Khalili Samani N, Esa Y, Amin SMN, Fatin Mohd Ikhsan N. 2016. Phylogenetics and population genetics of *Plotosus canius* (Siluriformes: Plotosidae) from Malaysian coastal waters. PeerJ 4:e1930 https://doi.org/10.7717/peerj.1930

Phylogenetics and Population Genetics of *Plotosus canius* (Siluriformes: Plotosidae) from Malaysian Coastal Waters

Nima Khalili Samani*, Yuzine Esa, Natrah Fatin Mohd Ikhsan, S.M Nurul Amin

Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

*Corresponding Author:

Nima Khalili Samani

Graduate Researcher, Department of Aquaculture, Faculty of Agriculture

Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

nima.khalili.samani@gmail.com

Tel: +60129175937. Fax: +603-8940 8311

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

ABSTRACT: *Plotosus canius* (Hamilton, 1822) is a significant marine species in Malaysia from nutritional and commercial perspectives. Despite numerous fundamental researches on biological characteristics of *P.canius*, there are various concerns on the level of population differentiation, genomic structure, and the level of genetic variability among their populations due to deficiency of genetic-based studies. Deficiency on basic contexts such as stock identification, phylogenetic relationship and population genetic structure would negatively impact their sustainable conservation. Hence, this study was conducted to characterize the genetic structure of *P.canius* for the first time through the application of mitochondrial Cytochrome Oxidase I (COI) gene, cross amplification of *Tandanus tandanus* microsatellites, and a total of 117 collected specimens across five selected populations of Malaysia. The experimental results of the mitochondrial analysis revealed that the haplotype diversity and nucleotide diversity varied from 0.395 to 0.771 and 0.033 to 0.65 respectively. Moreover, the statistical analysis of microsatellites addressed a considerable heterozygote insufficiency in all populations, with average observed heterozygosity (*H0*) value of 0.2168, which was lower than the standard heterozygosity in marine populations (*H0*= 0.79). This alongside the high *Fis* values estimation, high pairwise differentiation among populations and low within population variations are supposed to be associated with small sample size, and inbreeding system. Besides, the significant finding of this study was the sharing of common haplotype KR086940 at which reflects a historical genetic connectivity between Peninsular Malaysia and Borneo populations due to the geological history of Southeast Asia during Pleistocene era. To put it briefly, the current study has managed to provide an initial genomic database toward understanding of the genetic characterization, phylogenetic, molecular diversification and population structure in *P.canius*, and should be necessary highlighted for appropriate management and conservation of species. Though, further studies must be carried out involving more geographical and sampling sites, larger population size per site, and utilization of more COI genes and nuclear hypervariable markers.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

INTRODUCTION

Plotosus canius (Hamilton,1822) that is known as grey eel-catfish, black eel-tail catfish, canine catfish or Indian catfish (Khan *et al*., 2002; Riede, 2004; Usman *et al*., 2013; Prithiviraj, 2014), were diagnosed as a member of genus *Plotosus*, family Plotosidae (Froese & Pauly, 2015). They are being mainly distributed in estuaries, freshwater rivers, lagoons, and shallow waters of Australia and Southeast Asia (Carpenter, 1999; Prithiviraj & Annadurai, 2012). The species is an amphidromous and demersal bony fish that can live in marine, brackish and freshwater habitats. Their relocation in about 100 km range were described as cyclical and frequent horizontal movement on which could not be categorized as breeding migration (Riede, 2004). However, the species might be recently endured genetic destruction mostly due to overexploitation similar to other fish species (Pauly *et al*., 2002; Collette *et al*., 2011; Usman *et al*., 2013), their population structure could be considered as the reliable indicator in detection of sustainable and healthy marine environments (Thomsen *et al*., 2012; Bourlat *et al*., 2013).

Population structure is the direct consequence of biogeography (Leffler *et al*., 2012), which provides invaluable statistics on patterns of species dynamics, colonization, and isolation (Costello *et al*., 2003). As species become accustomed into new habitat, effective size of population extends through its dispersal, resulting in intensification of genetic variation (Charlesworth & Willis, 2009). However, deterioration of environmental equations alongside with ecological fluctuations such as recent re-treatment of Pleistocene era have changed species extensive genetic patterns (Krishnamurthy & Francis, 2012). Adding to complication, the accuracy of associated conservation strategies can be successively restrained by deficiency of reliable knowledge on biodiversity, conservation resolution, and extent of biological destruction among taxonomic levels (Wright *et al*., 2008; Butchart *et al*., 2010; Magurran *et al*., 2010; Pereira *et al*., 2010; Hoffmann *et al*., 2011). Such scarcities confidently offered a viable incentive to regulate the sustainable species variation (Primack, 2002; Duvernell *et al*., 2008; Appeltans *et al*., 2012; Bourlat *et al*., 2013; Leray & Knowlton, 2015) through advancing genomic protocols to challenge the genetic intimidations such as distraction of local traits, genetic drift and inbreeding effects (Tallmon *et al*., 2004).

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

The same conservation obstacle is hypothetically threatening *P.canius* populations in Malaysia, since there is not any comprehensive documentation nor a single initial research on their genetic characterization. Indeed, regarding to their regional significance in Oceania and Southeast Asia (Usman *et al*., 2013), a few regional studies have been merely carried out on basic biological perceptions of *Plotosus canius* including their morphology and fisheries (Kumar, 2012; Usman *et al*., 2013), fecundity (Khan *et al*., 2002; Usman *et al*., 2013), feeding behaviour (Leh *et al*., 2012), protein structure (Prithiviraj & Annadurai, 2012) and pharmacology (Prithiviraj, 2014). Such deficiencies have raised some severe concerns on the level of population structure, genetic variation and the consequences of genetic differentiation among populations of *P.canius* especially in Malaysia. Thus, this study was performed to genetically characterize *P.canius* through the utilization of the mitochondrial Cytochrome Oxidase I (COI) gene and *Tandanus tandanus* microsatellites, in order to examine the accuracy of employed markers in phylogenetic study, genetic variation assignment, and population genetic structure of *P.canius* in Malaysia.

MATERIALS AND METHODS

Sample Collection and DNA Isolation

Total of 130 catfish samples demonstrating 2 species of family Plotosidae were directly collected including of 117 samples of *Plotosus canius* and 13 samples of *Plotosus lineatus*. Sample collection of *P.canius* were performed in 5 various districts throughout Malaysia including: Negeri Sembilan (NSN), Selangor (SGR), Johor (JHR), Sarawak (SWK) and Sabah (SBH) (Figure.1) from May to December 2014, while *P.lineatus* samples were only collected from Selangor. Sampling was carried out directly from fishermen in commercial fishing docks of Port Dickson (Negeri Sembilan), Kuala Selangor (Selangor), Kukup (Johor), Bintulu (Sarawak) and Putatan (Sabah). DNA extraction protocol was performed upon sample collection in laboratory via the Wizard® SV Genomic DNA Purification System (Promega, USA), according to manufacturer's protocol instructions by using roughly 20 mg of specimens.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

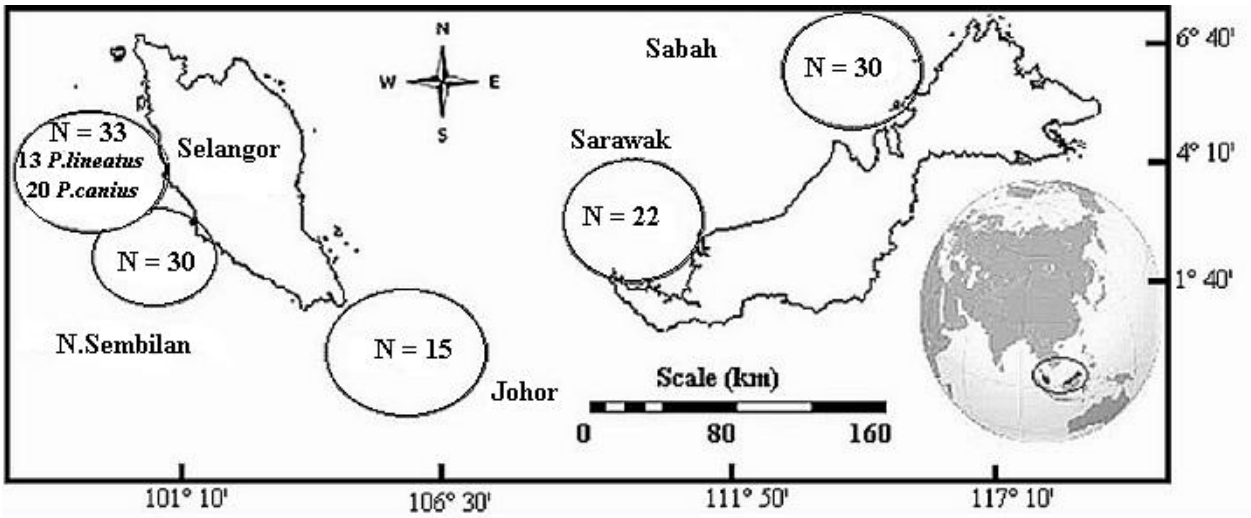

Figure. 1) Sampling sites and sample size (N) diagram of *P.canius* and *P.lineatus* in Malaysia.

PCR Amplification and Sequencing of Mitochondrial DNA

To accurately amplify a 655 bp fragment of mitochondrial DNA, PCR amplification were performed by using C_Fish1 primer set (Ward *et al*., 2005) on the 5′ end of COI gene. The amplification protocol were conducted in an overall volume of 25 µl at which contained 0.6 µl of each deoxynucleotide triphosphate (dNTPs), 0.7 µl Taq polymerase, 1.2 µl MgCl2, 0.25 µl of each primers, 2 µl of concentrated genomic DNA, 2.5 µl of Taq buffer and 18 µl of distilled H2O as stated by Ward *et al*. 2005, with slight modification. The PCR reaction was carried out using an Eppendorf Mastercycler based on the following thermal regime: 2 min of 95 °C initial denaturation step; 35 cycles of 94 °C denaturation step for 30 s, a 54°C annealing temperature for 45 s and a 72°C extension period of 1 min; followed by 72 °C final extension step for 10 min and a routine 4 °C final hold (Ward *et al*., 2005). In order to confirm that the PCR reaction generated sufficient amplicon proportions, PCR amplification products were visualized using a 2.0 per cent laboratory grade agarose gel containing 5 µl GelRed stain. Amplified products were subsequently isolated and purified upon their visualization and documentation. DNA purification from gel was commonly carried out using the Wizard® SV Gel and PCR Clean-up System (Promega, USA). Purified DNA samples were finally sent to private sector institution (1st Base laboratories Sdn Bhd) for sequencing to generate associated trace files and continuous read lengths intended for genetic and statistical analysis of mitochondrial DNA.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

Statistical Analysis of Mitochondrial DNA

Trace files were manually end-trimmed using BioEdit software 7.2.5 (Hall, 1999) regarding to their homologous section. Afterwards, ClustalX 2.1 (Thompson *et al*., 1997) was applied to progressively manipulate, align and analyze the DNA sequences. Finally, haplotypes were detected with DnaSP software 5.10.01 (Librado & Rozas, 2009) and deposited into BOLDSYSTEM (international barcode of life) and GenBank.

To comparatively analyze mitochondrial DNA sequences, MEGA 6 program (Tamura *et al*., 2013) were used intending to understand the phylogenetic outlines of COI gene and species, generate sequence alignment and perform evolutionary analysis. Calculation of the pairwise distance was obtained through 1000 bootstrap variance estimation and Tamura-Nei model (Tamura & Nei, 1993). Moreover, overall mean nucleotide distance of sequences was computed using same configuration at each codon positions separately. Subsequently, construction of phylogenetic tree from the highest grade aligned sequences of *P.canius* and *P.lineatus* was prompted in comparison to one haplotype of African sharp tooth catfish *Clarias gariepinus* (ANGBF8254-12) from Thailand as an outgroup through Neighbor-Joining (NJ) and Maximum Likelihood (ML) methods using a mutual 1000 replication bootstrap. Next, Minimum Spanning Network (MSN) was computed using PopART (Bandelt *et al*., 1994) application among obtained sequences of *P.canius*.

