CHAPTER 4

RESULTS

4.1. DNA Extraction

The concentration of genomic DNA extracted via GF-1 Tissue Extraction Kit (Vivantis, Malaysia) was in the range of 0.010 – 0.035 µg/µL and has an A_{260}/A_{280} ratio between 1.7 and 1.9. Successful PCR amplifications were obtained for all of the DNA samples.

4.2. Microsatellite markers characterisation

A total of fifty-five pairs of microsatellite primers were optimised in this study. The primers were designed to flank perfect microsatellites with the exception of CS-6 (Salah, 2010; Adamson, 2010). The repeat motifs for these primers vary from di- to hexa- repeats. After subjecting the optimised primers to preliminary polymorphism screening, only eight markers showed promising results based on PCR amplification success and were subsequently retained for further analyses. The selected primer sets were T113-11, BP6-2, BP6-4, PCT6-6, BP13-6, BP13-14, CS-4 and CS-5. Figure 4.1 summarises the microsatellite repeat types isolated from the eight chosen loci.

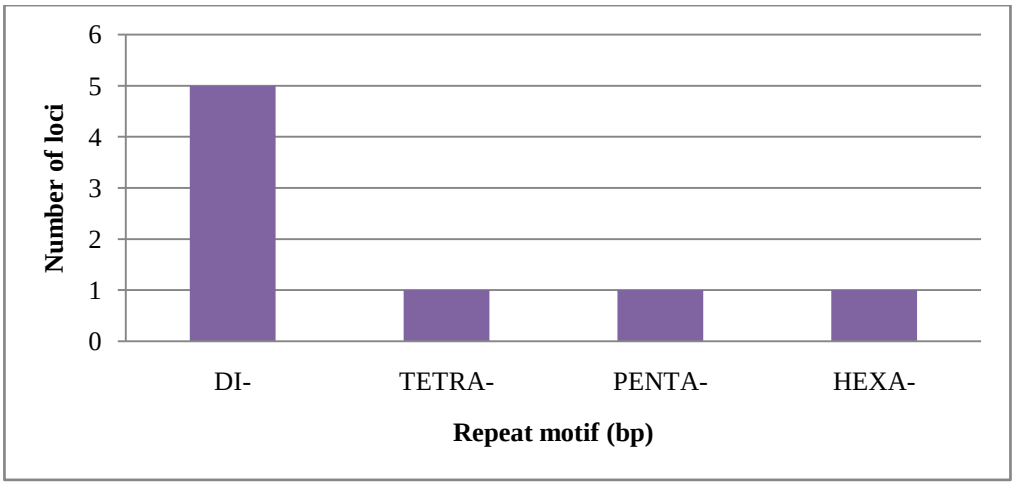


Figure 4.1. Type of microsatellite repeats contained within the eight potentially polymorphic loci of *C. striata* employed in this study.

The degree of polymorphism of a microsatellite marker is determined by the level of variation of banding patterns. For a polymorphic marker, the number of genotypes that can be screened by the marker is a requisite. The alleles are designated according to PCR product size relative to the molecular size marker. Hence, a range of banding patterns with assorted sizes were viewed on a stained electrophoresed gel. Figure 4.2 shows the initial amplification patterns obtained with the putatively polymorphic tri-nucleotide locus (BP13-14 primer set) which was assayed on randomly selected individuals from the study populations. Note the variation in PCR product sizes and the multiple products in lanes 2, 6, 10, 11, 14, and 15, suggesting that the primers are amplifying multi-allelic genotypes that contain sequence length variants.

