Journal of FISH BIOLOGY

fsbi An International Society for Fish Biology

Journal of Fish Biology (2016) **88**, 1932–1948
doi:10.1111/jfb.12956, available online at wileyonlinelibrary.com

Genotyping of microsatellite markers to study genetic structure of the wild striped snakehead *Channa striata* in Malaysia

M. P. Tan*†‡§, A. F. J. Jamsari‡ and M. N. Siti Azizah‡

*School of Fisheries and Aquaculture Sciences, Universiti Malaysia Terengganu, 21030, Kuala Terengganu, Terengganu, Malaysia, †Institut Akuakultur Tropika (AKUATROP), Universiti Malaysia Terengganu, 21030, Kuala Terengganu, Terengganu, Malaysiaand ‡School of Biological Sciences, Universiti Sains Malaysia, 11800, Minden, Penang, Malaysia

(Received 31 March 2015, Accepted 18 February 2016)

Genetic variability and differences in wild striped snakehead *Channa striata* from Malaysia were analysed by genotyping nine novel nuclear microsatellite loci. Analysis revealed moderate-to-high genetic diversity in most of the populations, indicative of large effective population sizes. The highly diversified populations are admixed populations and, therefore, can be recommended as potential candidates for selective breeding and conservation since they each contain most of the alleles found in their particular region. Three homogenous groups of the wild populations were identified, apparently separated by effective barriers, in accordance with contemporary drainage patterns. The highest population pairwise F_{ST} found between members of the same group reflects the ancient population connectivity; yet prolonged geographical isolation resulted in adaptation of alleles to local contemporary environmental change. A significant relationship between genetic distance and geographical isolation was observed ($r = 0{\cdot}644$, $P < 0{\cdot}01$). Anthropogenic perturbations indicated apparent genetic proximity between distant populations.

Key words: anthropogenic factor; conservation; freshwater fish; genetic diversity; isolation by distance; nuclear DNA.

INTRODUCTION

Uncovering the genetic structure of biota is as important as understanding environmental changes since both are integrally and naturally related. It becomes increasingly apparent that many of the genetic patterns at the species level could reflect geomorphological historical processes (Adamson *et al.*, 2012; Tan *et al.*, 2012). Investigating connections among populations of freshwater taxa provides unique opportunities since these species can only disperse through waterways, whereas other terrestrial taxa are capable of migrating *via* numerous overland routes. The genetics pattern of freshwater taxa frequently explains the historical drainage realignment.

In Malaysia, both the Peninsula and in Northern Borneo (Sarawak and Sabah), are renowned as rich bio-diverse zoogeographic regions where rivers and tributaries are dominant characteristics. The most prominent topographical feature of Peninsular Malaysia is the central mountain range, running from north to south. This range is

§Author to whom correspondence should be addressed. Tel.: +609 668 4917; email: mptan@umt.edu.my

commonly known as the Titiwangsa Mountains and forms the backbone of Peninsular Malaysia. The mountains originate in southern Thailand and run *c.* 480 km south-east before ending near Jelebu in Negeri Sembilan, effectively dividing the eastern region from western Peninsular Malaysia. Within east and west Peninsular Malaysia, many river systems, lakes and other water bodies, which are not interconnected, support a rich aquatic fauna. Migration of aquatic biota varies from occasional to frequent, increasing among adjacent rivers when floods occur. Natural or anthropogenic physical barriers, such as mountain ranges, dams, drainage rearrangements and agriculture development, affect the mobility of the biotic distribution as a whole.

The striped snakehead, also known as chevron snakehead or snakehead murrel *Channa striata* (Bloch 1793), is indigenous to many Asian countries, particularly the southern region of the Asian continent (Froese & Pauly, 2010). In Malaysia, the names haruan, aruan, toman paya or ruan (Mohsin & Ambak, 1983) are applied to this freshwater fish, which is economically valuable in both the capture and culture sectors. The fish has well-known biomedical properties (Mat Jais *et al.*, 1994; Baie & Sheikh, 2000), good flavour (Inger & Chin, 1962), and wide acceptance as a food fish in Southeast Asian countries, particularly Thailand and Vietnam (Yusoff *et al.*, 2006). These features make the species a candidate for large-scale aquaculture. Production of the species in Malaysia, however, is still at an artisanal level (Mazuki, 2008), where the major source of stock is from wild captures.

Previous reports on the population genetics of *C. striata*, from the Peninsular Malaysia used the random amplified polymorphic DNA method (Ambak *et al.*, 2006), from the Mekong River used both mitochondrial DNA and nuclear microsatellite markers (Adamson *et al.*, 2012) and from Sundaland used mitochondrial DNA markers (Tan *et al.*, 2012). There is significant genetic differentiation among the surveyed populations and a vast genetic differentiation among fish from various sampling sites (Adamson *et al.*, 2012; Tan *et al.*, 2012). High genetic fragmentation is expected for *C. striata*, being restricted by waterways as the only dispersal channel for this obligate and relatively non-migratory freshwater fish (Halls *et al.*, 1998). The genetic fragmentation may be increased as a result of geographic barriers. Research investigating other freshwater fish species (Knight *et al.*, 2009; Huey *et al.*, 2010) bears this out.

In this study, nine novel microsatellite markers were used to analyse the population genetic structure and diversity of *C. striata.* The intention of this study was to investigate genetic diversity, population connectivity and contemporary gene flow that may be present at the independent sampling sites. Due to the economic value of *C. striata*, human interference may also play a role in the species' contemporary genetic distribution and diversity throughout the region. The specific aims were to (1) characterize genetic diversity across all wild populations and (2) define population structure of *C. striata* associated with geophysical barriers (phylogeography). The working hypothesis is that *C. striata* populations may show significant genetic differentiation because their movement is impeded by geophysical barriers, but adjacent populations might share common alleles, because of gene flow and local adaptation.

MATERIALS AND METHODS

ETHICS STATEMENT

Many live specimens from wild populations at all sampling locations were collected from local fishermen. The origins and localities of the specimens were determined by interview of

local fishermen and confirmed to have originated from a single source prior to collection. A small portion of one set of the dorsal or caudal fin rays (*c.* 0·2 cm × 2 cm) from each fish was preserved in 1·5 ml tubes containing a solution of 95% ethanol, while another set was held in TNES-urea (100 mM Tris–HCl, pH 7·5, 125 mM NaCl and 10 mM EDTA, pH 7·5, 1% SDS, 3 M urea) modified from Valles-Jimenez *et al.* (2004). Tissue samples were stored at room temperature (*c.* 25° C) until use. The fish were then returned to the dealers or brought back to the Aquatic House of Universiti Sains Malaysia, Penang, for further aquaculture research. This study was approved by the University's Ethics Committee. All practical steps were taken to ameliorate suffering of the fish throughout the study.

SPECIMEN COLLECTION, DNA EXTRACTION AND PCR OPTIMIZATION AMPLIFICATION

Random sampling of 19 wild populations throughout Peninsular and Malaysian Borneo (Sarawak and Sabah) involving 486 individual *C. striata* was conducted from 2007 to 2010 (Table I and Fig. 1). Sampling locations were divided into five regions following Tan *et al.* (2012): (1) north-west Peninsular, (2) central west Peninsular, (3) southern Peninsular, (4) east Peninsular and (5) Malaysian Borneo. The north-west Peninsular (1) is the area of the west Peninsular confined by the Titiwangsa Mountain Range in the east and the Bintang Mountain Range in the south. The Bintang Mountain Range is located within the Perak state and runs from southern Thailand in the north to the south of Perak. Central west Peninsular (2) includes populations of the west Peninsular region and south of the Bintang Mountain Range, up to the southernmost tip of the Titiwangsa Mountain Range. The southern Peninsular (3) includes populations situated at the south, without confinement by the Titiwangsa Mountain Range. East Peninsular (4) includes populations east of the Titiwangsa Mountain Range. Malaysian Borneo (5) includes island states of Sarawak and Sabah.