Extraction of genetic features from assembly of sequences based on some rudimentary implemented analytical tests was performed through Arlequin software 3.5 (Excoffier *et al*., 2005). As the most crucial objective was to compute the genetic structure, hierarchical analysis of molecular variance (AMOVA) and pairwise F_{ST} values of chi square test, population differentiation was successively calculated among five populations of *P.canius* in Malaysia. Analysis of molecular variance was carried out using 1000 permutation to compute distance matrix of sequences, while the same structure were implemented for comparison of all available pair samples and populations through calculation of F_{ST} with 0.05 significance level.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

164 Microsatellite Genotyping

165

166 To run measured PCR protocol, optimization of PCR composition were performed at which
167 made positive outcomes based on initial PCR regime using the main protocol for the
168 amplification of *Tandanus tandanus* (Rourke *et al*., 2010) with minor regulation in the amount
169 and concentration of primers, *Taq* DNA polymerase, MgCl2 and dNTPs to enhance the accuracy
170 of the protocol. However, the optimum annealing temperature were practically calculated 55°C
171 similar as original species. To consciously amplify DNA fragments, PCR amplification were
172 performed by using five cross-amplified microsatellites of *T.tandanus* (Rourke *et al*., 2010;
173 Rourke & Gilligan, 2010) on the 5′ end, presented in Table.5. The feasible amplification
174 protocol were conducted in a total volume of 25 µl solution inclosing 0.6 µl of each
175 deoxynucleotide triphosphate (dNTPs), 0.7 µl *Taq* DNA polymerase, 1.2 µl MgCl2, 0.25 µl of
176 each primers, 2 µl of concentrated genomic DNA, 2.5 µl of Taq buffer and 18 µl of distilled
177 H2O as stated by Rourke & Gilligan (2010) with slight modification.

178

179 The PCR reaction was carried out using a gradient Eppendorf Mastercycler based on the
180 following thermal adjusted protocol: 2 min of 95 °C initial denaturation step; 35 cycles of 95 °C
181 denaturation step for 30 s, a 55°C annealing temperature for 45 s and a 72°C extension period of
182 1 min; followed by a 72 °C final extension (elongation) step for 10 min and a routine 4 °C final
183 hold (Rourke *et al*., 2010; Rourke & Gilligan, 2010). Afterwards, PCR amplification products
184 were visualized using a 4.0% MetaPhor agarose gel containing 5 µl GelRed staining solution.
185 Subsequently, gel images were subjected to microsatellite screening and approximately 15 µl of
186 the florescent label products were packed and sent to First Base laboratories (private institution)
187 for fragments analysis.

188

189 Genetic Analysis of Microsatellite markers

190

191 In order to verify any null alleles and scoring error, MICROCHECKER 2.2.3 (Van Oosterhout *et*
192 *al*., 2004) were applied using diploid data obtained from CONVERT software (Glaubitz, 2004).
193 GENEPOP 4.2 (Rousset, 2008) was employed in order to evaluate the conformity to the "Hardy-
194 Weinberg expectations" (HWE) with 10000 permutations for test of exact probability. Observed

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

heterozygosity *(HO)* was estimated using GENALEX 6.5 (Peakall & Smouse, 2012). Successively, FSTAT 2.9.3.2 (Goudet, 1995) was applied to calculate the expected heterozygosity (*HE*), with 15000 permutation and MolKin 3.0 (Gutiérrez *et al*., 2005) to validate the genetic analysis among genetic dataset using Polymorphism Information Content (PIC). Afterwards, ARLEQUIN 3.0 (Excoffier *et al*., 2005) were used to analyze the genetic configuration, hierarchical molecular analysis (AMOVA) and pairwise F_{ST} estimations among all five involved populations, using reflection of 95% significance level and 10000 permutations. Assignment of each individual to their genetic groups (*K*) employing admixture model and its associated frequency of allelic data was carried out using the STRUCTURE program 2.1 (Pritchard *et al*., 2000). Next, GENECLASS2 2.0 (Piry *et al*., 2004) was implemented to conduct assignment of individuals into the most plausible inheritance group. Finally, probability of current bottleneck was tested using BOTTLENECK software 1.2.02 (Piry *et al*., 1999).

RESULTS

Phylogenetic and Population Analysis inferred from Mitochondrial DNA

Among overall number of 130 studied specimens using one COI gene, a full mass of 118 reliable sequences (approximately 91%) were identified. Once the sequencing was completed, the C_Fish1 accordingly confirmed as suitable primer set amplifying in *P.canius* and *P.lineatus* DNA samples. Nevertheless, in some cases there were some uncertain base calls, there was no observation of any stop codons nor any instances of insertion or deletion in sequences. Deficiency of structural stop codon more likely supposed to be associated with every amplified mitochondrial sequences, and all that, alongside with the read length of amplified sequences implies that nuclear sequences initiated from vertebrate mitochondrial DNA are not sequenced. Such occasions are based on the fact that nuclear sequences have typically read lengths less than 600-bp (Zhang & Hewitt, 1996). Therefore, the selected COI gene alone was considered for phylogenetic and population structure analysis of *P.canius* and *P.lineatus* in advance.

Preliminary evaluation of verified sequences was mainly generated 20 haplotypes in five populations of *P.canius* and 3 Haplotypes in one population of *P.lineatus* (Table.1). Based on the presented data of obtained haplotypes, it can be evidently seen that in *P.canius* samples,

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

KR086940 was found as the most common haplotype in the entire populations from Malaysia, however it was not found in the Negeri Sembilan and the Sabah populations. Moreover, KR086939 was detected as the second common haplotype in *P.canius* populations. The N.Sembilan and the Selangor populations had the most unique haplotypes, while the Johor population had just two shared haplotypes each. In other words, the N.Sembilan and the Selangor population possessed the highest number of identified haplotypes (n = 6), while the N.Sembilan and the Sabah populations had five and four haplotypes respectively.

Species	BOLDSYSTEM Index	GenBank Accession Number	Sampling Site
Plotosus lineatus	NUPM017-14	KP258659	**Selangor**
	NUPM016-14	KP258657	
	NUPM015-14	KP258658	
Plotosus canius	NUPM001-14	KP258648	**Negeri Sembilan**
	NUPM002-14	KP258651	
	NUPM006-14	KP258655	
	NUPM023-15	KR086935	
	NUPM003-14	KP258650	**Sabah**
	NUPM004-14	KP258649	
	NUPM005-14	KP258656	
	NUPM022-15	KR086936	
	NUPM007-14	KP258654	**Selangor**
	NUPM008-14	KP258653	
	NUPM009-14	KP258652	
	NUPM020-15	KR086938	
	NUPM021-15	KR086937	
	NUPM010-14	KP221601	**Sarawak**
	NUPM011-14	KP221602	
	NUPM012-14	KP221603	
	NUPM013-14	KP221604	
	NUPM014-14	KP221605	
	NUPM018-15	KR086940	**Selangor** **Sarawak** **Johor**
	NUPM019-15	KR086939	**Johor** **Negeri Sembilan**

Table. 1) Overview of haplotypes, their sampling sites and accession numbers.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

For the 655 available COI nucleotides, 509 sites (roughly 78 %) were detected as conserved sites, 146 (22%) as variable sites and 136 (21%) identified as parsim-informative sites. The average nucleotide composition in *P.canius* was 29% T, 27.6% C, 25.2%, A, and 18.3% G, while the average C+G content of selected positions was calculated as 45.9%, in *P.lineatus* though, calculation was 28, 28.6, 25, 18.4 and 47% respectively. Translation of all 23 haplotypes for conserved 655 bp fragment was produced 165 amino acids, which presented no signal of pseudogene in their structure. However sample size noticeably varied ranging from 13 to 30 in collected samples of *P.lineatus* and *P.canius* from five different districts in Malaysia, and regarding to the fact that original sample sizes were moderate and some sequences were failed, the actual sample collection is in the desired range recommended by Zhang *et al*. (2010). Basically, in order to detect approximately 80% genetic variation, a collection of 31.9 to 617.8 samples could be needed; although it is far from real laboratory and field work measures; hence it is suggested that at least 10 sample might be desirably sufficient to accurately identify the genetic variability in real phylogenetic studies (Zhang *et al*., 2010). An overview of most crucial outcomes in polymorphism analysis (Table. 2) indicated that the number of variable sites are moderately fluctuating from 2 in *P.lineatus* to 54 in *P.canius* samples from Sarawak, While the degree of nucleotide diversity was relatively low (0.00067-0.0391) and as its consequence, the level of polymorphism and genetic variation in populations reasonably presented small portion. Besides, the degree of haplotypes diversity waved from 0.395 (Sabah) to 0.771 (Sarawak).

The Tamura-Nei pairwise distance matrix (Figure. 2) indicated a comparatively high overall interspecies pairwise divergence of 25.2 %, while the least interspecific distance was 0.2 %. The Tamura-Nei intraspecific distance however, ranged from 0.2% to 9.7 % between *P.canius* from Sarawak and Selangor. Nevertheless, the majority of *P.canius* pairwise distances displayed low levels of conspecific divergence roughly around 1 %. The greatest genetic differences was observed between the Selangor and Sarawak (KR086937-KP221604) samples (9.7%), which is moderately reasonable due to their geographical distance. The next significantly high variances was detected between the Sabah-Sarawak pair (KP258656-KP221604) and Negeri Sembilan-Sarawak (KP258654-KP221604), although in later occasion, distance between Sarawak-Sabah sites is extraordinarily closer than Sarawak-Negeri Sembilan.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

Haplotype	*P.canius*					*P.lineatus*
GenBank Accession Number	Selangor *n*= 20	Negeri Sembilan *n*= 18	Johor *n*= 15	Sabah *n*= 30	Sarawak *n*= 22	Selangor *n*= 13
KP258659						0.078
KP258657						0.153
KP258658						0.769
KP258648		0.056				
KP258651		0.166				
KP258655		0.111				
KR086935		0.056				
KP258650				0.033		
KP258649				0.033		
KP258656				0.766		
KR086936				0.168		
KP258654	0.05					
KP258653	0.1					
KP258652	0.05					
KR086938	0.05					
KR086937	0.1					
KP221601					0.045	
KP221602					0.318	
KP221603					0.136	
KP221604					0.045	
KP221605					0.090	
KR086940	0.65		0.6		0.366	
KR086939		0.611	0.4			
Nucleotide Diversity (Pi JC)	0.00457	0.00184	0.00306	0.00134	0.0391	0.00067
Number of Haplotypes	6	5	2	4	6	3
Haplotype Diversity (Hd)	0.642	0.614	0.667	0.395	0.771	0.410
Number of Polymorphic site	14	5	3	4	57	2