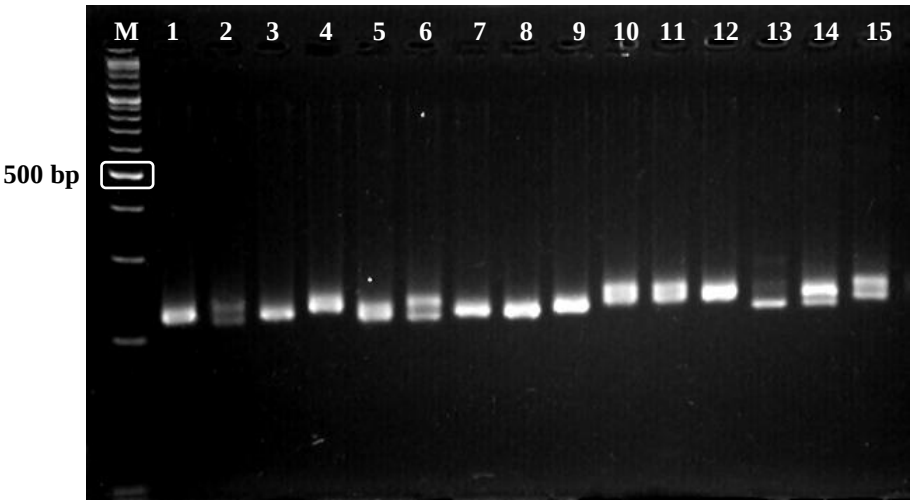
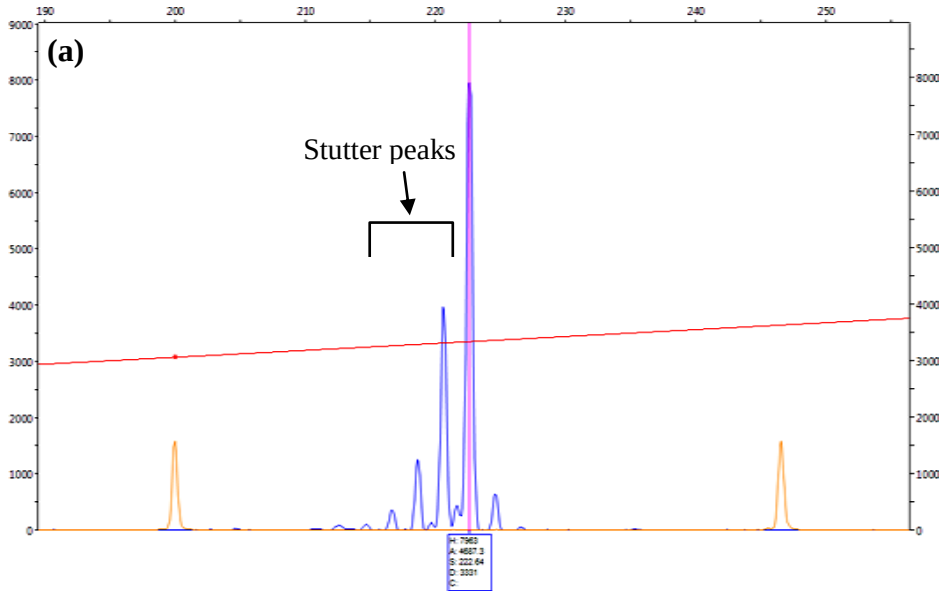


Figure 4.2. Preliminary screening for polymorphism using microsatellite primer BP13-14 on *C. striata* at 60 °C. M, 100 bp DNA size marker (Vivantis); lanes 1-3, three individuals from Johore; 4-6, three individuals from Kedah; 7-9, three individuals from Pahang; 10-12, three individuals from Malacca; 13-15, three individuals from Rawang.

To further elucidate the genetic variability as well as polymorphic competency of the identified eight loci, fragment sizes of PCR products were analysed with the GeneScan™-250 LIZ™ Size Standard via fragment analyser. Figures 4.3 (a) and (b) illustrate size variation in PCR products amplified using the BP13-14 primer set.



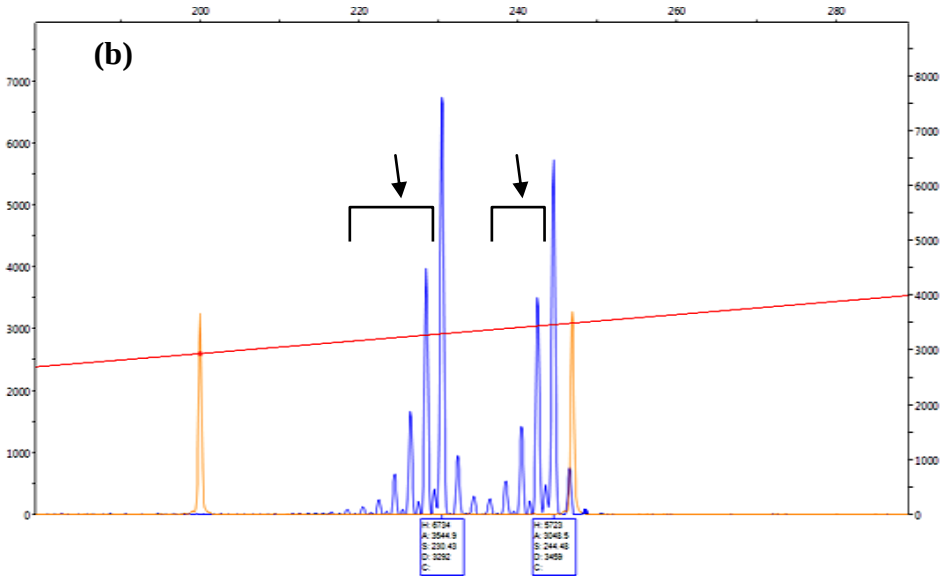


Figure 4.3. Electropherograms of individuals amplified using primer BP13-14. The x-axis and y-axis corresponds to the size of fragments (bp) and fluorescence intensity of each peak respectively. The yellow peaks represent the GeneScan™-250 LIZ™ Size Standard and the pink bar indicates an offscale fluorescence peak (exceeding maximum detectable range). Stutter peaks are indicated by black arrows. (a) Result showing a homozygote genotype (allele size 222bp) amplified from individual J10 from Johore; (b) Electropherogram of a heterozygote genotype (allele sizes; 230bps and 244bp) observed in individual KJ40 from Kajang.

