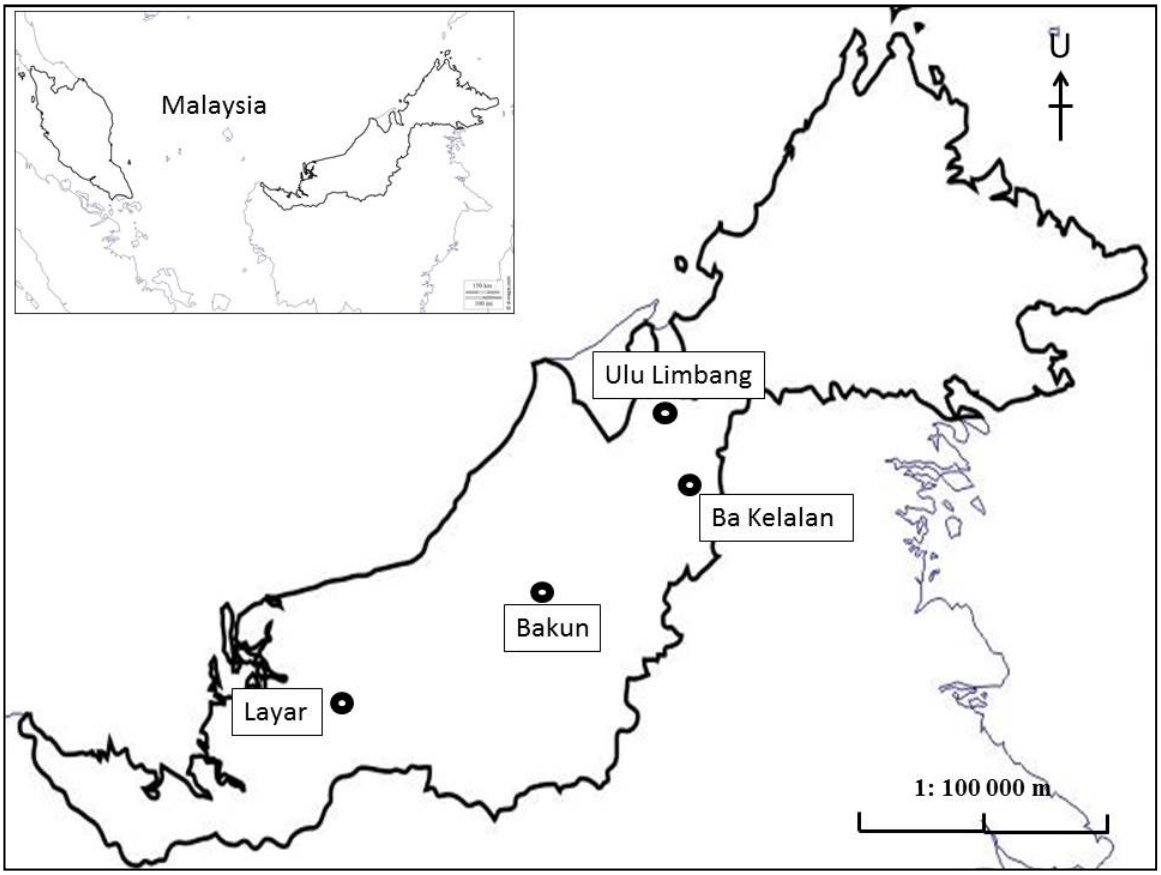

Figure 1. Location of *T. douronensis* populations involved in this study (Source: Google Map).

All samples were collected using cast nets or pole nets. The muscle tissues from each sample were collected and preserved in 95% ethanol and later transported back to Aquatic Molecular Laboratory, Faculty of Resource Science and Technology, UNIMAS for further analyses. The samples were identified using keys provided by Mohsin & Ambak (1983), Kottelat *et al.* (1993) and Inger & Chin (2002).

DNA Extraction and Polymerase Chain Reaction (PCR) and Sequencing

Total genomic DNA of each sample was extracted using the modified Cetyl-trimethyl Ammonium Bromide (CTAB) method (Doyle & Dickson, 1987) with the addition of proteinase K. The amplification of COI gene fragment was conducted using oligonucleotide primers COIf (5'-CCT GCA GGA GGA GGA GAY CC-3', forward) and COIe (5'-CCA GAG ATT AGA GGG AAT CAG TG-3', reverse) (Palumbi *et al.*, 1991). PCR for COI mtDNA gene was carried out following Esa *et al.* (2008) in a programmable gradient-enabled thermocycler (Bio-Rad MyCycler^{TM} Thermal Cycler). Approximately, 50-100 ng of the template DNA was amplified in a 25 µL reaction mixture containing 50 mM 10x buffer, 2 mM $MgCl_2$, 0.4 mM of dNTPs (Promega), 0.2 mM of each primer, and 0.5 U of *Taq* DNA polymerase (Promega). The cycle parameters included pre denaturation step at 95^{o}C for 5 min, followed by 35 cycles of denaturation at 95^{o}C for 30 s, annealing at 47^{o}C for 30 s, and extension at 72^{o}C for 60 s with final extension step at 72^{o}C for 5 min followed by 1% agarose gel electrophoresis. Reverse and forward sequencing were conducted by First Base Laboratories Sdn Bhd, Kuala Lumpur.

The procedures to align the sequences, obtain comparative data to construct the phylogenetic tree, and assign a degree of confidence to the trees was followed from Matsui *et al.* (2005). The sequence was aligned with Clustal X (Thompson *et al.*, 1997) and translated to amino acid sequences in MEGA 4.1 (Kumar *et al.,* 2008).