DNA templates were isolated by use of AQUAGENOMIC kits (MultiTarget Pharmaceuticals; www.multitargetpharm.com), according to the manufacturer's protocol. Seven microsatellite markers (developed by Jamsari *et al.*, 2011) and two additional GenBank loci (accession numbers GU321680 and HQ993085) were amplified successfully. PCR amplification followed two methods: multiplexing (simultaneous amplification of five loci) and individual PCR, due to the low reproducibility of some loci when amplified in multiplex PCR. Multiplex PCR amplified loci $A05^{(-NED)}$, $E12^{(-VIC)}$, $E11^{(-PET)}$, $C07^{(-VIC)}$ and $F05^{(-6FAM)}$ with the following conditions: 20 ng template DNA, 6·25 µl Master Mix (Qiagen; www.qiagen.com), 1·25 µl Primer Mix and 1·25 µl Q-solution (Qiagen) in 12·50 µl total volume. The following conditions governed individual PCR amplification analysis for loci $A11^{(-NED)}$, $B07^{(-NED)}$, $H09^{(-VIC)}$ and $H02^{(-PET)}$: 20 ng template DNA, 1xPCR buffer (iNtRON; www.intronbio.com), 1·04 mM $MgCl_2$ (iNtRON), 0·04 mM dNTP (iNtRON), 0·42 mM of each primer, and 0·04U *Taq* polymerase (iNtRON) in a total volume of 12·00 µl. Thermal cycling conditions for both PCR amplifications were: 95° C for 3 min, followed by 35 cycles of 94° C for 45 s, 57° C for 45 s, 72° C for 1 min and a final extension at 72° C for 10 min. Satisfactory PCR products were sent for fragment analysis (First BASE Laboratories; www.base-asia.com).

DATA ANALYSIS

Genetic diversity

Diploid genotypic data scoring was performed with GS500LIZ as the internal size standard by using Peak Scanner v 1.0 software (Applied Biosystems; www. appliedbiosystems.com); allele peaks were scored in electrophoretograms as in Arif *et al.* (2010). Initial screening of all genotypic data to check for the presence of null alleles and detection of scoring bias due to stuttering or large allele dropouts was performed with Micro-Checker v 2.2.3 (Van Oosterhout *et al.*, 2004). The raw data were converted into several specific data formats by using the programme CONVERT (Glaubitz, 2004) for various software analyses. Likelihood ratio test of linkage disequilibrium based on the expectation-maximization algorithm (Slatkin & Excoffier, 1996), with 10 000 permutations, was applied to all pairwise comparisons of loci by using Arlequin v 3.1 (Excoffier *et al.*, 2005) to test for significant

TABLE I. Sampling locality, co-ordinate and sample size of *Channa striata* populations in this study (see Fig. 1)

Region	Population	Latitude (north)	Longitude (east)	*n*
North-west Peninsular	1) Kuala Nerang (KN), Kedah	6° 14′ 38″	100° 36′ 22″	25
	2) Seberang Prai (SP), P. Pinang	5° 22′ 08″	100° 23′ 03″	33
	3) Teluk Kumbar (TK), P. Pinang	5° 17′ 04″	100° 14′ 27″	27
	4) Kerian (KR), Perak	4° 59′ 22″	100° 32′ 49″	34
Central-west Peninsular	5) Tanjung Rambutan (TR), Perak	4° 40′ 23″	101° 08′ 52″	30
	6) Tapah (TP), Perak	4° 11′ 50″	101° 15′ 48″	28
	7) Kajang (KJ), K. Lumpur	2° 59′ 42″	101° 47′ 51″	31
Southern Peninsular	8) Linggi (LG), N. Sembilan	2° 35′ 07″	102° 02′ 15″	31
	9) Yong Peng (YP), Johor	2° 14′ 39″	103° 02′ 28″	27
	10) Mersing (MS), Johor	2° 30′ 21″	103° 49′ 06″	31
East Peninsular	11) Kota Bahru (KB), Kelantan	6° 07′ 05″	102° 14′ 23″	26
	12) Binjai (BJ), Terengganu	4° 13′ 43″	103° 22′ 03″	32
	13) Kubang Bujuk, Marang (KT), Terengganu	5° 16′ 38″	103° 02′ 55″	19
	14) Kuala Krau, Mentakap (KK), Pahang	3° 37′ 17″	102° 23′ 07″	17
Malaysian Borneo	15) Kota Belud (SB), Sabah	6° 21′ 03″	116° 25′ 58″	13
	16) Kampung Kesapang (KS), Sabah	6° 21′ 55″	116° 26′ 50″	14
	17) Papar (PP), Sabah	5° 36′ 18″	115° 57′ 42″	16
	18) Sungai Sibuti, Miri (SS), Sarawak	4° 00′ 43″	113° 46′ 31″	27
	19) Serian (SW), Sarawak	1° 02′ 43″	110° 45′ 03″	25
			Total	486

n, sample size.

association between alleles at two loci. An exact test was performed to test for statistically significant departure from linkage equilibrium with 10 000 permutations followed by false discovery rate (FDR) adjustment (Benjamini & Hochberg, 1995) at a significance level of 0·05.

Allelic richness (A_R), number of alleles (N_A) and inbreeding coefficient (F_{IS}) for each locus and population (Weir & Cockerham, 1984) were assessed in FSTAT v 2.9.3 (Goudet, 1995). Allelic richness is a measure of the number of alleles per locus, corrected in this study for variation in population sizes by a rarefaction technique (Kalinowski, 2005); the smallest population sample size was 13 in SB (Table I and Fig. 1). The inbreeding coefficient was calculated based on the difference between observed and expected heterozygosity, with a range of −1 (outbreeding) to +1 (totally identical) to test for global deviation from Hardy-Weinberg equilibrium (HWE). Mean genetic heterozygosity over all loci, and observed (H_O) and expected (H_E) heterozygosities per locus and population were calculated, followed by exact tests with 10 000 steps in a Markov chain and 10 000 dememorization steps, were performed with Arlequin v 3.1.

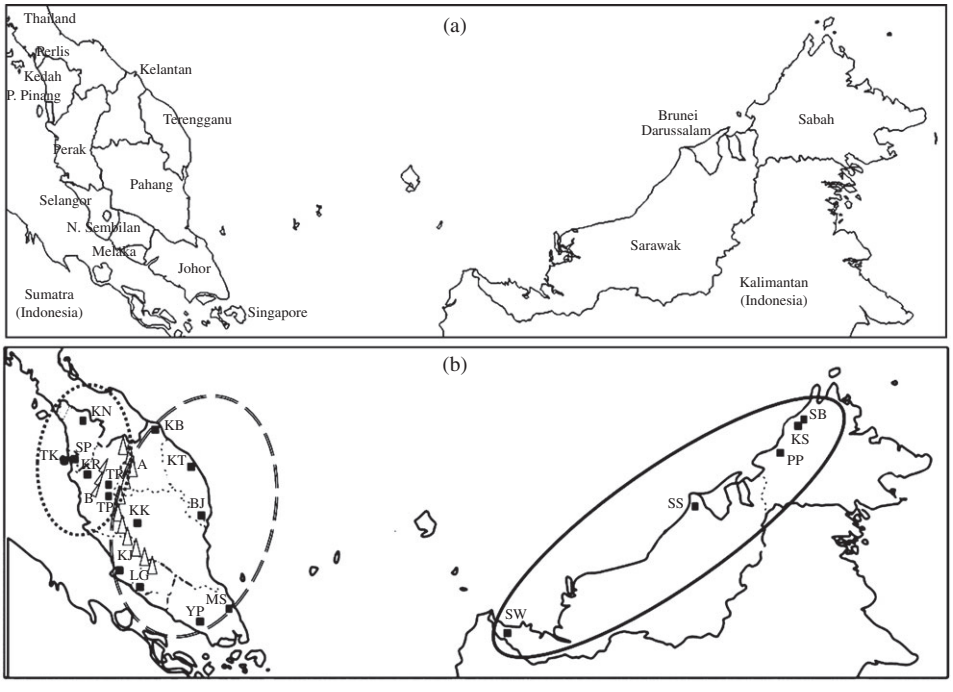

FIG. 1. Map of Malaysia. (a) political and (b) clustering of *Channa striata* populations into three groups as defined in STRUCTURE: - - - -, group 1; — –, group 2; ——, group 3; △, mountains' location; ■, sampling site. A, the Titiwangsa Mountain Ranges; B, the Bintang Mountain Range. KN, Kuala Nerang; SP, Seberang Prai; TK, Teluk Kumbar; KR, Kerian; TR, Tanjung Rambutan; TP, Tapah; KJ, Kajang; LG, Linggi; YP, Yong Peng; MS, Mersing; KB, Kota Bahru; BJ, Binjai; KT, Kubang Bujuk; KK, Kuala Krau; SB, Kota Belud; KS, Kampung Kesapang; PP, Papar; SS, Sungai Sibuti; SW, Serian.