Table. 2) Summary of 23 observed mitochondrial DNA haplotypes and their distribution, nucleotide diversity, number of haplotypes, haplotype diversity and number of polymorphic sites.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	KP221604																							
2	KP221603	0.002																						
3	KP221605	0.003	0.002																					
4	KP258655	0.091	0.089	0.087																				
5	KP258654	0.093	0.091	0.089	0.002																			
6	KR086940	0.091	0.089	0.087	0.089	0.002																		
7	KP258653	0.090	0.088	0.086	0.002	0.003	0.002																	
8	KR086937	0.097	0.094	0.092	0.005	0.006	0.005	0.006																
9	KP258652	0.096	0.094	0.092	0.006	0.008	0.006	0.008	0.005															
10	KR086938	0.089	0.087	0.084	0.002	0.003	0.002	0.003	0.006	0.006														
11	KP221601	0.091	0.089	0.087	0.003	0.005	0.003	0.005	0.008	0.009	0.005													
12	KR086936	0.089	0.087	0.084	0.005	0.006	0.005	0.006	0.009	0.011	0.006	0.008												
13	KP258650	0.090	0.088	0.086	0.006	0.008	0.006	0.008	0.011	0.012	0.008	0.009	0.002											
14	KP258649	0.091	0.089	0.087	0.006	0.008	0.006	0.008	0.011	0.012	0.008	0.009	0.002	0.003										
15	KP258656	0.093	0.091	0.089	0.005	0.006	0.005	0.006	0.009	0.011	0.006	0.008	0.003	0.005	0.005									
16	KP258648	0.087	0.084	0.082	0.006	0.008	0.006	0.008	0.011	0.012	0.008	0.009	0.005	0.006	0.006	0.005								
17	KR086935	0.087	0.084	0.082	0.006	0.008	0.006	0.008	0.011	0.012	0.008	0.009	0.005	0.006	0.006	0.005	0.000							
18	KP258651	0.089	0.087	0.084	0.005	0.006	0.005	0.006	0.009	0.011	0.006	0.008	0.003	0.005	0.005	0.003	0.002	0.002						
19	KR086939	0.089	0.087	0.084	0.005	0.006	0.005	0.006	0.009	0.011	0.006	0.008	0.003	0.005	0.005	0.003	0.002	0.002	0.000					
20	KP221602	0.078	0.076	0.078	0.028	0.029	0.028	0.029	0.032	0.032	0.026	0.028	0.026	0.027	0.027	0.026	0.024	0.024	0.022	0.022				
21	KP258657	0.241	0.239	0.235	0.238	0.240	0.238	0.240	0.246	0.246	0.235	0.245	0.240	0.242	0.237	0.240	0.237	0.237	0.241	0.241	0.252			
22	KP258658	0.237	0.235	0.231	0.235	0.237	0.235	0.237	0.243	0.243	0.232	0.241	0.236	0.239	0.234	0.236	0.234	0.234	0.237	0.237	0.248	0.002		
23	KP258659	0.237	0.235	0.231	0.235	0.237	0.235	0.237	0.243	0.243	0.232	0.241	0.236	0.239	0.234	0.236	0.234	0.234	0.237	0.237	0.248	0.003	0.002	

Figure. 2) Pairwise Tamura-Nei genetic distance in 23 employed haplotypes of *P.canius* and *P.lineatus*.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

Phylogenetic analysis using the 23 haplotypes of the genus *Plotosus* and one haplotype of *C.gariepinus* showed monophyletic status between *P.canius* and *P.lineatus* via the Neighbour-Joining method (Figure. 3). As stated, the three haplotypes from Sarawak population formed a basal clade for *P.canius* using Neighbour-Joining algorithm. Moreover, constructed topology precisely proved the pairwise genetic distances of shared haplotypes, highlighting that KR086940 (SGR, SWK, JHR) have the lowest distance to KP258654 (SGR) and highest to KP221604 (SWK), while KR086939 (JHR, NSN) have the greatest divergence from KP221604 (SWK) and smallest amount from KP258648 (NSN), as their subdivision clades illustrated.

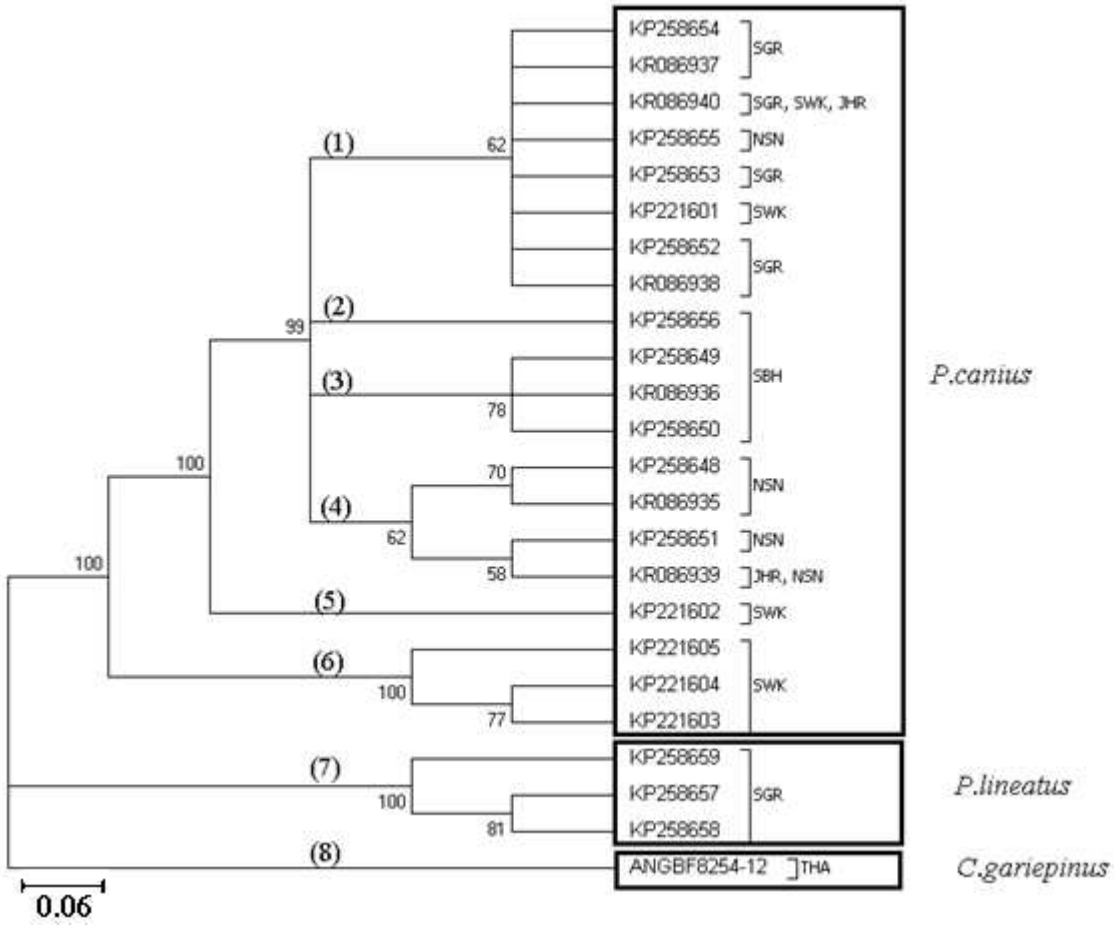

Figure. 3) Summary of Neighbour-Joining relationship in 24 employed sequences of *P.canius*, *P.lineatus* and *C.gariepinus* (clades have been indicated by bold numbers in round brackets).

Similarly, application of Maximum Likelihood algorithm (Figure. 4) have been subsequently confirmed the calculation of pairwise genetic distances among sequences of *P.canius* and *P.lineatus* with exception of some negligible changes in topology of clades and branches.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

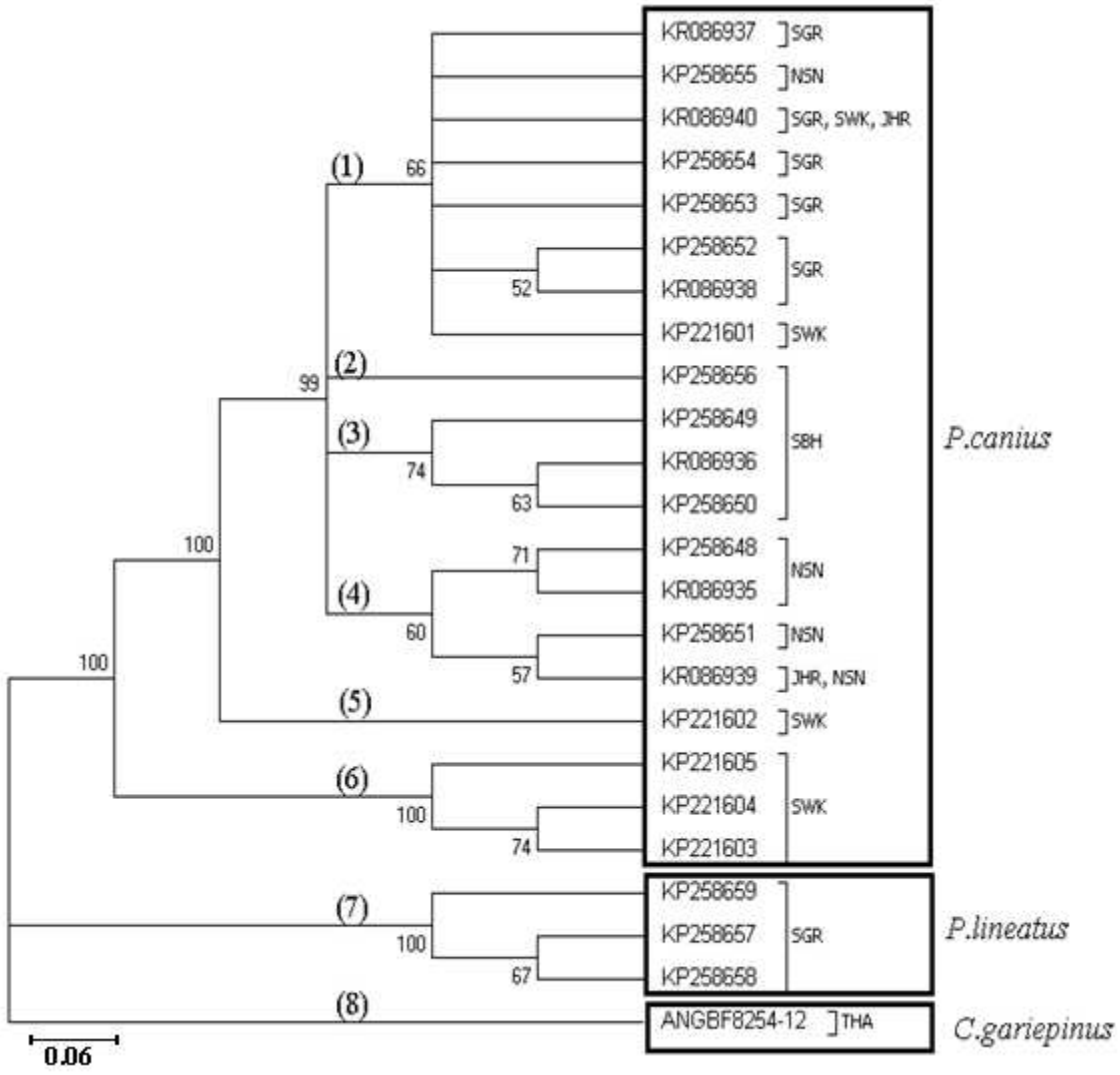

Figure. 4) Summary of maximum likelihood relationship in 24 employed sequences of *P.canius*, *P.lineatus* and *C.gariepinus* (clades have been indicated by bold numbers in round brackets).

The Minimum Spanning Network (MSN) of 20 haplotypes of *P.canius* (Figure. 5) in Malaysia presented more haplotype variability in Sarawak and Negeri Sembilan populations with six and five haplotypes respectively. Indeed, the Sarawak and Negeri Sembilan sequences illustrated a fairly high diversity, while the two haplotypes of Johor possessed the lowest variability. However, the phylogram revealed two relatively irrefutable geographical clades (Sarawak and Negeri Sembilan), occurrence of mix haplotypes with other clades indicating that no accurate geographical genetic structure have been certainly detected in studied populations of *P.canius*.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

Analysis of potential geographical clades have been similarly suggested that all populations are moderately mixed except in Sarawak and Negeri Sembilan. Although there are some possible clades, it was not precisely feasible to clustering the population based on their geographical divisions due to existence of exclusively one connecting mutational steps for most sequences. Hence, analysis was not capable of showing precise separation of geographical clades.