Maximum-likelihood (ML) and Bayesian analyses were based on the substitution model and phylogenetic parameters (HKY+G) identified as optimal by the Bayesian information criterion (BIC) in jModeltest (Posada, 2008). The confidence levels of the maximum-parsimony (MP) and ML analyses were analysed using PAUP4.0b10 (Swofford, 2000). Only bootstrap values of $\geq$ 70% was regarded as sufficiently resolved topologies (Huelsenbeck & Hillis, 1993) and those of 50%-70% as tendencies. For the Bayesian analysis using MrBayes (Huelsenbeck & Ronquist, 2001), the same substitution model as used for the ML analysis was used with two simultaneous metropolis-coupled Monte-Carlo Markov chains that was run for 106,000 generations or until the probability of split frequencies, $p < 0.01$. A tree was sampled every 100 generations, and a consensus topology was calculated for 795 tree by omitting the first 265 trees (as burn-in). The confidence level of tree nodes was indicated by posterior probabilities which represent the true probabilities of the clades (Rannala & Yang, 1996). Posterior probabilities of $\geq$ 98% was considered significant (Leache & Reeder, 2002).

The measurement of population genetic parameters such as genetic diversity (the probability that two randomly chosen mtDNA sequences differed in the sample) and nucleotide diversity (per nucleotide site, *i.e.*, the probability that two randomly chosen homologous nucleotides differ in the sample (Nei, 1987) was estimated from the mtDNA dataset using DNASP 4.0 (Rozas *et al.*, 2003). Estimates of nucleotide divergence among populations, after accounting for nucleotide diversity within populations (Da), were also generated using DNASP 4.0.

The demographic history was examined by Tajima's test of neutrality, D (Tajima, 1989), and the deviation of sequence variation from evolutionary neutrality was tested using Fu's Fs statistics test (Fu, 1997). A population that has experienced population expansion may result in rejection of the null hypothesis of neutrality of Tajima's D, or a large negative value of Fu's Fs. All analyses were computed in Arlequin vers. 3.00 (Excoffier *et al.*, 2005).

The extent of population differentiation investigation was conducted by dividing the individual into broad geographical groups, based on the ML topology of the mtDNA phylogeny. These groupings were compared by

analysis of molecular variance (AMOVA, Excofficer *et al.*, 2005).

The population expansion events were performed using a mismatch distribution analysis (Roger & Harpending, 1992; Rogers, 1995) using Arlequin vers. 3.0 with 1000 permutations (Excofficer *et al.*, 2005) and site-frequency spectra (Donnelly *et al.*, 2001) as implemented in DNaSP 4.0 (Rozas *et al.* 2003). The parsimony criterion was used to construct haplotype relationships of *T. douronensis,* assuming that differences at any given site between two randomly drawn haplotypes was unlikely to have arisen from more than one mutational step (Alexandrino *et al.*, 2002). In addition, a minimum-spanning network (MSN) was generated using Network 4.5.0.2 (Bandelt *et al.*, 1999) to illustrate these relationships.

RESULTS AND DISCUSSION

A total 465 bp of CO1 gene of *T. douronensis* was successfully amplified using CO1 gene primers. The overall frequency distributions of nucleotides at the 1st, 2nd and 3rd codon position were as follows: T=29.0%, C=19.5%, A=34.0%, and G=17.7%; T=30.0, C=25.4%, A=23.3%, and G=20.9%; T=38.0%, C=23.9%, A=20.3%, and G=18.3%, respectively. A compositional nucleotide bias analysis revealed no significant bias (p=1.29) across the *T. douronensis* populations. There were 23 (4.9%) variables sites with seven singletons, leaving 16 (69.6%) potentially parsimoniously informative characters, indicating that the gene is a reliable marker to infer genetic variations at the population level and 442 (95.1%) site were conserved. Within the dataset, transition occurred more than transversion, in agreement with Briolay *et al.* (1998).

The genetic distance among CO1 sequence of *T. douronensis* ranged from 0.0% to 0.9% (intra-population) and 0.0% to 3.6% (inter-population) indicating high disparities in differentiation in intra-population and inter-population (Table 1). According to Nguyen *et al.* (2006), high inter-population genetic distance could be explained by several factors including small population size, past bottleneck event and physical barriers among populations. In addition, Palumbi (1994) claimed that genetic variation and structure may also increase due to geographical distance. Hence, a population genetic analysis was conducted to further understand the genetic structure of *T. douronensis* from Sarawak.

Among the 11 individuals sequenced and eight GenBank sequences (EF192453, EF192454, EF192455, EF192446, EF192447, EF192444, EF192445), 13 haplotypes were identified with no haplotype shared among them (Table 2). Bakun population consisted of five haplotypes from 11 individuals, none of which was shared with the other three populations (Table 3).

Moreover, low intra population nucleotide diversity (π) was observed ranging from 0.04% to 0.12% (Table 3). Low inter population nucleotide diversity (π) were recorded ranging 0.3% to 1.7% and between 0.00% to 2.72% in net nucleotides divergence (DA) (Table 4). The Da (Table 4) values were high when comparing Ba Kelalan with Bakun (2.72%) than Ba Kelalan with Layar (2.29%) even though both sites are separated by 470 km. A lack of differentiation (zero value of net nucleotide divergence) (Table 5) can be observed between Ulu Limbang and and Ba Kelalan which are separated about 98 km from each other. Similarly, Esa *et al.* (2006) reported population of *T. douronensis* of Peninsular Malaysia was particularly closer to Sarawak population compared to Sabah population although Peninsular and Sarawak are separated by longer distance due to geographical isolation of interconnected river. Thus, this study also suggested that geographical distance between populations is very unlikely to influence the net nucleotide divergence.