To account for multiple testing in HWE, all calculated probability values were corrected with FDR with a global significance level of 0·05.

The programme BOTTLENECK (Cornuet & Luikart, 1996) was used to detect whether the populations had experienced recent reductions in population size. A two-phased mutational model (95% single stepwise mutation and 5% infinite allele mutation) and a two-tailed Wilcoxon sign-rank test were implemented in this analysis to test for deviation from HWE, followed by FDR adjustment. The qualitative descriptor of allele frequency (mode-shift indicator) analysed in BOTTLENECK differentiated shifted mode populations (bottleneck) from stable populations (Luikart *et al.*, 1998).

Population structure

A hierarchical analysis of molecular variance was performed on all populations to infer the relative attribution of variance among groups (F_{CT}), among populations within groups (F_{SC}), and within populations (F_{ST}) by use of Arlequin v 3.1. The calculation of two pairwise estimates of population differentiation, based on two different mutational models, allowed examination of the differences between each pair of populations. Estimates Rho (ρ), an unbiased estimator for Slatkins' R_{ST} statistic that assumes stepwise mutation, was calculated across total populations and all pairwise population comparisons by use of R_{ST}Calc (Goodman, 1997). All R_{ST} estimates that departed from zero were tested with 1000 permutations. Pairwise estimates using the unbiased estimator theta (θ) (Weir & Cockerham, 1984) of Wright's F_{ST} statistics, which assumes an infinite alleles mutation process were calculated in FSTAT v 2.9.3. Significant departure from

zero was evaluated by 1000 random permutations. Significance levels of all pairwise estimates were adjusted with FDR procedure at $\alpha = 0{\cdot}05$. The relationship between genetic and geographical distances was assessed by use of a Mantel's test in IBD v 1.52 (isolation by distance) (Mantel, 1967; Bohonak, 2002). Population pairwise θ values were used for genetic distance, and approximate geographical distances between sample locations were measured as linear distance in km by using Google Earth. Geographic distance was ln transformed, and the strength of the relationship was examined with reduced major axis regression (10 000 randomizations) in IBD v 1.52.

A neighbour-joining (NJ) tree was constructed to view the relationship between populations based on D_A distance (Nei *et al.*, 1983) in POPTREE2 (Takezaki *et al.*, 2010), and confidence levels at each node were assessed with 1000 bootstrap replications (Felsenstein, 1985). D_A distance appears to be more efficient than F^*_{ST} (Latter, 1972), F'_{ST} (Slatkin, 1993; Rousset, 1997), D_S (Nei, 1972) and $(\delta\mu)^2$ (Goldstein *et al.*, 1995) in obtaining the correct tree topology for multiple loci in microsatellite data (Takezaki & Nei, 1996, 2008).

After determining the probability of the number of groups, k, individuals were assigned to their respective source populations based on multilocus genotypic data by use of the programme STRUCTURE v 2.3 (Pritchard *et al.*, 2000). Correlated allele frequencies among populations (Falush *et al.*, 2003) and admix model assumptions corresponded with the burn-in period and Markov chain Monte Carlo length, each at 10 000 and 10 iterations, were used. An examination of $k = 1–10$ determined the true number of groups, k, as in Evanno *et al.* (2005). The highest peak of $\Delta k = k$ was chosen, where k is the most probable number of groups for the entire dataset. Posterior probability of any admixed individuals belonging to the original population was calculated to assign them back to the original population. The probabilities of genotype assignment into each individual group were performed across replicates by use of CLUMPP v 1.1.2 (Jakobsson & Rosenberg, 2007), and the graphical presentation was carried out using DISTRUCT v 1.1 software (Rosenberg, 2004). A NJ tree (Saitou & Nei, 1987) was constructed to demonstrate the genetic divergence among groups of populations based on the allelic frequency (net nucleotide distance) by use of DRAWTREE software (Felsenstein, 2005).

RESULTS

A total of 19 populations, consisting of 486 individuals of *C. striata* in Malaysia were genotyped for all nine microsatellites DNA loci. Population size ranged from 13 individuals in SB to 33 in SP, with an average of 26 individuals per population (Table I), after eliminating individuals with missing data for at least one locus. Only 8·0% (55 of 684) of pairwise comparisons displayed linkage disequilibria. Dropout of large alleles produced negligible scoring error. A significant shortage of heterozygotes with one unit repeat difference (Van Oosterhout *et al.*, 2004) indicated stuttering scoring errors at a few loci in random populations. Similarly, a few loci had null alleles, as indicated by a general excess of homozygotes. Mutations in flanking regions were not likely to have been the cause for this observation (Dewoody *et al.*, 2006; Neville *et al.*, 2007) as there was no consistent trend in linkage disequilibrium involving individual loci across all populations. Thus, it appears that the departure from HWE relates to the population sampling stochasticity and not on inherent characteristics of the markers used.

GENETIC DIVERSITY

The number of alleles per locus ranged from six at *CS-F05* ($A_R = 4{\cdot}6$) to 24 at *CS-E12* ($A_R = 12{\cdot}9$) (Table II). Within populations, the average number of alleles over the nine loci ranged from 3·2 in SS to 10·1 in TR, with a moderate average allele number of 5·7 per locus. The highest genetic diversity, as assessed by all related measures (private alleles, highest mean number of alleles, mean expected heterozygosity), was recorded

in TR followed by KJ, KR and LG. In contrast, the SB, SS, KN and KT populations were among the least diverse. A moderate mean H_E over all loci and populations (0·58) ranged from 0·33 to 0·79 (Table III).

A total of 47 out of 171 HWE tests (27·5%) departed significantly from random mating, with all but two of them having reduced observed heterozygosity. Specific loci across populations displayed no consistent linkage pattern, which suggests that all the loci used in this study provided independent assessments of genetic variation among populations. Within populations, seven of nine loci of MS population had a significant heterozygosity deficit from HWE. The relatively high F_{IS} mean value in MS ($F_{IS} = 0{\cdot}42$) also reflects this deficit. Four populations (KS, SB, PP and KT) were in HWE at all loci, whereas other populations showed departures from HWE at one to five loci (except for MS, which had seven). Only the KN population had lower inbreeding probability than expected ($F_{IS} = -0{\cdot}08$). Mean F_{IS} value in each population was low-to-moderate and ranged from −0·08 to 0·42 (Table III).

The Wilcoxon sign-rank test indicated that only one population was significantly in mutation-drift disequilibrium ($\alpha < 0{\cdot}05$). Population SS displayed significant excess heterozygosity and had probably undergone recent reduction in population size, as indicated by the shifted mode, whereas other populations had a pattern of mutation-drift equilibrium, as indicated by the typical L-shaped distribution allele frequency. All populations, except SS, retained a relatively high number of rare alleles (allele frequency <0·1), which resulted in L-shaped graphs, indicating that these populations had not experienced a population-size bottleneck (Cornuet & Luikart, 1996; Luikart *et al.*, 1998; MacAvoy *et al.*, 2007).