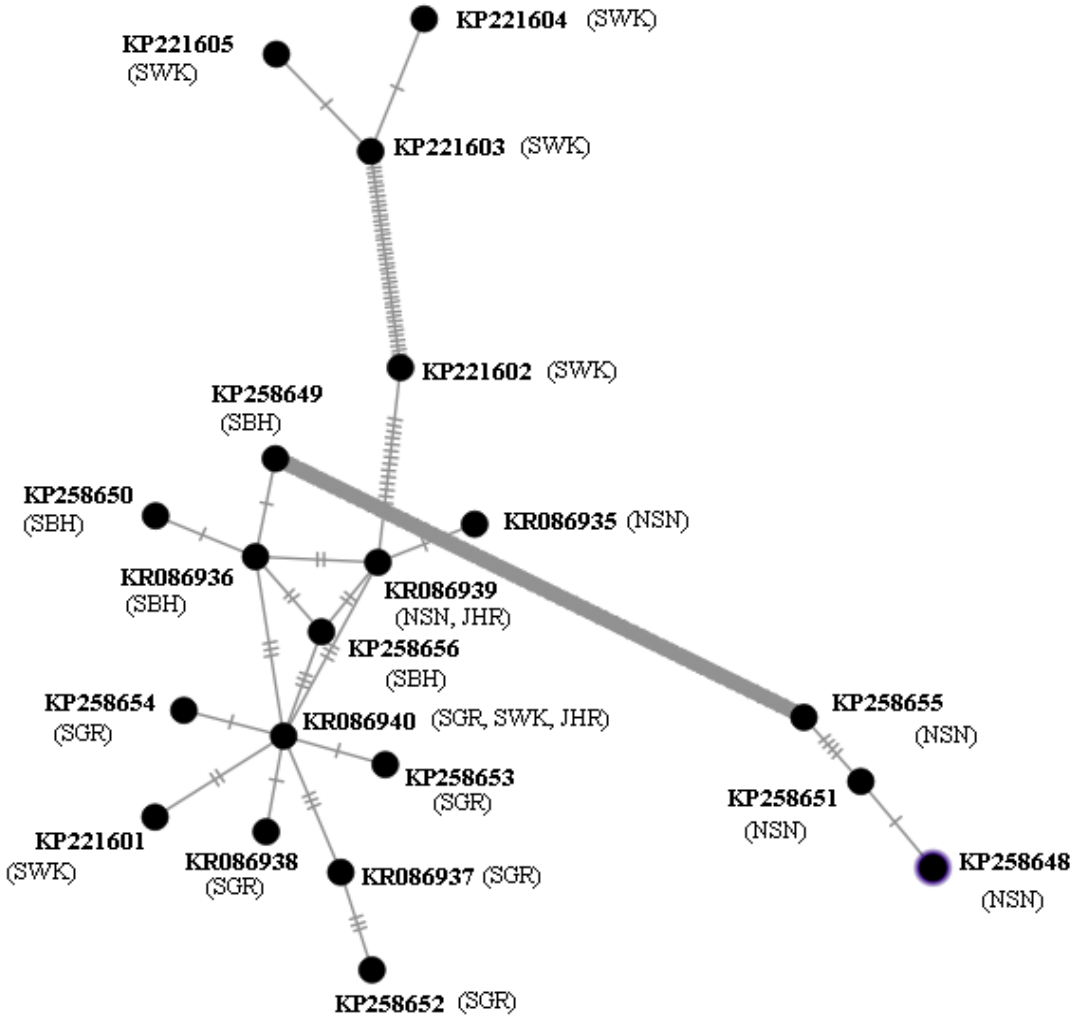

Figure. 5) Minimum Spanning Network (MSN) of 20 haplotypes of *P.canius*.

Analysis of population differentiation inferred from pairwise distance of F_{ST} and Chi-square among studied populations of *P.canius* is displayed in Table. 3. Significant genetic variations were detected in all assessments within *P.canius* sequences and two considered species of eel-tail catfishes ($p < 0.001$). However, there were significant distance diversity in genetic variations of

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

almost entire evaluations among populations of Malaysia especially between the Sabah and the Negeri Sembilan populations. As expected, the most diversity were identified between *P.canius* populations from the Negeri Sembilan and the Sabah with the rate of 0.62504, which basically means they are nearly genetically divided due to their distance and subsequent decrease in gene flow. However, the lowest genetic distance was detected between Selangor cluster and the Johor clade by the F_{ST} values of 0.05417. Hence, it was considered that maximum sharing of genetic material occurred between Johor and Selangor Populations, while the minimum genetic similarity identified among Sabah and Negeri Sembilan sequences.

	***P.canius* SBH**	***P.canius* JHR**	***P.canius* SGR**	***P.canius* NSN**	***P.canius* SWK**
***P.canius* SBH**	0.00000				
***P.canius* JHR**	0.60156	0.00000			
***P.canius* SGR**	0.43390	0.05417	0.00000		
***P.canius* NSN**	0.62504	0.44097	0.09806	0.00000	
***P.canius* SWK**	0.41533	0.27777	0.29359	0.31333	0.00000

Table. 3) Population pairwise (F_{ST}) values of chi square test for population variation originated with 1000 permutations.

Hierarchical statistics of AMOVA (Table. 4) was clearly suggested that roughly 36 % of experimental deviations were inter-population variations, while within population variations were merely responsible for approximately 64 % of overall differentiation.

Source of Variation	**Degree of Freedom**	**Sum of Squares**	**Variance Components**	**Percentage of Variation**
Among Populations	4	77.873	3.09472	35.55
Between Populations	18	100.969	5.60941	64.45

Table. 4) Hierarchical analysis of molecular variance (AMOVA) in *P.canius*.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

Population Genetic results inferred from Genotyping analysis

Fragment analysis were estimated DNA band sizes in *P.canius* that are illustrated alongside with size range in original species (*Tandanus tandanus*), the sequence of each primers and associate annealing temperatures in Table. 5.

Primer	Sequence	Size *T.tandanus*	Size *P.canius*	Annealing Temperature
Tan 1-2	F: 5´CCGACTGTCAGTGAAAAGGAG3´ R: 5´AGGGTCTGGGAGTGAATGAG3´	216-244	349-385	55°C
Tan 1-7	F: 5´TGTATGGAGCTACTAACAAAACAGG3´ R: 5´TACTCCAGCCCTGAAGGTG3´	181-227	114-125	55°C
Tan 1-10	F: 5´CCTGATTTCTCTCCCAAGG3´ R: 5´AGAAAGGTGGTGCATGTGTG3´	298-310	91-97	55°C
Tan 3-27	F: 5´TGTGGAAGGTTGGGGTTATG3´ R: 5´CGTGATCATGCAAACAGATG3´	215-269	167-168	55°C
Tan 3-28	F: 5´CCCCATTTGCTTTTTCTCTG3´ R: 5´TGTTGAAAGCGGCATGTTAG3´	289-301	280-299	55°C

Table. 5) Five engaged primer sets and their associated size and temperature in *T.tandanus* and *P.canius*.

Genotyping results did not found any signal of null allele nor large allele failure, hence nor scoring inaccuracies due to stuttering. All the five microsatellites loci showed accurate and successful amplification in all populations, although heterozygous alleles were not found in all populations. One microsatellite loci (Tan3-27), was evidently monomorphic in all populations while Tan 1-2, Tan 1-10, Tan 1-7 and Tan 3-28 were polymorphic in at least one population. After implementing the sequential Bonferroni adjustment (Rice, 1989), only four out of the 50 (8%) loci pairs were significant for linkage disequilibrium ($P < 0.05$). Thus, all the five microsatellites loci were considered useful for genetic applications based on the absence of consistent linkage disequilibrium in locus pairs among the studied populations.

Furthermore, after Bonferroni adjustment, nine out of the 25 (36%) microsatellite loci still showed deviation from Hardy-Weinberg Equilibrium (HWE), which might be owing to heterozygote deficiency effects (Table. 6). Heterozygote deficiency could be caused by population structuring, null alleles or inbreeding (Brook *et al*., 2002).

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

Locus	N.Sembilan	Sabah	Selangor	Sarawak	Johor	Total* Mean
N	30	30	20	22	15	117
Tan 1-2						
N_a	1	2	1	1	2	4
Ar	1.000	2.000	1.000	1.000	2.000	3.714
H_o	0.0000	0.0000	0.0000	0.0000	0.6000	0.0769
H_e	0.0000	0.4994	0.0000	0.0000	0.4345	0.5974
Fis	-	0.000	-	-	-0.400	
HW	-	1.0000	-	-	0.1791	
Tan 1-7						
N_a	2	2	3	2	1	6
Ar	2.000	2.000	3.000	2.000	1.000	5.844
H_o	1.000	0.0000	0.6000	0.0000	0.0000	0.3590
H_e	0.5085	0.4994	0.6769	0.4947	0.0000	0.7727
Fis	-1.000	1.000	0.116	1.000	-	
HW	0.0000	1.0000	0.5856	1.0000	-	
Tan 1-10						
N_a	1	1	1	1	2	3
Ar	1.000	1.000	1.000	1.000	2.000	2.564
H_o	0.0000	0.0000	0.0000	0.0000	0.4000	0.0513
H_e	0.0000	0.0000	0.0000	0.0000	0.3310	0.3472
Fis	-	-	-	-	-0.217	
HW	-	-	-	-	0.5395	
Tan 3-27						
N_a	1	1	1	1	1	2
Ar	1.000	1.000	1.000	1.000	1.000	1.988
H_o	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
H_e	0.0000	0.0000	0.0000	0.0000	0.0000	0.2242
Fis	-	-	-	-	-	
HW	-	-	-	-	-	
Tan 3-28						
N_a	3	2	3	2	1	8
Ar	3.000	2.000	3.000	2.000	1.000	7.592
H_o	0.4667	1.0000	0.3000	0.0000	0.0000	0.4274
H_e	0.6169	0.5885	0.4769	0.4947	0.0000	0.8393
Fis	0.247	-1.000	0.377	1.000	-	
HW	0.5862	0.0000	0.5771	1.0000	-	

Table. 6) Genetic variation at 5 microsatellite as of five populations of *P.canius* in Malaysia: sample size (*N*), Number of alleles (N_a), allele richness (*Ar*), observed heterozygosity (H_o), expected heterozygosity (H_e), inbreeding coefficient (*Fis*) ($p < 0.05$ symbolic accustomed nominal level (5%) 0.000042, and Hardy-Weinberg expectation (disequilibrium) (HW).

Data from Table. 6 showed that six out of nine deviations were related to Tan1-7 and Tan 3-28 among all the five *P.canius* populations. The fact that the two loci did not show any signal of deviation from HWE in three populations may imply the probability that the estimated deviations could have been originated from either the occurrence of uncertain structure or inbreeding among these three population divisions (Pritchard *et al*., 2000). Fairly high level of consistency

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

was detected toward inclusion or exclusion of Tan1-7 and Tan 3-28, accordingly these loci have been retained for further analysis. *Fis* ($P < 0.05$) estimations have been considerably diverse from zero, except in locus Tan 1-2 from Sabah. This alongside with substantial departure from HWE indicates the damaging effect of heterozygote deficiency within associated populations. However, the positive calculated estimations could be translated as decrease in heterozygous levels among offspring in population, mostly owning to non-random mating and its subsequent inbreeding. On the other hand, negative *Fis* estimates might be indication of increasing in heterozygosity level, which could usually occur as a result of random mating system, hence genes should be probably more different (Pritchard *et al*., 2000).