Phylogenetic analyses of *T. douronensis* from Sarawak produced similar tree topologies for Maximum Parsimony (MP) (Figure 2), Neighbour-joining (NJ) (not shown), Bayesian Inferences (BI) (Figure 3) and Maximum likelihood (ML) (not shown). They revealed a monophyly of *T. douronensis* from Sarawak with respect to outgroups *Neolssochilus stracheyi* and *Hampala dispar* with bootstrap value of 98% (MP), 99% (NJ), 100% (ML) and 1.0 Bayesin posterior probability of (BPP) in BI. Similarly, Esa *et al.* (2006) and Esa *et al.* (2008) also reported *T. douronensis* is monophyletic.

Table 1. Summary of genetic distance in percentage (%) within and among populations for CO1 gene sequences of *T. douronensis.*

	Bakun	Layar	Ba Kelalan	Ulu Limbang
Bakun	0.0-0.7	-		
Layar	1.1-2.7	0.4-0.9	-	
Ba Kelalan	2.9-3.6	2.7-3.1	0.0-0.4	-
Ulu Limbang	2.7-3.4	2.2-3.1	0.2-0.4	0.0-0.2

The phylogenetic tree re-construction based on CO1 gene variation also reveals that the *T. douronensis* is divided into three geographical clades. Clade I is comprises all *T. douronensis* from Bakun with, 84%, 86%, 83%, 0.57 for MP, NJ, ML and BI respectively. Clade II consists of population from Layar (86%, 74%, 100%, 0.63). and clade III consists of Ba Kelalan and Ulu Limbang population are well resolved with high bootstrap values (98%, 99%, 100%, 0.93). Both clade I and II are considered as sufficiently resolved (Huelsenbeck & Hillis, 1993) as the bootstrap values more than 70%.

From the network analysis (Figure 4), minimum-spanning network (MSN) model showed that there were three haplogroups among the population. Haplogroup I is consisted of Bakun population, haplogroup II consist of Layar samples and haplogroup III consist of Ba Kelalan and Limbang population with no sharing of haplotype although Layar and Bakun population were interconnected. Missing haplotypes (mv) can be observed in Figure 4, suggesting future work needs to involve more samples in the analysis and suggested the network was not fully resolved. The MSN result is consistent with the topologies of all trees (Figure 2 & 3).

The scatterplot of the mismatch distributions analysis in Bakun population illustrates a unimodal interpretation of the mismatch distribution, while *T. douronensis* population from Layar illustrate multimodal interpretation (Figure 5). A unimodal distribution is indicative of population expansion. Thus this suggest that the Bakun population passed through a recent demographic expansion (Rogers & Harpending, 1992; Hudson & Slatkin, 1991) or a range expansion with high level of migration between neighbouring demes (Ray *et al.* 2003; Excoffier, 2004). On the other hand, a multimodal distribution shows that populations are at demographic equilibrium (Rogers & Harpending, 1992) and reflects a highly stochastic shape of the evolutionary lineages as observed in Layar. In this study, mismatch distribution scatterplots for Ulu Limbang and Ba Kelalan populations could not be illustrated due to the small variance of the mismatch distribution thus no demographic parameters could be estimated for these two populations.

Furthermore, mismatch distribution of pairwise nucleotide differences among CO1 gene sequences in Bakun population showed small SSD value is 0.007, which is lack of significance value (p= 0.652) (Table 6) suggesting population expansion hypotheses following the expected distribution under a sudden expansion model (Rogers & Harpending, 1992; Hudson & Slatkin, 1991). However, higher SSD with significance value (SSD= 0.297, p= 0.04) showed that the Layar population has no recent sudden expansion and the population are at demographic equilibrium (Rogers & Harpending, 1992). Small values of Harpending's raggedness index from Bakun population (r= 0.099) with lack of significance (p=0.626), as shown on Table 6, revealed unimodal interpretation of the mismatch distribution in *T. douronensis*. In contrast, small Harpending's raggedness index in the Layar

Table 2. Samples of *T. douronensis* (*T. d*) analysed for (CO1) gene sequence variation with locality, GPS reading, field voucher and identified haplotype.

Species	Haplotype	Locality	GPS reading	Field Voucher/ GenBank
T. d	Hap_1	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK01
T. d	Hap_3	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK02
T. d	Hap_1	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK03
T. d	Hap_1	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK04
T. d	Hap_1	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK05
T. d	Hap_2	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK06
T. d	Hap_4	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK07
T. d	Hap_2	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK08
T. d	Hap_1	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK09
T. d	Hap_1	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK10
T. d	Hap_5	Bakun, Sarawak	N02°45′23″,E114°03′47″	BK11
T. d	Hap_6	Layar, Sarawak	NA	EF192454 (LA01)
T. d	Hap_9	Layar, Sarawak	NA	EF192454 (LA02)
T. d	Hap_12	Layar, Sarawak	NA	EF192455 (LA03)
T. d	Hap_13	Layar, Sarawak	NA	EF192456 (LA04)
T. d	Hap_7	Ba Kelalan , Sarawak	NA	EF192446 (BA01)
T. d	Hap_8	Ba Kelalan, Sarawak	NA	EF192447 (BA02)
T. d	Hap_10	Ulu Limbang, Sarawak	NA	EF192444 (LI01)
T. d	Hap_11	Ulu Limbang, Sarawak	NA	EF192445 (LI02)

Table 3. Segregating sites (23 bp) in the 465 bp of CO1 gene of *T. douronensis.*