POPULATION STRUCTURE

A hierarchical analysis of molecular variance performed on all populations without predefined group clustering revealed that 28·1% of the total genetic variance was among populations, while 71·9% was within populations. All pairwise comparisons of F_{ST} values revealed significant differentiation, except for the SB-KS pair (Table IV); these Sabah populations had a negative F_{ST} value of −0·014. In other pairwise comparisons, F_{ST} ranged from 0·048 (KK-LG) to 0·591 (SB-SW), with an average of 0·279. Pairwise comparisons F_{ST} of SW-SB and SW-KN also had relatively high differentiation. Inspection on intraregional populations showed relatively higher pairwise F_{ST} values of populations within Malaysian Borneo than with those of other regions. Specifically, high F_{ST} values were found between SW and all three Sabah populations (SB, KS and PP); F_{ST} values ranged from 0·489 to 0·591, the highest amongst all pairwise comparisons. Pairwise F_{ST} between both Sarawak populations of SW and SS was 0·347. By comparison, the highest value within a region other than Malaysian Borneo reached only 0·282 (KN-KR from the north-west Peninsular).

The analogous pairwise estimates of R_{ST} revealed an additional nine pairs of populations that lacked differentiation; these populations were mainly intraregional populations. Unexpectedly, PP-TR also had a lack of significant differentiation. The other values were in agreement with the F_{ST} statistics. A significant pattern of isolation by distance was found among the samples ($r = 0{\cdot}644$, $P < 0{\cdot}01$; $r^2 = 0{\cdot}414$).

Despite the high genetic differentiation detected between SW and the other Malaysian Borneo populations (between SW and Sabah's populations SB, KS and PP), the NJ tree linked these populations genetically and clustered them as a single

TABLE II. Allelic richness (A_R), number of alleles (N_A) and number of unique alleles (in parenthesis) at each locus and each population (see Table I and Fig. 1)

Locus	Population	KN	SP	TK	KR	TR	TP	KJ	LG	YP	MS	KB	BJ	KT	KK	SB	KS	PP	SW	SS	Total
	n	25	33	27	34	30	28	31	31	27	31	26	32	19	17	13	14	16	25	27	486
C07	A_R	3·7	4·4	5·5	7·1	9·6	6·0	8·2	7·7	5·4	5·0	6·9	5·3	3·0	4·5	2·0	2·9	3·8	3·1	2·0	9·1
	N_A	4	6	6	9	13	8	10	9	8	6	8	6	3	5	2	3	4	4(1)	2	18
A05	A_R	2·6	3·6	3·0	5·7	6·5	3·8	5·9	5·4	3·0	3·1	6·6	5·4	3·0	3·8	4·0	5·9	3·0	2·3	2·0	6·5
	N_A	4	4	3	8(1)	7(2)	4	7	7	4	4	8	7	3	4	4	6	3(1)	3	2(1)	13
F05	A_R	2·0	3·0	4·2	4·9	4·6	4·8	3·0	3·0	1·9	2·7	3·3	1·8	2·0	3·0	3·0	3·0	3·0	2·3	3·0	4·6
	N_A	2	3	5	5	5	5	3	3	2	3	4(1)	2	2	3	3	3	3	3	3	6
E12	A_R	4·3	9·2	7·5	7·8	8·1	10·5	10·3	10·2	4·0	4·1	6·7	7·4	4·0	10·1	5·0	4·9	5·6	4·5	4·0	12·9
	N_A	5(1)	11	8	10	10	13	15	14	5(1)	5	8	9	5	11	5	5	6	5	4	24
E11	A_R	4·0	5·8	5·0	5·2	6·2	5·6	6·1	5·2	3·5	2·4	6·1	2·9	1·0	4·9	3·0	3·0	4·6	3·0	3·0	6·7
	N_A	4	7	5	6	8	6	7	6	4	3	7	3	1	5	3	3	5	4	3	10
H09	A_R	3·5	5·4	4·0	7·4	11·4	4·4	7·2	7·4	5·5	5·0	6·0	5·5	5·5	4·0	1·0	1·0	8·0	1·5	4·8	9·5
	N_A	4	7	5	12	17(2)	5	10	10	7	8	8	6	6	4	1	1	9(1)	2	5	21
B07	A_R	3·0	3·3	3·0	4·1	5·6	4·7	5·8	5·2	5·1	3·5	4·9	3·9	4·1	4·0	3·0	3·0	3·8	2·4	3·0	7·2
	N_A	3	4	3	6	8	6	6	6	6(1)	4	5	4	5	4	3	4(1)	4	3	3	11
A11	A_R	3·6	5·8	6·0	5·6	9·9	6·1	11·6	9·9	10·1	8·2	6·4	7·2	3·9	10·0	3·0	3·0	3·0	5·4	3·5	12·1
	N_A	4	7	8	8	14	8(1)	15	12(1)	13	10	8	10(1)	4	11	3	5	3	6	4	22
H02	A_R	2·0	5·2	3·7	5·1	6·8	4·6	6·2	3·8	3·9	4·4	6·1	6·1	2·9	3·8	3·0	3·0	3·8	4·5	2·7	6·6
	N_A	2	6	4(1)	8	9(1)	5	8(2)	4	6	5	8	7	3	4	3	3	4	5	3	17
Average N_A	3·6	6·1	5·2	8·0	10·1	6·7	9·0	7·9	6·1	5·3	7·1	6·0	3·6	5·7	3·0	3·7	4·6	3·9	3·2	5·7	
(N_A)	(32)	(55)	(47)	(72)	(91)	(60)	(81)	(71)	(55)	(48)	(64)	(54)	(32)	(51)	(27)	(33)	(41)	(35)	(29)		

BJ, Binjai; KB, Kota Bahru; KJ, Kajang; KK, Kuala Krau; KN, Kuala Nerang; KR, Kerian; KS, Kampung Kesapang; KT, Kubang Bujuk; LG, Linggi; MS, Mersing; *n*, sample size; PP, Papar; SB, Kota Belud; SP, Seberang Prai; SS, Sungai Sibuti; SW, Serian; TK, Teluk Kumbar; TP, Tapah; TR, Tanjung Rambutan; YP, Yong Peng.

TABLE III. Population genetic diversity as measured by observed (H_O) and expected (H_E) heterozigosities and inbreeding coefficient (F_{IS}) (see Table I and Fig. 1). Bold values indicate significant deviation from Hardy-Weinberg equilibrium after FDR correction at $\alpha = 0{\cdot}05$