Analysis of population genetic inferred from molecular coancestry information (Table. 7) revealed that Polymorphism Information Content (PIC) of the applied microsatellites varied from 19.86 to 73.99. However, loci with numerous allele numbers and a PIC value of 1 are considered as highly polymorphic and thus most desirable, lowest rates are also slightly informative (if and only if PIC > 0) (Botstein *et al*., 1980). Besides, the arithmetic values of heterozygosity (0.2235-0.8357) confirmed the consistent application of all the 5 microsatellite loci in population genetic study of *P.canius* in Malaysia.

Microsatellite	N_a	Heterozigosity	PIC (%)	Effective Allele NO.
Tan 1-2	4	0.5949	54.12	2.47
Tan 1-7	6	0.7694	73.99	4.34
Tan 1-10	3	0.3457	30.11	1.53
Tan 3-27	2	0.2235	19.86	1.29
Tan 3-28	8	0.8357	81.71	6.09

*Heterozygosity was estimated as arithmetic mean of expected and observed heterozygosity.

Table. 7) Analysis of population genetic using molecular coancestry information: Number of alleles (N_a) and Polymorphism Information Content (PIC).

Hierarchical results of microsatellites showed that approximately 64 per cent of experimental variations were originated from inter – population variations, while within individuals variations were only accountable for roughly 28.5 per cent of overall differentiation (Table. 8).

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

Source of variation	Sum of squares	Variance components	Variation %
Among Populations	191.456	1.02366	63.77361
Among individuals within populations	79.039	0.12422	7.73898
Within Individuals	53.500	0.45726	28.48741

Table. 8) Hierarchical analysis of molecular variance (AMOVA) in *P.canius*.

F_{ST} plot estimations of all involved microsatellites noticeably illustrated that all pairwise calculations presented a fairly high differentiations among populations ranging from 0.29711 between populations of Selangor and Negeri Sembilan to 0.80500 between populations of Sarawak and Johor (Table. 9). The current microsatellite experiment showed that the highest genetic differentiation was between the Johor population and the other populations from Peninsular Malaysia and Borneo. Indeed, the Johor population showed strong deviation from other collected populations, while displayed relatively low intra-population genetic variation (Table. 9), suggesting that the Johor population was less connected to the others during a sizeable period of its evolutionary phase. Likewise, the Sarawak population in the Southwest region of the South China Sea had high F_{ST} values between other populations, signifying the restricted gene flow between the Sarawak population and the other populations. The differentiation level among populations of Negeri Sembilan and Selangor (F_{ST} = 0.29711, P< 0.05) showed relatively lower values, even in comparison with their neighbouring populations, theoretically demonstrating the populations that endured inbreeding or genetic drift since their isolation from other populations. Similarly, the Negeri Sembilan and Sabah (F_{ST} = 0.39244, P< 0.05) populations also showed small but significant variances in relation to their close neighbour populations.

		N.Sembilan	Sabah	Selangor	Sarawak	Johor
1	**N.Sembilan**	0.00000				
2	**Sabah**	0.39244	0.00000			
3	**Selangor**	0.29711	0.41039	0.00000		
4	**Sarawak**	0.71934	0.68086	0.72437	0.00000	
5	**Johor**	0.73211	0.68561	0.74834	0.80500	0.00000

Table. 9) Pairwise F_{ST} estimations through *P.canius* populations generated from five microsatellites loci and inclusion of five populations. All calculations were fairly significant ($P < 0.05$) using 10000 permutations.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

Levels of genetic variations seem to be widely fluctuated between *P.canius* populations owning to H_e and *Ar* oscillation on which H_e extending from 0.0000 to 0.6769 (Selangor) and *Ar* varied from 1 to 3 (Negeri Sembilan and Selangor). Obviously, two significant clusters could be seen among established populations of *P.canius* in this study, one cluster with low allelic richness and H_e estimations (Johor samples using locus Tan 1-2 and locus Tan 1-10), and another cluster with relatively acceptable H_e and allelic richness (the other four populations using locus Tan 3-28 and locus Tan 1-7). However, the genetic variation in the current study was highly reliant on microsatellites and their sequences as the engaged loci did not specifically develop for *P.canius.* Moreover, allelic frequencies among virtually each combination of population pairs showed highly significant differentiation (F_{ST} < 0.05) (Table. 9), implying that gene flow might be highly restricted among studied populations.

Sampled populations of *P.canius* were basically distributed into five minor clusters using Bayesian analysis. Consequently, the initial highest membership value (q) of the studied populations including Negeri Sembilan, Sabah, Selangor, Sarawak, Johor was estimated as 0.941, 0.983, 0.968, 0.988, and 0.986 respectively (Table. 10). The application of STRUCTURE program subsequently illustrated 4 major *K* (isolated clusters) (Figure. 6). Regarding the fact that assessing the expected value of *K* might not be straightforward (Evanno *et al*., 2005), Bayesian structure analysis of the current study revealed the highest probability of *K* for *P.canius* in *K*=2. The four estimated clusters were included Cluster 1: Johor, Cluster 2: N.Sembilan and Selangor, Cluster 3: Sabah, and Cluster 4: Sarawak (Figure. 7).

Populations	Cluster Membership				
	1	2	3	4	5
N.Sembilan	0.005	0.041	0.941	0.010	0.004
Sabah	0.005	0.004	0.005	0.983	0.003
Selangor	0.005	0.968	0.019	0.005	0.004
Sarawak	0.003	0.003	0.003	0.003	0.988
Johor	0.985	0.004	0.004	0.004	0.003

Table. 10) Membership ratio estimated for each population of *P.canius.*

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

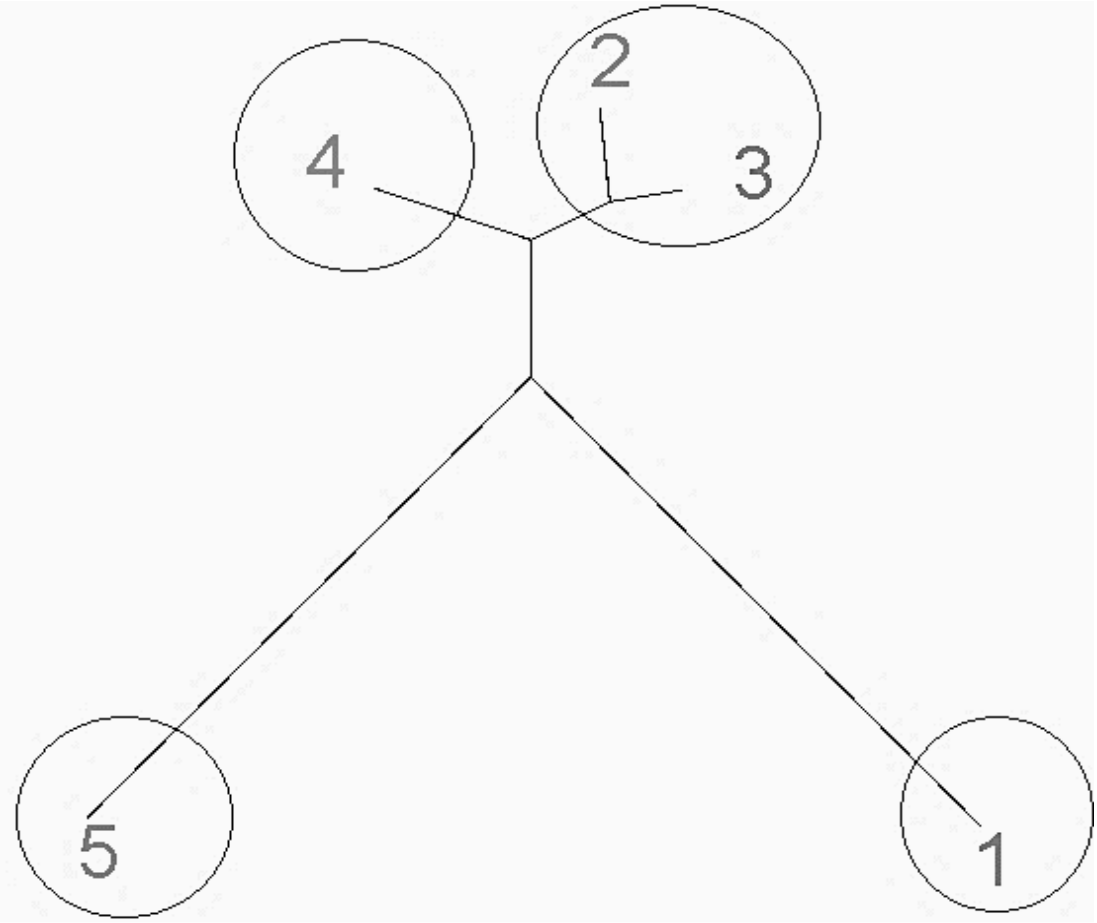

Figure. 6) Tree plot scheme of five engaged populations of *P.canius*.

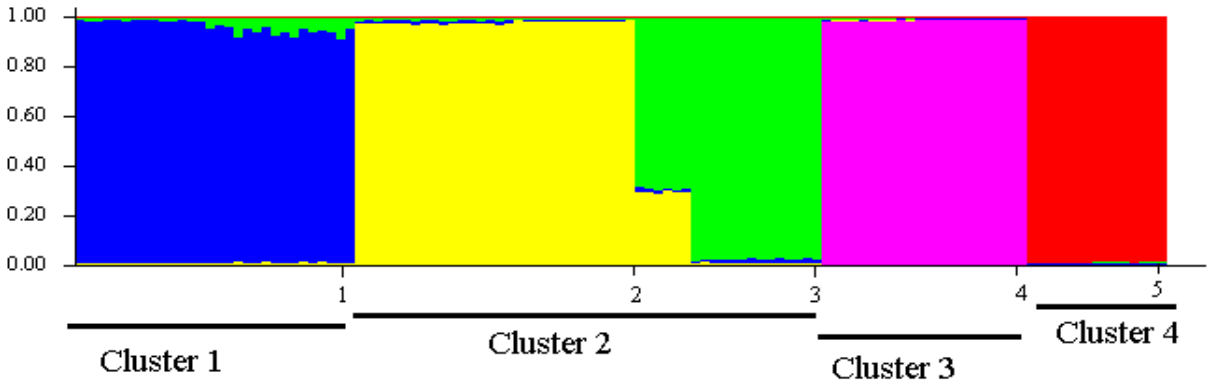

Figure. 7) Population structure of five *P.canius* populations in Malaysia.

Population assignment outcomes evidently revealed that almost all individuals were assigned to their original populations with the probability rate of P> 0.05 (Table. 11). However, the estimation of individual assignment to their populations with the same probabilities revealed the

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

closer rates in comparison levels to other sampling sites. For instance, the Negeri Sembilan population had a relatively closer assignment ratio to the Selangor population.

Assigned Population	CA%	Original Location				
		N.Sembilan (N=30)	Sabah (N=30)	Selangor (N=20)	Sarawak (N=22)	Johor (N=15)
N.Sembilan	100	4.219	70.330	35.873	117.792	92.839
Sabah	100	70.143	4.338	58.286	132.093	96.994
Selangor	100	35.937	58.537	4.715	101.049	83.156
Sarawak	100	115.433	129.921	98.626	3.133	101.070
Johor	100	88.554	92.895	78.807	99.143	2.761

Table. 11) Population assignment based upon five microsatellite loci frequencies in *P.canius* (CA: correct assignment).