Haplotype	Nucleotide position																								Locality				HG
	1	2	3	4	5	6	7	8	9	1 0	1 1	1 2	1 3	1 4	1 5	1 6	1 7	1 8	1 9	2 0	2 1	2 2	2 3		B	L	Li	Bl	
Hap_1	T	T	T	G	T	T	C	G	C	C	C	G	C	C	C	C	G	A	T	A	T	C	C		6	-	-	-	1
Hap_2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G		2	-	-	-	1
Hap_3	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		1	-	-	-	1
Hap_4	.	.	.	.	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		1	-	-	-	1
Hap_5	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		1	-	-	-	1
Hap_7	C	.	G	A	.	.	A	.	T	T	T	.	.	.	.	.	.	G	C	G	C	T	G		-	-	-	1	3
Hap_8	C	C	G	A	.	.	A	.	T	T	T	.	.	.	.	.	.	G	A	G	C	T	G		-	-	-	1	3
Hap_10	C	.	G	A	.	.	A	.	T	T	T	.	.	.	.	.	.	G	A	G	C	T	G		-	-	1	-	3
Hap_11	C	.	G	A	.	.	A	.	T	T	T	.	.	.	.	.	.	G	C	G	C	T	G		-	-	1	-	3
Hap_6	.	.	A	G	.	.	.	.	.	T	.	.	.	.	T	.	.	C	.	C	.	G			-	1	-	-	2
Hap_9	C	.	A	G	.	.	.	.	.	T	.	T	.	.	T	.	.	C	.	C	.	G			-	1	-	-	2
Hap_12	C	.	A	G	.	.	.	.	.	T	.	T	T	.	T	A	.	C	.	C	.	G			-	1	-	-	2
Hap_13	C	.	A	G	.	.	.	.	.	T	.	T	T	T	T	.	.	C	.	C	.	G			-	1	-	-	2

B = Bakun; L = Layar; Li = Limbang. Bl = Ba Kelalan, HG = Haplogroup.

Table 4. Measures of haplotype and nucleotide diversity within populations of *T. douronensis* analysed by location.

Locality	H	Percent (%) pairwise divergence[a]	Gene diversity	Nucleotide diversity
Bakun	5	0.0-0.7	0.0071 ± 0.0013	0.0005 ± 0.03
Layar	4	0.4-0.9	0.0010 ± 0.0018	0.0012 ± 0.09
Ba Kelalan	2	0.4	0.0010 ± 0.0050	0.0008 ± 0.11
Ulu Limbang	2	0.2	0.0010 ± 0.0050	0.0004 ± 0.06

H = number of haplotypes.

Table 5. Measures of nucleotide diversity (π) and net nucleotide divergence (Da) among populations of *T. douronensis* analysed by location.

Locality	Distance (km)	Nucleotide diversity (π)	Net Nucleotide divergence (Da)
Bakun-Layar	248	0.0093	0.0141
Bakun-Ulu Limbang	280	0.0096	0.0262
Bakun-Ba kelalan	215	0.0102	0.0272
Layar-Ulu Limbang	540	0.0156	0.0208
Layar-Ba kelalan	470	0.0175	0.0229
Ulu Limbang-Ba kelalan	98	0.0032	0.0000

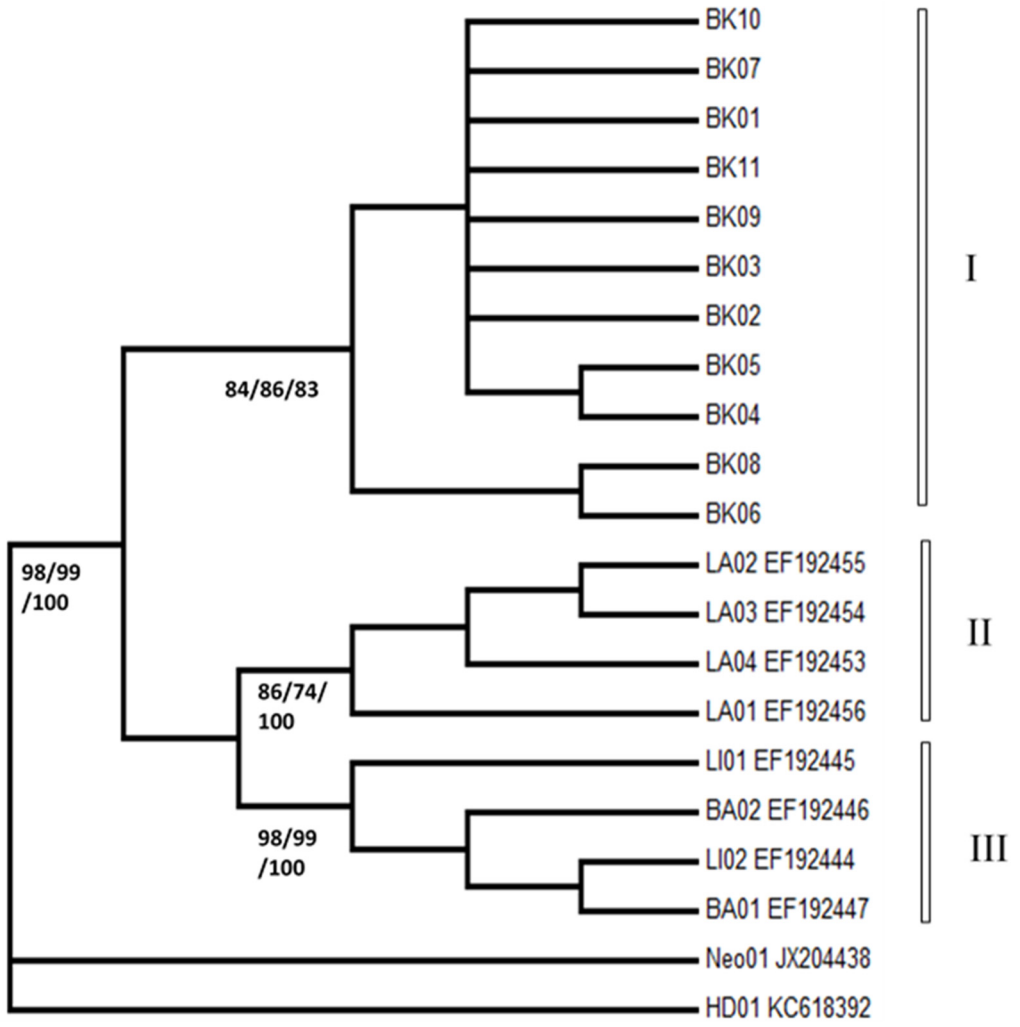

Figure 2. A maximum parsimony 50% majority rule consensus tree constructed of CO1 gene sequence of *T. douronensis* from Sarawak with *H. dispar and N. stracheyi* as outgroups. Bootstrap values are indicated above branch corresponding to MP, NJ and ML, respectively. Tree length is 341 with consistency index (CI) = 0.9724 and retention index (RI) = 0.8601.