Locus		Pop *n*	KN 25	SP 33	TK 27	KR 34	TR 30	TP 28	KJ 31	LG 31	YP 27	MS 31	KB 26	BJ 32	KT 19	KK 17	SB 13	KS 14	PP 16	SW 25	SS 27
C07	H_O		**0·60**	0·30	0·70	0·85	0·67	**0·46**	0·77	0·90	0·52	**0·48**	**0·73**	0·59	0·63	0·53	0·08	0·14	0·31	0·12	0·33
	H_E		**0·56**	0·33	0·72	0·78	0·82	**0·72**	0·80	0·80	0·47	**0·74**	**0·83**	0·64	0·58	0·56	0·08	0·14	0·38	0·19	0·51
	F_{IS}		−0·08	0·01	0·02	−0·09	0·19	0·36	0·03	−0·13	−0·12	0·35	0·12	0·07	−0·08	0·06	0·00	−0·02	0·19	0·37	0·35
A05	H_O		0·12	0·55	0·52	0·62	0·83	0·61	0·55	0·39	0·18	**0·10**	0·69	0·72	0·58	0·65	0·46	0·43	0·38	0·12	0·19
	H_E		0·12	0·56	0·58	0·74	0·79	0·52	0·71	0·49	0·18	**0·26**	0·75	0·66	0·59	0·53	0·40	0·61	0·41	0·12	0·28
	F_{IS}		−0·02	0·03	0·10	0·16	−0·06	−0·06	0·23	0·21	−0·05	0·64	0·07	−0·09	0·02	−0·22	−0·15	0·31	0·09	−0·03	0·35
F05	H_O		0·28	0·48	**0·96**	0·65	**0·37**	**0·32**	0·48	0·42	0·04	**0·23**	0·58	0·09	0·36	0·35	0·54	0·57	0·56	0·12	0·56
	H_E		0·25	0·58	**0·70**	0·72	**0·63**	**0·52**	0·63	0·44	0·11	**0·35**	0·54	0·09	0·37	0·43	0·59	0·59	0·66	0·12	0·67
	F_{IS}		−0·14	0·17	−0·39	0·11	0·43	0·39	0·24	0·05	0·66	0·36	−0·07	−0·03	0·01	0·19	0·10	0·04	0·16	−0·03	0·17
E12	H_O		0·76	**0·36**	**0·26**	**0·15**	**0·10**	**0·43**	**0·39**	**0·58**	**0·07**	**0·35**	0·66	0·75	0·21	**0·53**	0·54	0·79	0·38	**0·40**	**0·41**
	H_E		0·62	**0·88**	**0·85**	**0·71**	**0·86**	**0·90**	**0·89**	**0·86**	**0·33**	**0·60**	0·79	0·78	0·25	**0·87**	0·58	0·63	0·55	**0·62**	**0·70**
	F_{IS}		−0·24	0·59	0·70	0·80	0·89	0·53	0·57	0·33	0·78	0·42	0·18	0·03	0·15	0·40	0·08	−0·26	0·33	0·36	0·42
E11	H_O		0·84	**0·73**	0·81	0·74	**0·40**	0·64	0·65	0·52	0·41	0·19	**0·35**	**0·19**	–	0·53	0·15	0·36	0·56	0·12	0·67
	H_E		0·64	**0·78**	0·78	0·73	**0·81**	0·72	0·65	0·56	0·56	0·23	**0·70**	**0·33**	–	0·57	0·28	0·37	0·57	0·25	0·63
	F_{IS}		−0·32	0·07	−0·05	−0·01	0·51	0·12	0·01	0·08	0·27	0·17	0·51	0·43	NA	0·07	0·47	0·04	0·02	0·53	−0·07
H09	H_O		0·68	**0·33**	0·70	**0·50**	0·73	0·61	**0·48**	**0·55**	0·67	**0·32**	0·73	0·72	0·63	0·71	–	–	0·69	0·04	0·59
	H_E		0·65	**0·69**	0·65	**0·60**	0·86	0·71	**0·79**	**0·79**	0·67	**0·65**	0·71	0·64	0·64	0·73	–	–	0·74	0·04	0·66
	F_{IS}		−0·05	0·52	−0·08	0·17	0·15	0·15	0·39	0·31	0·01	0·51	−0·04	−0·12	0·02	0·04	NA	NA	0·08	0·00	0·11
B07	H_O		0·52	0·33	0·37	0·50	0·67	0·57	0·58	0·68	0·63	0·58	0·81	0·50	0·47	0·65	0·46	0·43	0·75	0·08	0·67
	H_E		0·49	0·32	0·50	0·63	0·70	0·58	0·77	0·72	0·71	0·58	0·71	0·65	0·53	0·66	0·60	0·55	0·70	0·15	0·62
	F_{IS}		−0·07	−0·04	0·26	0·22	0·05	0·01	0·25	0·06	0·12	0·01	−0·15	0·23	0·11	0·02	0·24	0·23	−0·08	0·48	−0·07
A11	H_O		**0·04**	**0·55**	**0·30**	0·56	**0·63**	**0·39**	0·90	**0·68**	**0·74**	**0·32**	0·50	**0·59**	0·58	0·94	0·31	0·29	0·38	0·56	0·59
	H_E		**0·32**	**0·71**	**0·54**	0·60	**0·87**	**0·63**	0·92	**0·89**	**0·87**	**0·73**	0·78	**0·75**	0·63	0·89	0·34	0·27	0·48	0·73	0·61
	F_{IS}		0·88	0·24	0·46	0·07	0·27	0·38	0·02	0·24	0·17	0·56	0·36	0·21	0·08	−0·06	0·11	−0·06	0·22	0·24	0·03
H02	H_O		0·48	**0·27**	**0·19**	0·53	0·67	0·61	**0·45**	0·74	0·30	**0·26**	**0·42**	**0·53**	0·16	0·41	0·38	0·57	0·56	**0·28**	**0·15**
	H_E		0·37	**0·71**	**0·59**	0·59	0·77	0·64	**0·76**	0·70	0·40	**0·73**	**0·72**	**0·78**	0·32	0·67	0·38	0·44	0·47	**0·77**	**0·48**
	F_{IS}		−0·30	0·62	0·69	0·11	0·14	0·05	0·41	−0·06	0·27	0·65	0·41	0·33	0·52	0·40	0·00	−0·31	−0·22	0·64	0·69
Mean	H_E		0·45	0·62	0·66	0·68	0·79	0·67	0·77	0·69	0·48	0·54	0·72	0·59	0·44	0·66	0·36	0·40	0·55	0·33	0·57
Mean	F_{IS}		−0·08	0·30	0·19	0·17	0·29	0·23	0·24	0·13	0·18	0·42	0·17	0·12	0·08	0·11	0·11	0·01	0·08	0·39	0·20

BJ, Binjai; KB, Kota Bahru; KJ, Kajang; KK, Kuala Krau; KN, Kuala Nerang; KR, Kerian; KS, Kampung Kesapang; KT, Kubang Bujuk; LG, Linggi; MS, Mersing; *n*, sample size; PP, Papar; SB, Kota Belud; SP, Seberang Prai; SS, Sungai Sibuti; SW, Serian; TK, Teluk Kumbar; TP, Tapah; TR, Tanjung Rambutan; YP, Yong Peng.

TABLE IV. Population pairwise divergence (F_{ST} lower left; R_{ST} upper right) of *Channa striata* populations (see Table I and Fig. 1) as inferred from multilocus microsatellite markers

Region	North-west Peninsular				Central-west Peninsular			Southern Peninsular			East Peninsular				Malaysian Borneo				
Population	KN	SP	TK	KR	TR	TP	KJ	LG	YP	MS	KB	BJ	KT	KK	SB	KS	PP	SS	SW
KN		0·157	0·122	0·230	0·228	0·282	0·338	0·415	0·563	0·482	0·346	0·474	0·486	0·468	0·354	0·383	0·343	0·406	0·575
SP	0·252		0·077	0·246	0·245	0·275	0·370	0·410	0·527	0·502	0·367	0·534	0·551	0·438	0·418	0·435	0·339	0·417	0·550
TK	0·280	0·055		0·187	0·160	0·218	0·280	0·332	0·432	0·407	0·337	0·450	0·484	0·339	0·301	0·319	0·256	0·398	0·507
KR	0·282	0·120	0·122		0·024†	0·039†	0·115	0·142	0·315	0·198	0·095	0·187	0·150	0·227	0·223	0·247	0·132	0·193	0·371
TR	0·284	0·166	0·132	0·086		0·048†	0·079	0·101	0·209	0·165	0·105	0·147	0·118	0·154	0·107	0·127	0·035†	0·165	0·354
TP	0·296	0·181	0·178	0·092	0·072		0·183	0·210	0·350	0·286	0·178	0·307	0·233	0·288	0·279	0·296	0·167	0·254	0·379
KJ	0·321	0·202	0·175	0·124	0·055	0·145		0·144	0·292	0·112	0·094	0·104	0·122	0·198	0·250	0·266	0·132	0·199	0·442
LG	0·346	0·264	0·233	0·152	0·098	0·159	0·120		0·129	0·069	0·124	0·105	0·189	0·030†	0·186	0·168	0·093	0·150	0·322
YP	0·479	0·401	0·374	0·296	0·267	0·341	0·249	0·190		0·282	0·316	0·343	0·434	0·046†	0·378	0·351	0·224	0·342	0·386
MS	0·473	0·381	0·361	0·306	0·238	0·322	0·207	0·166	0·219		0·183	0·120	0·258	0·149	0·348	0·327	0·208	0·297	0·470
KB	0·375	0·268	0·241	0·188	0·095	0·186	0·077	0·144	0·295	0·254		0·068	0·042†	0·226	0·216	0·239	0·135	0·046	0·325
BJ	0·442	0·326	0·309	0·219	0·178	0·281	0·113	0·125	0·221	0·210	0·126		0·107	0·221	0·306	0·309	0·197	0·196	0·453
KT	0·497	0·401	0·375	0·310	0·220	0·351	0·177	0·242	0·389	0·354	0·189	0·188		0·331	0·402	0·416	0·221	0·179	0·503
KK	0·382	0·278	0·247	0·174	0·112	0·190	0·122	0·048	0·221	0·202	0·165	0·138	0·253		0·249	0·223	0·142	0·238	0·351
SB	0·537	0·387	0·356	0·318	0·269	0·336	0·279	0·376	0·518	0·491	0·324	0·403	0·529	0·393		−0·025†	0·049†	0·232	0·478
KS	0·530	0·377	0·346	0·306	0·258	0·325	0·261	0·356	0·500	0·470	0·310	0·384	0·509	0·373	−0·014†		0·052†	0·247	0·472
PP	0·441	0·312	0·278	0·245	0·175	0·246	0·190	0·275	0·419	0·402	0·240	0·315	0·374	0·274	0·130	0·133		0·146	0·424
SS	0·446	0·328	0·298	0·245	0·197	0·270	0·186	0·279	0·381	0·393	0·229	0·307	0·398	0·296	0·182	0·181	0·147		0·362
SW	0·590	0·484	0·463	0·384	0·352	0·438	0·354	0·382	0·453	0·496	0·382	0·413	0·548	0·444	0·591	0·586	0·489	0·347	