Analysis of population bottleneck did not identified substantiating signals of recent population drop in all populations studied using the two phase model (T.P.M) estimations (Table. 12). Furthermore, calculation of the infinite-allele model (I.A.M) comprehensively implied that there was no bottleneck evidence among the studied populations, while the parallel statement was assumed using the stepwise mutation model (S.M.M). Moreover, the shift-mode estimation of allele frequencies was perceived in all five populations, while altogether none of the mutation models were broadly illustrated consistent signals of bottleneck in engaged populations. Therefore, the current experiment could not detect any signals of bottleneck within the *P.canius* populations based on three applied models. Nevertheless, outcomes of mutational models consistently suggested the extension in populations due to absence of genetic drift and /or an invasive allele originating from different populations (Piry *et al*., 1999).

	I.A.M	T.P.M			S.M.M	Mode Shift
		60	70	80		
N.Sembilan	0.22672	0.24435	0.26030	0.26596	0.28595	Y
Sabah	0.07656	0.09247	0.09295	0.09683	0.10790	Y
Selangor	0.28119	0.31416	0.34221	0.32656	0.64363	Y
Sarawak	0.17976	0.21249	0.21558	0.20907	0.24057	Y
Johor	0.18105	0.21600	0.24444	0.26510	0.22513	Y

Table. 12) *P* values originated from bottleneck analysis within five populations of *P.canius* (I.A.M: infinite allele model, T.P.M: two phase model, S.M.M: stepwise mutational model, estimation indicate the mutation in stepwise mutational model: Y: yes, N: no) (*Significant values $P < 0.05$).

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

DISCUSSION

Phylogenetic Analysis of COI Gene

This study has confirmed the efficiency of the COI barcode in identification of eel-tail catfish species. Barcoding fragment has been effectively sequenced mitochondrial DNA isolated from two species of family Plotosidae. Current study provided the first sequence database of *P.canius* to be submitted into barcoding data set. The first and the most common outcome of the undergone experiment could be the involvement of two common haplotype. While the most common sequence was KR086940 between populations of Selangor (*n= 7*), Sarawak (*n= 6*), and Johor (*n= 9*), the KR086939 identified as second common haplotype sharing between Johor (*n= 6*) and Negeri Sembilan (*n= 4*). Nevertheless, the most significant finding should be the occurrence of shared haplotype between Selangor, Johor (Peninsular Malaysia) populations and the Sarawak (Borneo) population. The haplotype sharing and their consequent gene flow could practically happen due to several reasons such as breeding migration, mutation, pelagic larvae, and sharing of common ancestors (Frankham, 1996).

Migration is a common behaviour in nearly 3% of all extant fish species (Binder *et al*., 2011). However, there is practically no record on migration and migratory behaviour of family Plotosidae, thus the first assumption of dispersal via migration and ocean current might be highly unlikely since majority of the catfish species cannot endure long distance swimming (more than 500 kilometers) due to their body shape and structure (Jónsson, 1982). In addition, marine dispersal of eggs, larvae and even juveniles of *P.canius* between two separate ocean currents comprising Straits of Malacca (Selangor and Johor) and South China Sea currents (Sarawak) might also be questionable. In the Strait of Melaka, circulation of currents (particularly in surface) are due to effects of tidal patterns and wind, while the route of both surface and deep layer currents are shown to be relatively the same and toward northwest (Rizal *et al*., 2010). In Johor, however, currents are highly depend on strong winds during monsoon seasons. Indeed, if the monsoon is in its northeast route, the current streams toward South along the coastal region of Malaysia, otherwise, the current will be northward (Mohd Fadzil & Chuen, 2011). Finally, pattern of ocean currents in western South China Sea are largely influenced by season. The circulation in South segment of western current tends to be stable and northward where after

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

separation from coastal region, it forms a northeastern pattern in summer. In fall, however, it strongly flows toward southwest (Fang *et al*., 2012). Therefore, general patterns of aquatic circulation in Strait of Melaka, Johor maritime territory and western area of South China Sea might not strongly implies the probable distribution of grey eel-tail catfish eggs, larvae and juveniles and its consequent gene flow and genetic connectivity.

Considering all possible expectation on genetic variability and gene flow of *P.canius* in Peninsular Malaysia, the second hypothesis of sharing common ancestor due to historical geographic events may reflect the most plausible explanation. Southeast Asia is believed to endured simultaneous glaciation and consequent deglaciation along with its associated decreasing and increasing of marine levels during the Pleistocene period which greatly influenced continental and oceanic configuration (Voris, 2000). During such variations, some regions might be preserved their stable environmental conditions that is nowadays called a refugium on which can greatly affect the gene flow and genetic variability particularly in endemic species (Hobbs *et al*., 2013). Moreover, geographical information proposed that the Pacific and Indian ocean were initially connected directly before the formation of Sundaland (nowadays submerged forming shallow ocean of most Southeast Asia with less than 100m depth) during the Triassic up to the Pleistocene period (Esa *et al*., 2008), hence made such gene flow possible between these comparatively distant locations.

Hardy-Weinberg Equilibrium and Genetic Diversity

Overall allelic richness revealed quite lower rates using the cross amplified primers ranging from 2 – 8 among the sampled populations in comparison with original species (Rourke *et al*., 2010). Tan 3-28 demonstrated the highest overall allelic rate fluctuating from 1-3 among five populations of *P.canius,* while the lowest level was detected in Tan 3-27. Moreover, the Selangor population showed maximum number of alleles (9), whereas the Johor and Sarawak populations exhibited the lowest (6). Similar instance of low allelic variation have been described in *Bolbometopon muricatum* (Priest *et al*., 2014), *Schizothorax biddulphi* (Palti *et al*., 2012) and *Prosopium cylindraceum* (Mccracken *et al*., 2014). A possible reason for the occurring of low levels of allelic richness might be due to the small employed population size. Hale *et al*. (2012)

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

pointed out the positive effects of sampling size between 25 to 30 individuals per population, however they also mentioned the necessity of 5-100 samples per collection to avoid rare uninformative alleles. Marine vertebrates are believed to present greater allele difference at their microsatellite primers comparing to freshwater populations, which is mostly consistent with their higher population evolutionary size (Neff & Gross, 2001). Their research later revealed that the difference in microsatellite polymorphism among classes and species could be highly dependent upon changes in life history and population biology and moderately to differences happening to microsatellite functions during natural selection. Therefore, it would be probably reasonable that fewer allele number of *P.canius* might be owning to variation in its biology and historic traits, however, the correlation of allelic richness and sample size should not be overlooked.

The average value of observed heterozygosity (H_O) estimated in the five tested microsatellites in *P.canius* were as low as 0.2168, which showed high difference levels in comparison to standard heterozygosity in marine populations (H_O= 0.79) and anadromous fish species (H_O= 0.68) (DeWoody & Avise, 2000). In fact, considerable heterozygote deficiencies were observed in the engaged populations with the exception of the Tan 1-7 and Tan 3-28 loci. Similar temporal pattern of low genetic diversity have been reported for *Pleuronectes platessa* in Island (Hoarau *et al*., 2005) and *Clarias macrocephalous*(Na·Nakorn *et al*., 1999), while in most catfish species higher levels of heterozygosity have been documented as in *Mystus nemurus* (H_O= 0.44-0.57) (Usmani *et al*. 2003). Several decisive issues might influence the genetic variability of marine species through the variation of Hardy-Weinberg including migration, genetic drift, sample size, over-exploitation, effective size of population and patterns of mating (DeWoody & Avise, 2000). Certainly, *P.canius* should not be presumably considered as long distance migratory marine fish species due to its body structure (Riede, 2004). Alternatively, a possibility of genetic drift in the current study is also suspicious as it basically happens only in small effective size populations that experiencing a period of bottleneck (DeWoody & Avise 2000) at which is completely invalidated in marine species studies like current research.

Small sample size of collected populations might also be measured as a major parameter in detection of low heterozygosity variation because of the failure to accurately detect the entire extant alleles of the selected populations, hence, deficiency in heterozygote identification

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

(Na·Nakorn *et al*., 1999). Indeed, the current collection size for *P.canius* used for population genetic analysis purposes should be quite small based on Kalinowski (2005), therefore, the hypothesis of deficiency in heterozygote detection due to the low level of sampled specimen could be accepted. The last cause of a low heterozygosity levels and its consequent genetic variation is non-random system of mating behaviour among populations (Brook *et al*., 2002; Balloux *et al*., 2004). Estimation of HWD for the current study however, showed considerable deviation for approximately 36% of the primer/population pairs, which might be due to heterozygote deficiency effects. However, Balloux *et al*. (2004) highlighted that the positive correlation of inbreeding and heterozygosity needs to be examined through application of more polymorphic markers on which demonstrates greater proportion of linkage disequilibrium. Alternatively, the correlation of *Fis* values and inbreeding have been practically assessed and documented in many studies (Balloux *et al*., 2004; Abdul-Muneer, 2005; O'Leary *et al*., 2013). The positive calculated estimations could be translated as a decrease in heterozygous levels among offspring in a population, mostly due to absence of random mating and its subsequent inbreeding. The current study showed considerable significance levels ($P < 0.05$) of *Fis* estimations. This alongside with substantial departure from HWE would indicate the damaging effect of heterozygosity deficit within the populations.

Analysis of Population Structure

A remarkably high levels of genetic structure were detected among populations of *P. canius* ranging from 0.05417 to 0.62504, showing significantly high structuring among studied populations except differences between Johor – Selangor samples (F_{ST} = 0.05417) and Selangor – Negeri Sembilan (F_{ST} = 0.09806). Moreover, AMOVA statistics evidently revealed that approximately 64 % of genetic variations were due to within population variations. Hence, the fairly high F_{ST} rates, significant hierarchical molecular results and consequent higher genetic variances among *P.canius* populations in Peninsular Malaysia and their relatives in Borneo, in addition to the detection of only one sharing haplotype (KR086940), would suggest the absence of contemporary gene flow among them most probably due to the geological modification, consequential rise in marine water levels and historical continental separation during the

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

589 Pleistocene era (Esa *et al*., 2008; Song, 2012). However, exceptional cases between Selangor –
590 Negeri Sembilan and Selangor – Johor might be inversely interpreted as occurrence of gene flow
591 or migration regarding to fairly close distances rather than extraordinary distance between
592 Borneo and Peninsular Malaysia. The sequential genetic diversity presented in this study could
593 be caused by high haplotype frequencies among the five populations of *P.canius* in Malaysia, in
594 addition to identification of unique sequences in each population (except in Johor). The present
595 patterns of differentiation among catchments is believed to be significantly as a consequence of
596 the Pleistocene associated historical and geological continental and sea level distraction and its
597 subsequent isolation of lands and populations (Esa *et al*., 2008).
598
599 The calculated F_{ST} values of five microsatellites in *P.canius* showed significant estimation,
600 indicating substantial genetic structure and differentiation among the studied populations. All
601 populations also showed significantly high assignment rates, followed by a low membership
602 recorded for other population clusters. High rates of proper assignment might indicates strong
603 population differentiation among the studied populations (Paetkau *et al*., 2004). Although the
604 Sabah population demonstrated a close pairwise distance with the Selangor and Negeri Sembilan
605 populations, the Negeri Sembilan and Selangor populations showed the lowest differentiation
606 level (0.29711), and also the highest cluster membership in comparison with other populations.
607 Surprisingly, the highest level of pairwise *Fst* differentiation has been estimated between the
608 Johor population and the other four populations, in contrast to the closer geographical distance
609 between the Johor and the Negeri Sembilan populations. Indeed, microsatellite analysis made a
610 relatively counter-outcome in comparison with mitochondrial results, where F_{ST} estimation of
611 former populations was estimated as the lowest among the *P.canius* samples. Discrepancies
612 between genetic differentiation detection through microsatellite loci and mitochondrial DNA is
613 believed to be related to three factors: (1) high sensibility of mitochondrial COI gene in detection
614 of variation (Shaw *et al*. 2004), (2) weaker nuclear-based subpopulation detection (Cano *et al*.
615 2008) and (3) technical complications of microsatellite like homoplasy (Estoup *et al*., 2002).
616
617 One of the most common practical problems, which is believed to be mostly associated with
618 microsatellite primers (due to higher mutation rate) is well-known as homoplasy (Balloux &
619 Lugon-Moulin, 2002). Homoplasy might diminish the microsatellite-based population

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

differentiation signals. The existence of homoplasy is highly dependent on the occurrence of different identical locus copies, while such identical character is not consequent of mutual ancestor. In fact, the occurrence of homoplasy might be correlated with combination effects of high rates of mutation, and outsized population together with strong restriction in allele size (Estoup *et al*., 2002). However, the effective number of alleles on which presented in Table. 6 and Table. 7 showed a low level of allele size frequency, the current population size of *P.canius* used for population genetic analysis is ostensibly quite small based on Kalinowski (2005) rather than being oversized. Hence, the later effective cause of homoplasy is somehow nullified in this study. Furthermore, several microsatellite based studies have pointed out the significance of Stepwise Mutation Model (S.M.M) on possibility of homoplasy in different taxa (Angers & Bernatchez, 1997; Culver *et al*., 2001; Estoup *et al*., 2002; Anmarkrud *et al*., 2008), which was invalidated by provided statistics on bottleneck analysis of recent study in Table.11.