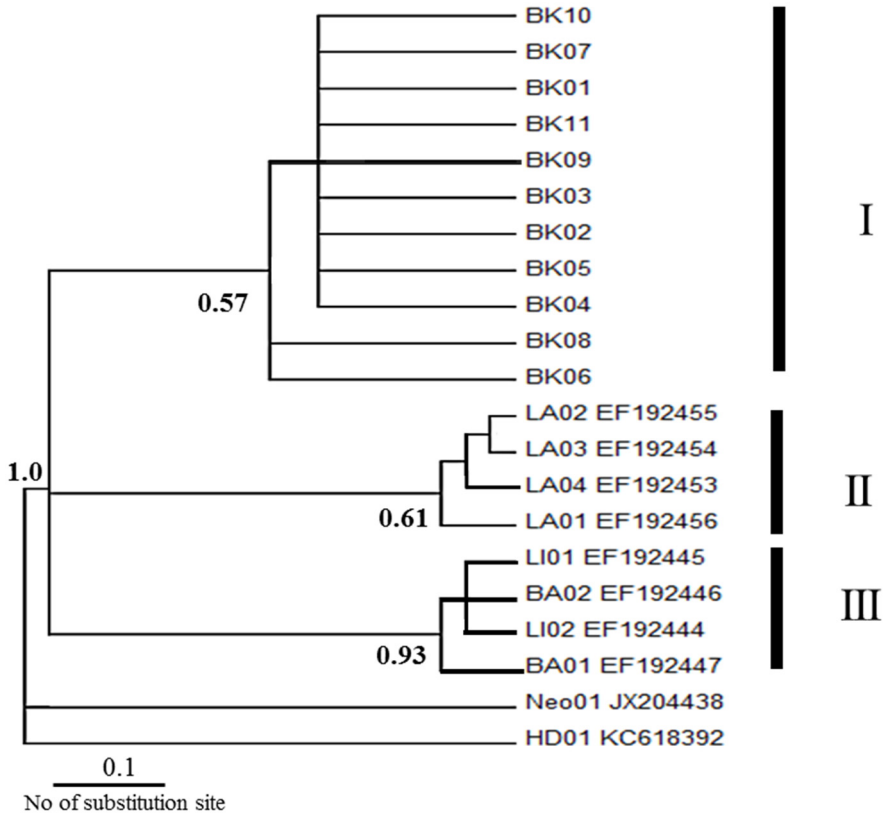

Figure 3. Bayesian inference of the 50% majority rule consensus tree of CO1 gene sequences of *T. douronensis* with *H. dispar and N. stracheyi* as outgroup.

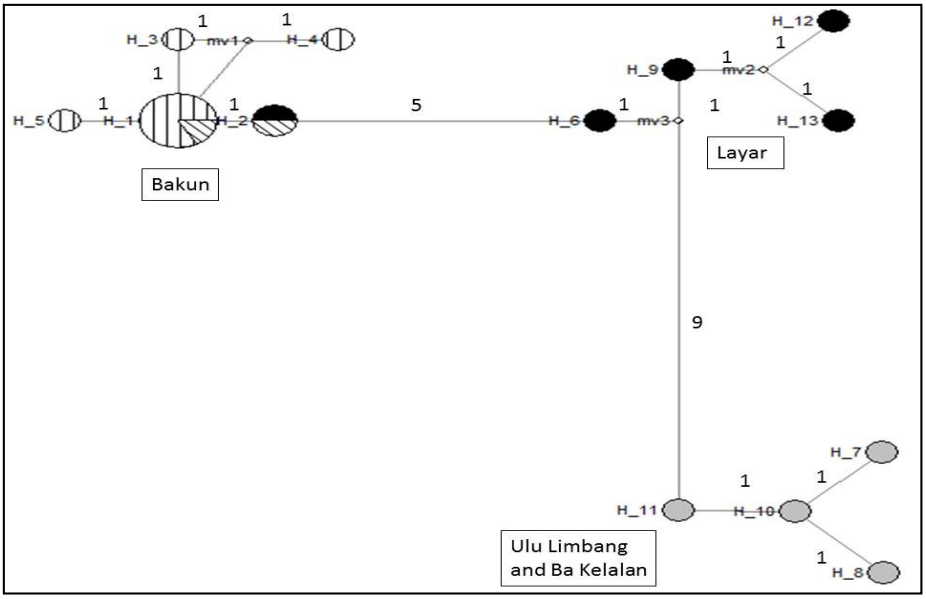

Figure 4. The minimum-spanning network (MSN) generated by Network 4.5.0.2 illustrating the relationships of *T. douronensis* in Sarawak. Each circle represents a haplotype, and the diameter is scaled to the haplotype frequency. Note that diagonal circles indicate Bakun haplotypes, black circles indicate Layar haplotype, and dark gray indicates Ulu Limbang and Ba Kelalan haplotypes. Numbers on the lines connecting haplotypes indicate number of mutational steps.