BJ, Binjai; FDR, false discovery rate; KB, Kota Bahru; KJ, Kajang; KK, Kuala Krau; KN, Kuala Nerang; KR, Kerian; KS, Kampung Kesapang; KT, Kubang Bujuk; LG, Linggi; MS, Mersing; PP, Papar; SB, Kota Belud; SP, Seberang Prai; SS, Sungai Sibuti; SW, Serian; TK, Teluk Kumbar; TP, Tapah; TR, Tanjung Rambutan; YP, Yong Peng.

†Non-significant values after FDR correction at $\alpha = 0{\cdot}05$, all other values are significant.

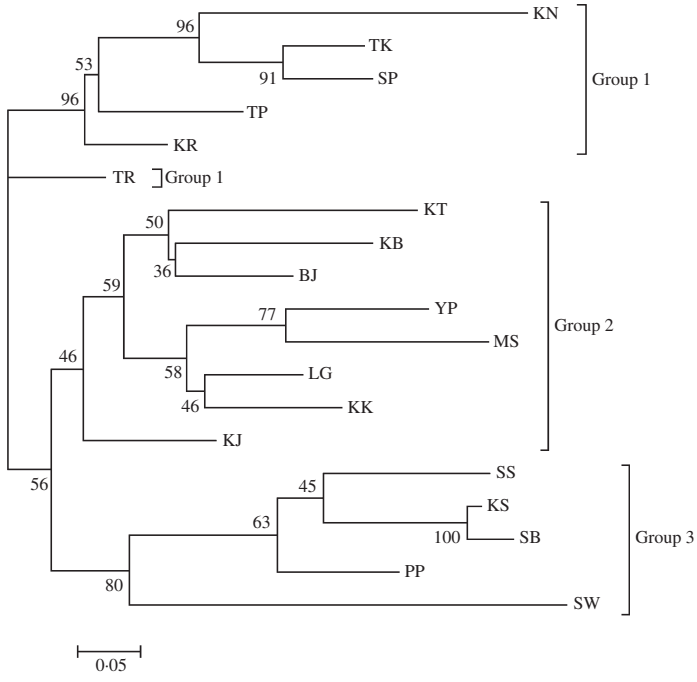

FIG. 2. The unrooted neighbur-joining (NJ) tree inferred from multilocus microsatellite DNA based on D_A distance. Each branch has a denoted NJ bootstrap value. Group 1 (all north-west Peninsular with two central-west Peninsular populations), group 2 (all east Peninsular, southern Peninsular and one central-west Peninsular populations) and group 3 (all populations of Malaysian Borneo) (see Fig. 1).

group (Fig. 2). Population SW, which appeared differentiated from its group members, simply had a longer branch length in the NJ tree. Separation of populations was supported by the Bayesian clustering analysis, which detected three phylogenetic groups ($\Delta k = 3$) among the wild populations of *C. striata* in Malaysia (Fig. 3). Assignment of the populations into respective groups based on multi-locus genotypic data revealed that the segregation of individual populations correlated with geographical distribution. STRUCTURE clustered all north-west Peninsular populations (KN, SP, TK and KR) with two populations from the central west Peninsular (TR and TP) populations as one group (group 1); all east Peninsular, southern Peninsular, and one of the central west Peninsular populations (KJ) belonging to group 2, and all Malaysian Borneo populations, group 3 (Fig. 1). The algorithm revealed a relatively high membership coefficient of the individuals with their respective groups. Groups 1 and 3 (0·1712) had the highest genetic divergence, followed by groups 2 and 3 (0·1633), then groups 1 and 2 (0·1228).

DISCUSSION

GENETIC DIVERSITY

The overall mean expected heterozygosity and mean number of alleles per locus for *C. striata* were $H_E = 0{\cdot}58$ and $N_A = 5{\cdot}7$, in agreement with a comprehensive study of

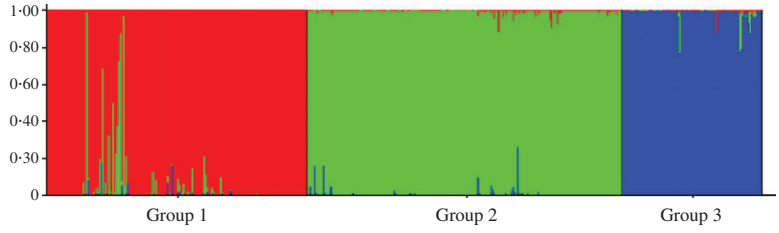

FIG. 3. Bar plots show the three-clustered population structure of *Channa striata* into three groups. The different coloured fine lines in the major groups indicate admixed individuals. Group 1 (all north-west Peninsular with two central west Peninsular populations), group 2 (all east Peninsular, southern Peninsular and one central-west Peninsular populations) and group 3 (all Malaysian Borneo populations) (see Fig. 1).

microsatellite variation in marine, freshwater and anadromous fishes, and other organisms (DeWoody & Avise, 2000). Those authors reported a population-level genetic variation in freshwater fishes of $H_E = 0{\cdot}46$ and $N_A = 7{\cdot}5$. All diversity analyses consistently indicated that TR, KJ, LG and KR were the most highly diversified populations, with an average number of alleles per locus ranging from 3·2 to 10·1 and expected heterozygosities of 0·33–0·79. The relatively moderate-to-high genetic diversity in those populations reflected large effective population sizes. Another plausible explanation was that of allele contribution by migrant individuals from genetically distinctive stocks. TR (pre-assigned to group 1) represented a more complex population, with up to 40% of its members significantly close to group 2. In this instance, it was believed that the geographically adjacent population, KJ (assigned as group 2 after analysis in STRUCTURE), other than TP, was the main gene contributor, and detailed inspection of diploid data supported this hypothesis (Fig. 3). Common alleles found between these two populations and a lower population pairwise F_{ST} value observed between them indicated the presence of gene flow between the populations. No severe bottleneck or reduction in population size existed except in SS, as noted by a significant reduction in its rare alleles. Translocated stock in the SS population (Tan *et al.*, 2012), from sources other than Sabah's populations (SB, KS and PP), explains the reduced effective population size, which reflects small numbers of mating pairs and expressed as inbreeding signature or founder effect (Lippé *et al.*, 2006).

POPULATION STRUCTURE

Three phylogenetic groups are associated with phylogeographic relationships. Peninsular Malaysia populations were divided into two discrete groups by the Titiwangsa Mountain Range to form north-west and east-south distinct lineages. This insurmountable physical impediment has limited genetic exchange between the west–east regions and, after a prolonged period of time, the isolated stocks evolved independently to form separate lineages. Similarly, the isolation of the Malaysian Borneo states of Sarawak and Sabah from Peninsular Malaysia since sea level rose *c.* 10 000 years ago towards the end of the Pleistocene Epoch, permitted accumulation of private alleles over time to form a single distinct region. Within and between groups, almost all of population pairwise comparisons showed a high level and significant genetic structuring and differentiation (particularly between groups); this observation is a typical freshwater fish

characteristic (Barson *et al.*, 2009; Huey *et al.*, 2010) and is supported by the significant correlation between genetic distance and geographical isolation in IBD analysis. Exceptions to this pattern were 10 population pairwise R_{ST} values; these involved adjacent localities (within the same state), a finding which suggests common ancestry.