O'Reilly *et al*. (2004) later pointed out that implications of homoplasy in identification of population structure using microsatellite loci are supposedly common in marine species. Nevertheless, further researches have been implied that implications of genetic drift and gene migration might have considerably greater effects on population differentiation analysis in comparison with homoplasy (Estoup *et al*., 2002). Basically, marine vertebrates supposed to have the higher population effective size (N_e) comparing to freshwater species (Hauser & Carvalho, 2008). Besides, genetic drift and effective size are believed to be greatly correlated, hence it is highly probable that neighbouring geographical populations demonstrate the imperceptible population differentiation and structures especially using neutral primers like microsatellites (Larmuseau *et al*., 2010).

CONCLUSION

The current genetic characterization of *P.canius* provided some basic results on their phylogeny and population structure. The phylogenetic and phylogeographic analysis of *P.canius* noticeably constructed accurate clusters in the five population of Malaysia, demonstrating the capability of the chosen mitochondrial COI barcoding gene to potentially assign the genus *Plotosus* into their biological taxa. Indeed, COI analysis resolved the phylogenetic relationships between *P.lineatus*

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

and *P.canius* populations, supporting their taxonomic status as different species. A low mitochondrial differentiation was found among *P.canius* populations with some indication of endemism (haplotype restricted only to a particular population) as observed in the Sabah population. Nevertheless, the COI gene revealed sufficient informative interpretation of relationships among the five populations, supporting by reasonable bootstrapping values (>85%). The sharing of haplotypes between a few samples from Peninsular Malaysia and their Sarawak counterpart of Borneo showed the sensitivity of the COI marker to infer the biogeographical history of *P. canius* and potentially other marine taxa in the region.

Microsatellites analyses on the population structure of *P.canius* showed slightly different patterns of structuring in comparison with the COI findings. Nevertheless, both markers detected higher level of among population differentiations than within population variability. In addition, four main clusters or genetic stocks of *P.canius* were identified using the cross species amplification study of *T. tandanus* microsatellites.

Finally, the results from this study has provided valuable understandings on the genetic characterization, molecular phylogeny, evolutionary kinship, and population structuring of *P.canius*, in particular, and the genus *Plotosus*, in general. However, further studies must be conducted using more geographical and sampling sites, larger population sizes per site, and more documented sequences from applicable mtDNA fragments. Furthermore, designing species specific hypervariable nuclear markers such as microsatellite for *P.canius* must be considered in order to accurately analyze the population structure and genetic diversity of *P.canius* before implementation of advanced fisheries and conservation strategies.

REFERENCES

Abdul-Muneer, P.M. (2005). Molecular genetic characterization of endemic yellow catfish, *Horabagrus brachysoma* (Gunther). Cochin University of Science and Technology. 238 pp.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

682 Angers, B., Bernatchez, L. (1997). Complex evolution of a salmonid microsatellite locus and its
683 consequences in inferring allelic divergence from size information. *Molecular Biology and*
684 *Evolution* **14**(3), 230 – 238.

685 Anmarkrud, J.A., Kleven, O., Bachmann, L., Lifjeld, J.T. (2008).Microsatellite evolution: Mutations,
686 sequence variation, and homoplasy in the hypervariable avian microsatellite locus HrU10. *BMC*
687 *Evolutionary Biology* **8**(1), 1-10.

688 Appeltans, W., Ahyong, S.T, Anderson, G., Angel, M.V., Artois, T., Bailly, N., Costello, M. J. (2012).
689 The magnitude of global marine species diversity. *Current Biology* **22**(23), 2189–2202.

690 Balloux, F., Amos, W., Coulson, T. (2004). Does heterozygosity estimate inbreeding in real populations?
691 *Molecular Ecology* **13**(10), 3021–3031.

692 Balloux, F., Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite
693 markers. *Molecular ecology* **11**(2), 155–65.

694 Bandelt, H., Forster, P., Röhl, A. (1994). Median-Joining Networks for Inferring Intraspecific
695 Phylogenies. *Molecuar Biology and Evolution* **16**(1), 37–48.

696 Binder, T.R., Cooke, S.J., Hinch, S.G. (2011). The Biology of Fish Migration.In "Encyclopedia of Fish
697 Physiology: From Genome to Environment", Academic Press, San Diego. pp 1921–1927.

698 Botstein, D., White, R.L., Skolnick, M., Davis, R.W. (1980). Construction of a genetic linkage map in
699 man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**(3),
700 317-341.

701 Bourlat, S.J., Borja, A., Gilbert, J., Taylor, M.I., Davies, N., Weisberg, S.B., Obst, M. (2013). Genomics
702 in marine monitoring: New opportunities for assessing marine health status. *Marine Pollution*
703 *Bulletin* **74**(1), 19–31.

704 Brook, B., Tonkyn D, O'Grady J, Frankham R. (2002). Contribution of inbreeding to extinction risk in
705 threatened species. *Consevation Ecology* **6** (1):16 [online].

706 Butchart, S., Baillie, J., Chenery, A., Collen, B., Gregory, R., Revenga, C., Walpole, M. (2010). Global
707 Biodiversity: Indicators of Recent Declines. *Sci*ence **328**(5982), 1164-1168.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

708 Cano, J., Mäkinen, H., Merilä, J. (2008). Genetic evidence for male-biased dispersal in the three-spined
709 stickleback (*Gasterosteus aculeatus*). *Methods in Ecology* **16** (14), 3234–3242.

710 Carpenter, K.E. (1999). The Living Marine Resource of Western Central Pacific. Volume4: Bony Fishes
711 (Mugilidae-Carangidae). (Carpenter, K.E. and Niem, V.H., Eds.) (Vol. 3), FAO, Rome.

712 Charlesworth, D., Willis, J. (2009). The genetics of inbreeding depression. *Nature Reviews Genetic*
713 **10**(11), 783–796.

714 Collette, B.B., Carpenter, K. E., Polidoro, B.A., Boustany, A., Die, D.J., Elfes, C., Nelson, R. (2011).
715 High Value and Long Life - Double Jeopardy for Tunas and Billfishes. *Sci*ence **333**(6040), 291–292.

716 Costello, AB., Down TE, Pollard SM, Pacas CJ, Taylor EB. (2003). The influence of history and
717 contemporary stream hydrology on the evolution of genetic diversity within species: an examination
718 of microsatellite DNA variation in bull trout, *Salvelinus confluentus* (Pisces: Salmonidae).
719 *Evolution; international journal of organic evolution* **57**(2), 328–344.

720 Culver, M., Menotti-Raymond, M.A., O'Brien, S. J. (2001). Patterns of Size Homoplasy at 10
721 Microsatellite Loci in Pumas (*Puma concolor*). *Molecular Biology and Evolution* **18** (6), 1151-1156.

722 DeWoody, J. A., Avise, J.C. (2000). Microsatellite variation in marine, freshwater and anadromous
723 fishes compared with other animals. *Journal of Fish Biology* **56**(3), 461–473.

724 Duvernell, D., Lindmeier J, Faust K, Whitehead A. (2008). Relative influences of historical and
725 contemporary forces shaping the distribution of genetic variation in the Atlantic killifish, *Fundulus*
726 *heteroclitus*. *Molecular Ecology* **17**(5), 1344–1360.

727 Esa, Y., Sirajm S., Daud, S. (2008). Mitochondrial DNA diversity of Tor *tambroides Valenciennes*
728 (Cyprinidae) from five natural populations in Malaysia. *Zoological studies* **3**(47), 360–367.

729 Estoup, A., Jarne, P., Cornuet, J.M. (2002). Homoplasy and mutation model at microsatellite loci and
730 their consequences for population genetics analysis. *Molecular Ecology* **11**(9), 1591–1604.

731 Evanno, G., Regnaut, S., Goudet, J. (2005). Detecting the number of clusters of individuals using the
732 software STRUCTURE: A simulation study. *Molecular Ecology* **14**(8), 2611–2620.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

733 Excoffier, L., Laval, G., Schneider, S. (2005). Arlequin (version 3.0): an integrated software package for
734 population genetics data analysis. *Evoltionary Bioinformatics Online* **1**, 47–50.

735 Fang, G., Wang, G., Fang, Y., Fang, W. (2012). A review on the South China Sea western boundary
736 current. *Acta Oceanologica Sinica* **31**(5), 1–10.

737 Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation*
738 *Biol*ogy **10**(6), 1500–1508.

739 Froese, R., Pauly, D. (2015). FishBase. World Wide Web electronic publication.

740 Glaubitz, J.C. (2004). CONVERT: A user-friendly program to reformat diploid genotypic data for
741 commonly used population genetic software packages. *Molecular Ecology Notes* **4**(2), 309–310.

742 Goudet, J. (1995). FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *Journal of*
743 *HERIDITY* **86**(6), 485–486.

744 Gutiérrez, J.P., Royo, L.J., Álvarez, I., Goyache, F. (2005). MolKin v2.0: a computer program for genetic
745 analysis of populations using molecular coancestry information. *Journal of Heredity* **96**, 718-721.

746 Hale, M.L., Burg, T.M., Steeves, T.E. (2012). Sampling for Microsatellite-Based Population Genetic
747 Studies: 25 to 30 Individuals per Population Is Enough to Accurately Estimate Allele Frequencies.
748 *PLoS ONE* **7**(9), 1–12.

749 Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for
750 Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**(1999), 91-98.

751 Hauser, L., Carvalho, G. (2008). Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by
752 beautiful facts. *Fish and Fisheries* **9**(4), 333–362.

753 Hoarau, G., Boon E, Jongma DN, Ferber S, Palsson J, Van der Veer HW, … Olsen JL. (2005). Low
754 effective population size and evidence for inbreeding in an overexploited flatfish, plaice
755 (*Pleuronectes platessa L.*). *Proceedings. Biological sciences / The Royal Society* **272**(1562), 497–
756 503.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

757 Hobbs, J.P.A., Herwerden, V., Jerry, D.R., Jones, G.P., Munday, P.L. (2013). High genetic diversity in
758 geographically remote populations of endemic and widespread coral reef angelfishes (genus:
759 *Centropyge*). *Diversity* **5**(1), 39–50.