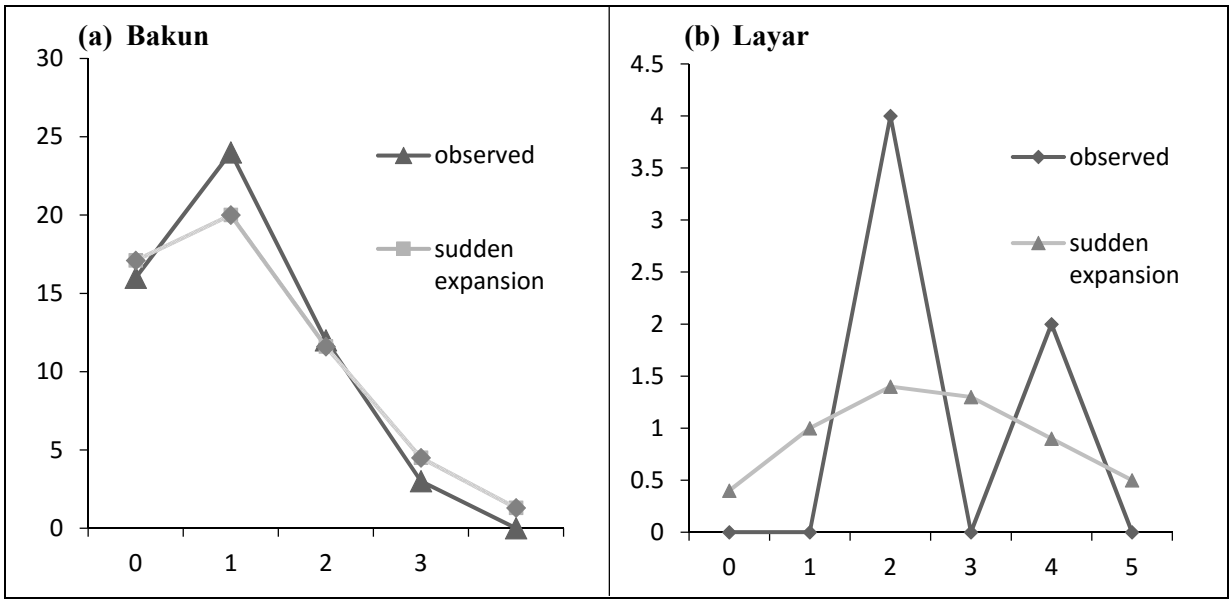

Figure 5. Mismatch distribution for (a) Bakun showing a unimodal distribution and (b) Layar showing multimodal distribution.

population (r=1.111) with significant value (p=0.05) (Table 6), revealed multimodal interpretation of the mismatch distribution in *T. douronensis*. A unimodal distribution is indicative of population expansion whereas a multimodal distribution indicates constant population size over a long time period (Rogers & Harpending, 1992).

To examine the demographic history and evolutionary neutrality of *T. douronensis* population, Tajima's test of neutrality (Tajima, 1989) and Fu's Fs statistical tests were conducted. Based on Table 7, the Tajima test was not significant with p more than 0.05 showing no deviation from evolutionary neutrality for all populations (Tajima, 1989). The scatterplot of the mismatch could not be generated for Ulu Limbang and Ba Kelalan populations might be due to no polymorphism was examined in the samples. The insignificant shown in p value from Ulu Limbang and Ba Kelalan (p=1.0) from Tajima's test (Table 5) also revealed that there had been no recent expansion occurring at Ulu Limbang and Ba Kelalan *T. douronensis* populations.

Moreover, the negative value of Fu's Fs statistics for Bakun population showed a negative (-1.20) and significant value (Table 6) indicating the presence of unique haplotype in the population (Zainudin *et al.*, 2010) and inferring the population had undergone expansion or genetic hitchhiking (Fu, 1997).

AMOVA analysis (Table 7) shows high and significant differentiation (86.88%) among *T. douronensis* populations in Sarawak. Furthermore, variation within population, although lower (13.12%) is still significant, indicating presence of variation among individuals within a (or a few) population(s). This is likely contributed by Bakun and Layar populations as the pairwise divergence and nucleotide diversity of these two particular populations are high.

Genetic differentiation matrix was estimated to observe the genetic structure among *T. douronensis* populations in Sarawak. The estimated ΦST values among populations of *T. douronensis* showed significance except for the Ulu Limbang and Ba Kelalan (Table 8) populations suggesting closed relationship and potential high gene flow occurrence among these two populations while other pairwise population comparisons were low indicative of relatively low gene flow occurrence.

The high levels of nucleotide subdivision, N_{st} and F_{ST} value of 0.00 with high level of migrant per generation (Nm=-1.00.) between Ba Kelalan and Ulu Limbang populations (Table 9) suggesting that there is high gene flow between them.

Table 6. Summary statistics of gene mtDNA sequence variation in four populations of *T. douronensis* in Sarawak.

Population	H	% sdiv	D	Fs	Sudden expansion	
					SSD	r
Bakun	5	0.0-0.7	-0.88	-1.20**	0.007	0.099
			(p=0.21)	(p=0.02)	(p=0.652)	(p=0.626)
Layar/Spak	4	0.4-0.9	-0.21	-1.41	0.297	1.111
			(p=0.55)	(p=0.06)	(p=0.041)	(p=0.052)
Ba Kelalan	2	0.4	0.0	0.69	NA	NA
			(p=1.00)	(p=0.35)		
Ulu Limbang	2	0.2	0.0	0.0	NA	NA
			(p=1.00)	(p=0.26)		

H = number of haplotypes, sdiv = pairwise sequence divergence (estimated using the HKY distance (Hasegawa *et al*., 1985), D = Tajima's statistic (P (D simul< D obs), (Tajima, 1989), SSD = sum of squared deviations of the observed and expected mismatch with p values in parentheses, r = raggedness statistic (Harpending, 1993 with p values in parentheses; Fs = Fu's statistic (Fu, 1997). ** Significance ($p< 0.05$) was determined using coalescent simulations in Excoffier (2005).