In contrast, within group 3, high population pairwise genetic distance comparisons between SW and group members (SB, KS, PP and SS) indicate SW's genetic distinction from the rest of the Malaysian Borneo populations in Sarawak and Sabah. This finding is similar to the previous study based on the mtDNA *ND5* gene by Tan *et al.* (2012). Compared with the findings in this study, however, SW was previously included with populations from the east-south Peninsular (Tan *et al.*, 2012). Esa *et al.* (2006) made a similar observation in their study on a freshwater fish species, *Tor douronensis* (Valenciennes 1842), within the Malaysian Borneo states of Sarawak and Sabah, which was based on data from the mtDNA *CO1* gene. The authors revealed distinct haplotype sequences in Sabah populations, with high genetic differentiation from the populations in Sarawak, a finding that suggested the Sabah populations were genetically isolated and physically disconnected from those of other Borneo regions. The authors further showed that *T. douronensis* populations from the northern part of Sarawak are significantly differentiated from populations in west Sarawak, a genetic pattern that was also observed in *C. striata* (SS is located at the northern part of Sarawak and SW to the west). These studies on mtDNA produce a more reliable range of historical signatures than do contemporary events detected by nuclear markers, such as microsatellites. In this study, despite the relatively high differentiation (based on pairwise F_{ST} values) detected between SW and the other Malaysian Borneo populations (SB, KS, PP and SS), concordant results of the NJ phylogenetic tree and STRUCTURE analyses linked all Malaysian Borneo populations as a homogenous group. These results corroborated the substantial genetic affinity among the populations as measured by the genotype frequency input for the latter analysis.

The relatively high populations pairwise comparisons F_{ST} and R_{ST} are similar to those in *C. striata* populations of Mekong (Adamson *et al.*, 2012) and other freshwater fishes such as the marble trout *Salmo marmoratus* Cuvier 1829 (Fumagalli *et al.*, 2002), the three-spined stickleback *Gasterosteus aculeatus* L. 1758 (Caldera & Bolnick, 2008) and the grass carp *Ctenopharyngodon idella* (Valenciennes 1844) (Liu *et al.*, 2009). Chauhan *et al.* (2007) postulated that low population structuring in several riverine species, such as the Indian major carp *Cirrhinus mrigala* (Hamilton 1822), resulted from common ancestry and gene transfer between different river basins. The significant genetic partitioning among the majority of *C. striata* populations, as shown in this study, indicates non-migratory behaviour (Halls *et al.*, 1998), which is an important consideration in planning conservation and breeding strategies.

Population pairwise R_{ST} showed that TR from the central west Peninsular and the Malaysian Borneo population, PP, is somewhat homogenous. The genetic homogeneity observed over these geomorphologically distant and presently physically disconnected regions, suggests gene transfer through ancient dispersal or recent translocation. Since there is no documentation of connectivity of an ancient river system or tributaries between these two regions (Voris, 2000; Esa *et al.*, 2006; Kamarudin & Esa, 2009), ancient dispersal is unlikely. On the other hand, there is evidence that the shipping of live *C. striata* was a common practice in the 1950s. Moreover, this species had the special capability to aestivate and stay alive when transported (Schuster, 1952; Courtenay *et al.*, 2004). Thus, available evidence indicates that anthropogenic translocation is

the most plausible explanation for the sporadic genetic propinquity between *C. striata* from TR and PP.

Our utmost appreciation is extended to G. K. Chambers from Victoria University of Wellington, New Zealand, and D. Hurwood and P. Mather from Queensland University of Technology, Australia, for their helpful guidance and advice. Thanks are also due to our colleagues at Universiti Sains Malaysia for their assistance throughout the study. Universiti Sains Malaysia funded this research under Research University Grant 1001/PBIOLOGI/8150123.

References

Adamson, E. A. S., Hurwood, D. A. & Mather, P. B. (2012). Insights into historical drainage evolution based on the phylogeography of the chevron snakehead fish (*Channa striata*) in the Mekong Basin. *Freshwater Biology* **57,** 2211–2229. doi: 10.1111/j.1365-2427.2012.02864.x

Ambak, M. A., Ambok, B. A. M., Ismail, P. & Bui, M. T. (2006). Genetic variation of snakehead (*Channa striata*) populations using random amplified polymorphic DNA. *Biotechnology* **5,** 104–110. doi: 10.3923/biotech.2006.104.110

Arif, I. A., Khan, H. A., Shobrak, M., Al Homaidan, A. A., Al Sadoon, M., Al Farhan, A. H. & Bahkali, A. H. (2010). Interpretation of electrophoretograms of seven microsatellite loci to determine the genetic diversity of the Arabian Oryx. *Genetics and Molecular Research* **9,** 259–265. doi: 10.4238/vol9-1gmr714

Baie, S. H. & Sheikh, K. A. (2000). The wound healing properties of *Channa striatus* cetrimide cream: tensile strength measurement. *Journal of Ethnopharmacology* **71,** 93–100. doi: 10.1016/S0378-8741(99)00184-1

Barson, N. J., Cable, J. & Van Oosterhout, C. (2009). Population genetic analysis of microsatellite variation of guppies (*Poecilia reticulata*) in Trinidad and Tobago: evidence for a dynamic source-sink metapopulation structure, founder events and population bottlenecks. *Journal of Evolutionary Biology* **22,** 485–497. doi: 10.1111/j.1420-9101.2008.01675.x

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B* **57,** 289–300. doi: 10.2307/2346101

Bohonak, A. J. (2002). IBD (Isolation By Distance): a program for analyses of isolation by distance. *Journal of Heredity* **93,** 153–154. doi: 10.1093/jhered/93.2.153

Caldera, E. J. & Bolnick, D. I. (2008). Effects of colonization history and landscape structure on genetic variation within and among three spine stickleback (*Gasterosteus aculeatus*) populations in a single watershed. *Evolutionary Ecology Research* **10,** 575–598.

Chauhan, T., Lal, K. K., Mohindra, V., Singh, R. K., Punia, P., Gopalakrishnan, A., Sharma, P. & Lakra, W. S. (2007). Evaluating genetic differentiation in wild populations of the Indian major carp, *Cirrhinus mrigala* (Hamilton-Buchanan, 1882): evidence from allozyme and microsatellite markers. *Aquaculture* **269,** 135–149. doi: 10.1016/j.aquaculture.2007.05.007

Cornuet, J. M. & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144,** 2001–2014.

Courtenay, W. R. Jr., Williams, J. D., Britz, R., Yamamoto, M. N. & Loiselle, P. V. (2004). Identity of introduced snakeheads (Pisces, Channidae) in Hawaii and Madagascar, with comments on ecological concerns. *BISHOP Museum Occasional Papers* **77,** 1–13.

DeWoody, J. A. & Avise, J. C. (2000). Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology* **56,** 461–473. doi: 10.1111/j.1095-8649.2000.tb00748.x

Dewoody, J., Nason, J. D. & Hipkins, V. D. (2006). Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes* **6,** 951–957. doi: 10.1111/j.1471-8286.2006.01449.x

Esa, Y., Siraj, S. S., 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 Agricultural Science* **29,** 47–55.