760 Hoffmann, M., Hilton-Taylor, C., Angulo, A., Böhm, M., Brooks, T. M., Butchart, S.H.M., Stuart, S.N.
761 (2011). The impact of conservation on the status of the world's vertebrates. *Sci*ence **330**(6010),
762 1503–1509.

763 Jónsson, G. (1982). Contribution to the Biology of Catfish (*Anarhichas lupus*) at Iceland. *Journal of*
764 *Marine Research Institute* **6** (4), 3–26.

765 Kalinowski, S.T. (2005). Do polymorphic loci require large sample sizes to estimate genetic distances?
766 *Journal of Heredity* **94**, 33–36.

767 Khan, M.S.A., Alam, M.J., Rheman, S., Mondal, S., Rahman, M.M. (2002). Study on the Fecundity and
768 GIS of Brackish Catfish *Plotosus canius*. *Online Journal of Biological Sciences* 4(2), 232–234.

769 Krishnamurthy, P.K., Francis, R. (2012). A critical review on the utility of DNA barcoding in biodiversity
770 conservation. *Biodiversity and Conservation* **21**(8), 1901–1919.

771 Kumar, B. (2012). Finding of (Plotosidae: Siluriforms) and its Abundance from Parangipettai Coastal
772 Area – A Review. *International Journal of Recent Scientific Research* **3**(6), 482–485.

773 Larmuseau, M.H.D., Raeymaekers JAM, Hellemans B, Van Houdt JKJ, Volckaert FAM. (2010). Mito-
774 nuclear discordance in the degree of population differentiation in a marine goby. *Journal of Heredity*
775 **105**(6), 532–542.

776 Leffler, E.M., Bullaughey, K., Matute, D.R., Meyer, W.K., Ségurel, L., Venkat, A., Przeworski, M.
777 (2012). Revisiting an Old Riddle: What Determines Genetic Diversity Levels within Species? *PLoS*
778 *Biology* **10**(9), e1001388.

779 Leh, M.U.C., Sasekumar, A., Chew, L.L. (2012). Feeding biology of Eel Catfish *Plotosus Canius*
780 Hamilton in a Malaysian Mangrove Estuary and Mudflat. *Raffles Bulletin of Zoology* **60**(2), 551–
781 557.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

782 Leray, M., Knowlton, N. (2015). DNA barcoding and metabarcoding of standardized samples reveal
783 patterns of marine benthic diversity. *Proceedings of the National Academy of Sciences* **112**(7), 2076-
784 2081.

785 Librado, P., Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism
786 data. *Bioinform*atics **25**(11), 1451–1452.

787 Mccracken, G.R., Wilson KL, Brewis HT, Mcbride MC, Paterson I, Perry R. (2014). Development of 26
788 novel microsatellite makers for the round whitefish (*Prosopium cylindraceum*) and successful
789 polymorphic cross-specific amplification of seven previously developed salmonid markers.
790 *Conservation Genetic Resources* **6**(4), 1023–1026.

791 Magurran, A.E., Baillie, S.R., Buckland, S.T., Dick, J.M., Elston, D.A.., Scottm E,M,, Wattm A,D,
792 (2010), Long-term datasets in biodiversity research and monitoring: Assessing change in ecological
793 communities through time. *Trends in Ecology and Evolution* **25**(10), 574–582.

794 Mohd Fadzil Mohd Akhir, Chuen, Y.J. (2011). Seasonal Variation of Water Characteristics during Inter-
795 Monsoon along the East Coast of Johor . *Journal of Sustainability Science and Management* **6**(2),
796 206-214.

797 Na·Nakorn, U., Taniguchi N, Nugroho E, Seki S, Kamonrat W. (1999). Isolation and characterization of
798 microsatellite loci of *Clarias macrocephalus* and their application to genetic diversity srudy.
799 *Fisheries Science* **65**(4), 520–526.

800 Neff, B.D., Gross, M.R. (2001). Microsatellite evolution in vertebrates: inference from AC dinucleotide
801 repeats. *Evolution; international journal of organic evolution* **55**(9), 1717–1733.

802 O'Leary, S.J., Hice LA, Feldheim KA., Frisk MG, McElroy AE, Fast MD, Chapman DD. (2013). Severe
803 Inbreeding and Small Effective Number of Breeders in a Formerly Abundant Marine Fish. *PLoS*
804 *ONE* **8**(6), 2–9.

805 O'Reilly, P.T., Canino MF, Bailey KM, Bentzen P. (2004). Inverse relationship between *F ST* and
806 microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*):
807 Implications for resolving weak population structure. *Molecular Ecology* **13**(7), 1799–1814.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

808 Paetkau, D., Slade R, Burden M, Estoup A. (2004). Genetic assignment methods for the direct, real-time
809 estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular*
810 *Ecology* **13**(1), 55–65.

811 Palti, Y., Luo A, Charlet GG, Hu Y. (2012). A Second Generation Integrated Map of the Rainbow trout
812 (*Oncorhynchus mykiss*) Genome: Analysis of Conserved Synteny with Model Fsh Genomes. *Marine*
813 *Biotechnology* **14**(3), 343-357.

814 Pauly, D., Christensen, V., Guénette, S., Pitcher, T.J., Sumaila ,U.R., Walters, C.J., Zeller, D. (2002).
815 Towards sustainability in world fisheries. *Nature* **418**(6898), 689–695.

816 Peakall, R., Smouse, P.E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software
817 for teaching and research-an update. *Bioinformatics* **28**(19), 2537–2539.

818 Pereira, H., Leadley P, Proença V, Alkemade R. (2010). Scenarios for global biodiversity in the 21st
819 century. *Science* **330**(6010), 1496–1501.

820 Piry, S., Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A. (2004). GENECLASS2: A software
821 for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**(6), pp.536–
822 539.

823 Piry, S., Luikart, G., Cornuet, J.-M. (1999). BOTTLENECK: A computer program for detecting recent
824 reductions in the effective population size using allele frequency data. *Journal of Heredity*
825 **90**(4),502–503.

826 Priest, M.A., Almany GR, Braun CD, Hamilton RJ, Lozano-cortés DF, Saenz-agudelo P, Berumen ML.
827 (2014). Isolation and characterization of 29 microsatellite markers for the bumphead parrotfish ,
828 *Bolbometopon muricatum* , and cross amplification in 12 related species. *Marine Biodiversity* **44**(4),
829 1-6.

830 Primack, R. (2002). *Essential of conservation genetics* 3rd ed., Sunderland, MA: Sinauer Association
831 Inc.603.pp.

832 Pritchard, J.K., Stephens, M., Donnelly, P. (2000). Inference of population structure using multilocus
833 genotype data. *Genetics* **155**(2), 945–959.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

834 Prithiviraj, N. (2014). Studies on Biological Properties and in Silico Approach of *Plotosus canius*, from
835 Parangipettai Coastal Waters. *International Journal of Modern Research Reviews* **1**(6), 24–37.

836 Prithiviraj, N., Annadurai, D. (2012). Studies on Bioactive Properties of the Catfish *Plotosus canius*
837 (Hamilton , 1822). *International Journal of Recent Scientific Research* **3**(6), 467–473.

838 Rice, W.R. (1989). Analyzing Tables of Statistical Tests. *Evolution* **43**(1),223–225.

839 Riede, K. (2004). Global Register of Migratory Species - from Global to Regional Scales. Final Report of
840 the R&D-Projekt 808 05 081, Federal Agency for Nature Conservation, Bonn, Germany.

841 Rizal, S., Setiawan, I., Iskandar, T., Ilhamsyam, Y., Musman, M., Whahid, M. (2010). Currents
842 Simulation in the Malacca Straits by Using Three-Dimensional Numerical Model. *Sains Malaysiana*
843 **39**(4), 519–524.

844 Rourke, M., Gilligan, D. (2010). Population Genetic Structure of Freshwater Catfish (*Tandanus*
845 *Tandanus*) in the Murray-Darling Basin and Coastal Catchments of New South Wales: Implications
846 for Future Re-stocking, Final report to the NSW Recreational Freshwater Fishing Trust for Project
847 No. DPIFT48. Industry & Investment – Fisheries Final Report Series No. 123. 73pp.

848 Rourke, M.L., Teske PR, Attard CRM, Gilligan DM, Beheregaray LB, Rourke ML, Gilligan DM. (2010).
849 Isolation and characterisation of microsatellite loci in the Australian freshwater catfish (*Tandanus*
850 *tandanus*) Murray – Darling Basin. *Conservation Genetic Resources* **2**(1), 245–248.

851 Rousset, F. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for
852 Windows and Linux. *Molecular Ecology Resources* **8**(1), 103–106.

853 Shaw, P., Arkhipkin, A., Al-Khairulla, H. (2004). Genetic structuring of Patagonian toothfish
854 populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and
855 deep-water troughs as barriers to genetic exchange. *Molecular Ecologgy* **13**(11), 3293–3303.

856 Song, L.M. (2012). Understanding Divergence in *Channa satiata*: A Genetic and Morphological
857 Approach. Master Thesis. University of Malaya, Kuala Lumpur, Malaysia.

858 Tallmon, D. A., Luikart, G., Waples, R.S. (2004). The alluring simplicity and complex reality of genetic
859 rescue. *Trends in Ecology and Evolution* **19**(9), 489–496.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

860 Tamura, K., Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of
861 mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10**(3), 512–526.

862 Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013). MEGA6: Molecular evolutionary
863 genetics analysis version 6.0. *Molecular Biology and Evolution* **30**(12), 2725–2729.

864 Thompson, J.D., Gibson, T. J., Plewniak, F., Jeanmougin, F., Higgins, D.G. (1997). The CLUSTAL X
865 windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis
866 tools. *Nucleic Acids Research* **25**(24), 4876–4882.

867 Thomsen, P.F., Kielgast, J., Iversen, L. L., Møller, P.R., Rasmussen, M., Willerslev, E. (2012). Detection
868 of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples. *PLoS ONE*
869 **7**(8), 1–9.

870 Usman, B., Amin, S.M., Arshad, A., Rahman, M. (2013). Review of Some Biological Aspects and
871 Fisheries of Grey-Eel Catfish *Plotosus canius* (Hamilton, 1822). *Asian Journal of Animal and*
872 *Veterinary Advances* **2**(8), 154–167.

873 Usmani, S., Tan SG, Siraj SS, Yusoff K. (2003). Population structure of the Southeast Asian river catfish
874 *Mystus nemurus*. *Animal Genetics* **34**(6),462–464.

875 Van Oosterhout, C., Hutchinson WF, Wills DPM, Shipley P. (2004). MICRO-CHECKER: Software for
876 identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*
877 **4**(3),535–538.

878 Voris, H.K. (2000). Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems and time
879 durations. *Journal of Biogeography* **27**(5), 1153–1167.

880 Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D.N. (2005), DNA barcoding Australia's
881 fish species. *Philosophical Transactions of the Royal Society of London. Series B, Biological*
882 *Sciences* **360**(1462), 1847–57.

883 Wright, L.I., Tregenza, T. & Hosken, D.J. (2008). Inbreeding, inbreeding depression and extinction.
884 *Conservation Genetics* **9**(4), 833–843.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016

885 Zhang, A.B., He, L.J., Crozier, R.H., Muster, C., Zhu, C.D. (2010). Estimating sample sizes for DNA
886 barcoding. *Molecular Phylogenetic and Evolution* **54**(3), 1035–1039.

887 Zhang, D.X., Hewittm G,M, (1996), Nuclear integrations: challenges for mitochondrial DNA markers.
888 Trends *Ecology & Evolution* **11**(6), 247–251.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1640v2 | CC-BY 4.0 Open Access | rec: 26 Jan 2016, publ: 26 Jan 2016