In contrast, Bakun and Limbang populations were the most differentiated and slightly isolated (N_{ST} or F_{ST} >0.9) with low gene flow (Nm=0.04) to (Table 9). The Bakun with Ba Kelalan and Ba Kelalan with Layar population pairwise are (N_{ST}= 0.89, F_{ST}= 0.89, Nm=0.06) and (N_{ST}= 0.79, F_{ST}= 0.79, Nm=0.14) respectively. The result also revealed that *T. douronensis* population from Layar with Ulu Limbang was closer (N_{ST}= 0.84, F_{ST}= 0.84, Nm=0.09) compare to Bakun with Limbang population although the distance between relationships of the four popualations *T.* Layar-Limbang is 540 km while Bakun- Limbang is only 280 km. Thus, the genetic *douronensis* are less likely related to geographical distance. But it is not fully resolved since the significance value cannot be generated due to low samples size.

Zainudin *et al.* (2010) reported that geographical barrier can also be an important factor for the high genetic divergence. The existence of large river, vast land area and isolated river reflect the species dispersal. This could also account for the observations in this study as evident by the AMOVA. Specifically, migration was hindered by geographical barrier except during the last Tertiary and Quaternary historical periods (Inger & Chin, 2002) where the southern and northern part of Sarawak river system and channel were interconnected.

CONCLUSION AND RECOMMENDATION

A total of 465 bp of the CO1 gene was successfully amplified from four populations of *T. douronensis* in Sarawak river systems revealing high genetic divergence values among populations. Three population subdivisions were observed representing central, southern, and northern population; 1st clade (haplogroup I) from Bakun, 2nd clade (haplogroup II) from Layar and 3rd clade (haplogroup III) from Ba Kelalan and Ulu

Table 7. Measures of geographical population differentiation in *T. douronensis* based on an analysis of molecular variance approach with cytochrome c oxidase subunit 1 data.

Source of variation	Variance component	Percent (%) variation	Fixation index, Φ	p^a
Among populations	4.72	86.88	0.87	0.00 ± 0.00*
Within populations	0.71	13.12	0.87	0.00 ± 0.00*

*Significant ($p < 0.05$). [a]Probability of finding a more-extreme variance component of the Φ index than that observed by chance alone after 1000 permutations.

Table 8. Genetic differentiation matrix of populations calculated by ΦST. *P* values are shown in parenthesis (below the diagonal).

	Bakun	Layar	Limbang	Ba Kelalan
Bakun	-			
Layar	0.83 (0.00 ± 0.00)*	-		
Limbang	0.92 (0.00 ± 0.00)*	0.81 (0.00 ± 0.00)*	-	
Ba Kelalan	0.92 (0.01 ± 0.00)*	0.81 (0.08 ± 0.01)	0.00 (0.00 ± 0.00)*	-

*Significant ($p < 0.05$) with 1000 permutations.

Table 9. Measures of nucleotide subdivision (N_{ST}), population subdivision (F_{ST}), and gene flow (number of migrants, Nm) among 4 populations of *T. douronensis.*

Locality	Distance (km)	Nucleotide subdivision (N_{ST})[a]	Estimate of population subdivision(F_{ST})[b]	Number of migrant per generation (N_m)[b]
Bakun-Layar	248	0.79	0.79	0.14
Bakun- Limbang	280	0.92	0.92	0.04
Bakun-Ba kelalan	215	0.89	0.89	0.06
Layar-Limbang	540	0.84	0.84	0.09
Layar-Ba kelalan	470	0.82	0.82	0.11

[a]Estimated using Lynch & Crease (1990). [b]Estimated using Hudson *et al.* (1992).

Limbang. Overall, there were 13 haplotypes and none was shared among populations, suggesting low level of inter-population gene flow has been observed. The small number of migrants per generation (Nm<1.0) among the population indicated that the small populations were separated, possibly by large geographical areas. All population had undergone expansion. Furthermore a large negative value and significant test of Fu' Fs in Bakun population suggested recent expansion. The result also suggested that all the populations do not deviate from evolutionary neutrality. Future studies should involve increased number and size of the populations in order to have a greater understanding of the genetic structure of *T. douronensis* in Sarawak.

ACKNOWLEDGEMENTS

Authors would like to thank UNIMAS for the field trip, chemicals and consumable, laboratory facilities and transportation. Special thank you to Prof Dr Lee Nyanti FRST UNIMAS for organizing field trip to Bakun Dam and useful discussions on the morphological aspects of *Tor* spp.

REFERENCES

Alexandrino, J, Jarntzen, J.W., & Ferrand, N. (2002). Nested clade analysis & the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: *Urodela*). *Heredity,* 88: 66-74.

Bandelt, H.J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16: 37-48.

Briolay, J., Galtier, N., Brito, R.M., & Bouvet, Y. (1998). Molecular phylogeny of Cyprinidae inferred from cytochrome *b* DNA sequences. *Molecular Phylogenetic Evolution*, 9(1): 100-108.

Donnelly, M.J., Licht, M.C., & Lehmann, T. (2001). Evidence for recent population expansion in the evolutionary history of the malaria vectors *Anopheles arabiensis* and *Anopheles gambiae. Molecular Biology and Evolution,* 18: 1353-1364.

Doyle, J.J., L. & Dickson, E.E. (1987). Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon,* 36: 715-722.

Esa, Y.B., Siraj SS, Daud, S.K. Rahim, K.A.A. Abdullah, M.T., Japning, J.R.R., & Tan, S.G. (2006). Mitochondrial DNA diversity of *Tor douronensis* Valenciennes (Cyprinidae) in Malaysian Borneo. *Pertanika Journal of Tropical Agriculture Science*, 29: 47-55.

Esa, Y.B., Siraj, S.S., Daud, S.K., Khairul Adha, A.R., Ryan, J.R.J., & Tan, S.G. (2008). Mitochondrial DNA diversity of *Tor tambroides* Valenciennes (Cyprinidae) from five natural populations. *Malaysian Zoology Study,* 47(3): 360-367.