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14,** 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

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

Falush, D., Stephens, M. & Pritchard, J. K. (2003). Inference of population structure: extensions to linked loci and correlated allele frequencies. *Genetics* **164,** 1567–1587.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39,** 783–791. doi: 10.2307/2408678

Fumagalli, L., Snoj, A., Jesenšek, D., Balloux, F., Jug, T., Duron, O., Brossier, F., Crivelli, A. J. & Berrerbi, P. (2002). Extreme genetic differentiation among the remnant populations of marble trout (*Salmo marmoratus*) in Slovenia. *Molecular Ecology* **11,** 2711–2716. doi: 10.1046/j.1365-294X.2002.01648.x

Glaubitz, J. C. (2004). CONVERT: a user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes* **4,** 309–310. doi: 10.1111/j.1471-8286.2004.00597.x

Goldstein, D. B., Ruiz Linares, A., Cavalli-Sforza, L. L. & Feldman, M. W. (1995). Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences of the United States of America* **92,** 6723–6727.

Goodman, S. J. (1997). RST CALC: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology* **6,** 881–885. doi: 10.1111/j.1365-294X.1997.tb00143.x

Goudet, J. (1995). FSTAT (version 1.2): a computer program to calculate *F*-statistics. *Journal of Heredity* **86,** 485–486.

Halls, A. S., Hoggarth, D. D. & Debnath, K. (1998). Impact of flood control schemes on river fish migrations and species assemblages in Bangladesh. *Journal of Fish Biology* **53**(Suppl. A), 358–380. doi: 10.1111/j.1095-8649.1998.tb01037.x

Huey, J. A., Baker, A. M. & Hughes, J. M. (2010). High levels of genetic structure in the Australian freshwater fish, *Ambassis macleayi*. *Journal of the North American Benthological Society* **29,** 1148–1160. doi: 10.1899/09-093.1

Jakobsson, M. & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23,** 1801–1806. doi: 10.1093/bioinformatics/btm233

Jamsari, A. F. J., Tan, M. P. & Siti Azizah, M. N. (2011). Isolation and multiplex genotyping of polymorphic microsatellite DNA markers in the snakehead murrel, *Channa striata*. *Genetics and Molecular Biology* **34,** 345–347. doi: 10.1590/S1415-47572011005000000

Kalinowski, S. T. (2005). HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5,** 187–189. doi: 10.1111/j.1471-8286.2004.00845.x

Kamarudin, K. R. & Esa, Y. (2009). Phylogeny and phylogeography of *Barbonymus schwanenfeldii* (Cyprinidae) from Malaysia inferred using partial cytochrome b mtDNA gene. *Journal of Tropical Biology and Conservation* **5,** 1–13.

Knight, J. T., Nock, C. J., Elphinstone, M. S. & Baverstock, P. R. (2009). Conservation implications of distinct genetic structuring in the endangered freshwater fish *Nannoperca oxleyana* (Percichthyidae). *Marine and Freshwater Research* **60,** 34–44.

Latter, B. D. H. (1972). Selection in finite populations with multiple alleles. III. Genetic divergence with centripetal selection and mutation. *Genetics* **70,** 475–490.

Lippé, C., Dumont, P. & Bernatchez, L. (2006). High genetic diversity and no inbreeding in the endangered copper redhorse, *Moxostoma hubbsi* (Catostomidae, Pisces): the positive sides of a long generation time. *Molecular Ecology* **15,** 1769–1780. doi: 10.1111/j.1365-294X.2006.02902.x

Liu, F., Xia, J. H., Bai, Z. Y., Fu, J. J., Li, J. L. & Yue, G. H. (2009). High genetic diversity and substantial population differentiation in grass carp (*Ctenopharyngodon idella*) revealed by microsatellite analysis. *Aquaculture* **297,** 51–56. doi: 10.1016/j.aquaculture.2009.09.008

Luikart, G., Allendorf, F. W., Cornuet, J. M. & Sherwin, W. B. (1998). Distortion of allele frequency distribution provides a test for recent population bottlenecks. *The Journal of Heredity* **89,** 238–247. doi: 10.1093/jhered/89.3.238

MacAvoy, E. S., McGibbon, L. M., Sainsbury, J. P., Lawrence, H., Wilson, C. A., Daugherty, C. H. & Chambers, G. K. (2007). Genetic variation in island populations of tuatara (*Sphenodon* spp) inferred from microsatellite markers. *Conservation Genetics* **8,** 305–318. doi: 10.1007/s10592-006-9170-5

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research* **27,** 209–220.

Mat Jais, A. M., McCulloch, R. & Croft, K. (1994). Fatty acid and amino acid composition in haruan as a potential role in wound healing. *General Pharmacology* **25,** 947–950.

Mohsin, A. K. M. & Ambak, M. A. (1983). *Freshwater Fishes of Peninsular Malaysia*. Penerbit: Universiti Pertanian.

Nei, M. (1972). Genetic distance between populations. *The American Naturalist* **106,** 283–291.

Nei, M., Tajima, F. & Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution* **19,** 153–170. doi: 10.1007/BF02300753

Neville, H., Isaak, D., Thurow, R., Dunham, J. & Rieman, B. (2007). Microsatellite variation reveals weak genetic structure and retention of genetic variability in threatened Chinook salmon (*Oncorhynchus tshawytscha*) within a Snake River watershed. *Conservative Genetics* **8,** 133–147. doi: 10.1007/s10592-006-9155-4

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

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4,** 137–138. doi: 10.1046/j.1471-8286.2003.00566.x

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145,** 1219–1228.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4,** 406–425.

Schuster, W. H. (1952). A provisional survey of the introduction and transplantation of fish throughout the Indo-Pacific region. In *Proceedings of the Third IPFC Fisheries Symposium* (Shuster, W. H., ed), pp. 187–196. Madras: FAO.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47,** 264–279. doi: 10.2307/2410134

Slatkin, M. & Excoffier, L. (1996). Testing for linkage disequilibrium in genotypic data using the expectation-maximization algorithm. *Heredity* **76,** 377–383.

Takezaki, N. & Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144,** 389–399.

Takezaki, N. & Nei, M. (2008). Empirical tests of the reliability of phylogenetic trees constructed with microsatellite DNA. *Genetics* **178,** 385–392. doi: 10.1534/genetics.107.081505

Takezaki, N., Nei, M. & Tamura, K. (2010). POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with windows interface. *Molecular Biology and Evolution* **27,** 747–752. doi: 10.1093/molbev/msp312

Tan, M. P., Jamsari, A. F. J. & Siti Azizah, M. N. (2012). Phylogeography pattern of the striped snakehead, *Channa striata* in Sundaland: ancient river connectivity, geographical and anthropogenic signatures. *PLoS ONE* **7,** e52089. doi: 10.1371/journal.pone.0052089

Valles-Jimenez, R., Cruz, P. & Perez-Enrique, R. (2004). Population genetic structure of Pacific white shrimp (*Litopenaeus vannamei*) from Mexico to Panama: microsatellite DNA variation. *Marine Biotechnology* **6,** 475–484. doi: 10.1007/s10126-004-3138-6

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. F. (2004). MICRO-CHECKER: for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4,** 535–538. doi: 10.1111/j.1471-8286.2004.00684.x

Voris, H. K. (2000). Map of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time duration. *Journal of Biogeography* **27,** 1153–1167. doi: 10.1046/j.1365-2699.2000.00489.x

Weir, B. S. & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38,** 1358–1370. doi: 10.2307/2408641

Yusoff, F. M., Shariff, M. & Gopinath, N. (2006). Diversity of Malaysian aquatic ecosystems and resources. *Aquatic Ecosystem Health & Management* **9,** 119–135. doi: 10.1080/14634980600713315

Electronic References

Felsenstein, J. (2005). *PHYLIP (Phylogeny Inference Package) Version 3.6*. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle. Available at http://evolution.genetics.washington.edu/phylip.html/

Froese, R. & Pauly, D. (Eds) (2010). *FishBase. World Wide Web electronic publication.* Available at www.fishbase.org (last accessed 16 September 2015).

Inger, R. F. & Chin, P. K. (1962). *The Fresh-Water Fishes of North Borneo. Fieldiana*. Chicago, IL: Chicago Natural History Museum. Available at https://archive.org/details/freshwaterfishes45inge/ (last accessed 16 September 2015).

Mazuki, H. (2008). *National Aquaculture Sector Overview Fact Sheets*. Rome: FAO. Available at http://www.fao.org/fishery/countrysector/naso_malaysia/en/ (last accessed 15 January 2011).