Excoffier, L. (2004). Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology*, 13(4): 853-864.

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin vers. 3.0: an integrated software package for population genetics data analysis. *Evolution Bioinformation Online*, 1: 4750.

Fu, Y.X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics,* 147: 915-925.

Harpending, H.C., Sherry, S.T., Rogers, A.R., & Stoneking, M. (1993). The genetic structure of ancient human populations. *Current Anthropology*, 483-496.

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating the human-ape split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution,* 22: 160-174.

Hudson, R.R., Boos, D.D., & Kaplan, N.L. (1992). A statistical test for detecting geographic subdivision. *Molecular biology and Evolution*, 9(1): 138-151.

Huelsenbeck, J.P. & Hillis, J.E. (1993). Success of phylogenetic methods in the four taxon case. *Systematic Biology*, 42: 247-264.

Huelsenbeck, J.P. & Ronquist, F.R. (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17: 754-755.

Inger, F.R. & Chin, P.K. (2002). *The freshwater fishes of North Borneo*. Kota Kinabalu: Natural History Publication (Borneo).

Ingram, B.A., Sungan, S., Gooley, G.J., Sim, Y.S., Tinggi, D., & De Silva, S.S. (2005). Induced spawning, larval development and rearing of two indigenous Malaysian Mahseer, *Tor tambroides* and *Tor douronensis*. *Aquaculture research*, 36: 1001-1014.

Kottelat, M., Whitten, A.J., Kartikasari, S.N., & Wirjoatmodjo, S. (1993*). Freshwater fishes of Western Indonesia and Sulawesi*. Singapore: Periplus Edition.

Kumar, S., Dudley, J., Nei, M., & Tamura, K. (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefing in Bioinformatics*, 9: 299-306.

Leache, A.D. & Reeder, T.W. (2002). Molecular systematics of the eastern fence lizard (*Sceloporus undulatus*): a comparison of parsimony, likelihood, and Bayesian approaches. *Systematic Biology,* 51: 44-68.

Litis, B.A., Sungan, S., Jugang, K., Ibrahim, M., & Bin, H.A. (1997). *Features of indigenous fish species having potential for aquaculture.* Sarawak: Inland Fisheries Division, Department of Agriculture.

Lynch, M. & Crease, T.J. (1990). The analysis of population survey data on DNA sequence variation. *Molecular Biology and Evolution*, 7(4): 377-394.

Matsui M., Shimada, T., Ota, H., & Tanaka-Ueno, T. (2005). Multiple invasions of the Ryukyu Archipelago by Oriental frogs of the subgenus *Odorrana* with phylogenetic reassessment of the related subgenera of the genus *Rana*. *Molecular Phylogenetic Evolution*, 37: 733-742.

Mohsin, A.K.M. & Ambak, M.A. (1983). *Freshwater fishes of Peninsular Malaysia*. Serdang: Universiti Pertanian Malaysia Publication.

Nei, M. (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.

Ng, C.K. (2004). *Kings of the Rivers: Mahseer in Malaysia and the Region*. Kuala Lumpur: Inter Sea Fishery.

Nguyen, T.T.T., Ingram, B., Sungan, S., Gooley, G., Sim, S.Y., Tinggi, D., & De Silva, S.S. (2006). Mitochondrial diversity of broodstock of two indigenous mahseer species*, Tor tambroides* and *Tor douronensis* (Cyprinidae) cultured in Sarawak. *Aquaculture,* 253: 259-269.

Nyanti, L., Yee, L.T., & Adha, K. (1999). *Freshwater fishes from Bario, Kelabit Highlands.* Asean review in biodiversity and environmental conservation. Pp 1-6.

Palumbi, S., Martin, A., Romano, S., McMillan, W.O., Stice, L., & Grabowski, G. (1991). *The simple fool's guide to PCR.* Hawaii: Department of Zoology & Kewalo Marine Laboratory, University of Hawaii.

Posada, D. (2008). jModelTest: phylogenetic model averaging. molecular biology and evolution. *Molecular Biology and Evolution,* 25: 1253-1256.

Rainboth, W.J. (1996). *Fishes of the Cambodia*, *Mekong.* FAD species identification field guide for fishery purposes. Rome: Food and Agriculture Organization (FAO) Publication.

Rannala, B. & Yang, Z. (1996). Probability distribution of molecular evolution trees: a new method of phylogenetic inference. *Journal of Molecular Evolution*, 43: 304-311.

Roberts, T.R. (1989). *The freshwater fishes of Western Borneo (Kalimantan Barat, Indonesia).* California: California Academy of Sciences.

Rogers, A.R. (1995). Genetic evidence for a Pleistocene population explosion. *Evolution,* 49: 608-615.

Rogers, A.R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, 9: 552-569.

Rozas, J., Sanches-DelBarrio, J.C., Messeguer, X., & Rozas, S. (2003). DNASP, DNA polymorphism by the coalescent & other methods. *Bioinformatics*, 19: 2496-3497.

Swofford, D.L. (2000). *PAUP*. Phylogenetic Analysis Using Parsimony (*& other methods). Vers. 4.0b10*. Sunderland, MA: Sinauer Associates.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123: 585-595.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., & Higgins, D.G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 24: 4876-4882.

Ray, N., Currat, M., & Excoffier, L. (2003). Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution*, 20(1): 76-86.

Slatkin, M. & Hudson, R.R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, 129(2): 555-562.

Zainudin, R., Nor, S.M., Ahmad, N., Md-Zain, B.M., & Rahman, M.A. (2010). Genetic structure of *Hyalarana erythraea* (Amphibia: Anura: Ranidae) from Malaysia. *Zoological Studies,* 49(5): 688-702.