4.3. Statistical results

A total of 239 individuals collected from 6 locations were genotyped at eight microsatellite loci. However, an aggregate of three PCR amplifications were unsuccessful, resulting in a final data set with ~0.0157% missing data. Allele frequencies for each locus are presented in Appendix 3.

4.3.1. Deviations from Hardy-Weinberg Equilibrium (HWE) and linkage disequilibrium (LD)

Among the forty-eight tests (eight loci, six populations) for Hardy-Weinberg equilibrium (HWE) (Table 4.0), significant departures were observed after False Discovery Rate (FDR) correction in all the cultured populations, with Malacca exhibiting HWE deviations at all eight loci. Of the wild populations, Kedah did not conform to HWE at five (PCT6-6, BP13-6, CS-5, BP6-2 and BP6-4) of the eight loci tested, and the Johore population showed deviation at only one locus (CS-5). Of the eight loci assessed, two (CS-5 and BP6-4) showed significant non-conformance to HWE in five populations whilst another two (BP13-14 and CS-4) revealed deviation in only the cultured populations (Kajang, Malacca and Rawang).

Table 4.0. Significance values for HWE tests assessed for each population at each locus. * denotes significance (α=0.05) of test before FDR correction, and bold values indicate significant departures after FDR correction.

Population /Locus	Johore	Kedah	Pahang	Kajang	Malacca	Rawang
PCT6-6	0.12119	**0.00088**	0.06693	**0.00663**	**0.00031**	**0**
BP13-6	1	**0.01166**	0.57492	0.41558	**0**	0.26022
BP13-14	0.04008*****	0.06687	0.08478	**0**	**0**	**0**
CS-4	0.03546*****	1	0.08518	**0**	**0**	**0**
CS-5	**0.00062**	**0**	1	**0.01184**	**0**	**0**
T113-11	0.05169	0.18771	**0.01743**	0.55336	**0.00074**	0.15552
BP6-2	0.52608	**0.00046**	**0.00328**	0.33198	**0.00129**	0.25259
BP6-4	0.2765	**0.02066**	**0**	**0**	**0**	**0**

After FDR adjustment, significant linkage disequilibrium (LD) was observed in 26 of the 168 pair-wise comparsions between individual loci at each site (15.48% of tests performed) (data not shown). Loci BP6-2 and BP6-4 displayed linkage associations consistently in four populations (Kedah, Pahang, Malacca and Rawang). In view of the fact that subsequent analyses involves the assumption that the loci being assayed are independent samples of the genome, one locus in the pair was discarded from further analysis. Since the locus BP6-4 showed consistently greater (Kedah, Pahang, Kajang, Malacca and Rawang) HWE deviations among populations compared to BP6-2 (Kedah, Pahang and Malacca), BP6-4 was excluded from data set in order to avoid increased Type 1 error in the succesive assesments (Selkoe and Toonen, 2006). Hence, from this point on, only data concerning the following seven markers; PCT6-6, BP13-6, BP13-14, CS-4, CS-5, T113-11 and BP6-2 was used in the analysis.

4.3.2. Genetic variation within populations and population bottlenecks

Table 4.1 summarises genetic variability for each population at each locus. All loci were variable (100% polymorphic) in all populations, and the levels of genetic diversity varied relative to the locus. The number of alleles per locus in the 6 populations ranged from 28 (BP13-6) to 53 (BP13-14). Table 4.2 lists the PIC values of the seven microsatellite markers used. Co-dominant markers exhibiting PIC values higher than 0.5 are regarded as highly informative (Botstein et al., 1980). Across the six populations of *C. striata*, the most informative marker was BP13-14 (0.869) whereas the least informative marker was T113-11 (0.481), the only one below the level of 0.5.

Private alleles were observed at all seven loci and accounted for 24 of the 264 individual alleles (~9.09%) present at all loci among all six populations. In general, the cultured populations, Rawang, Malacca and Kajang showed considerably greater numbers of site-specific alleles (9, 7 and 5 private alleles respectively) than the wild populations Johore, Pahang and Kedah with 2, 1 and 0 private alleles respectively. The two microsatellite loci with the highest PIC values, BP13-14 and CS-5, each possessed 6 private alleles among all the population samples.

Table 4.1. Genetic diversity statistics at 7 microsatellite loci assessed for the wild and cultured populations of *C. striata*, Pop. = population name, n = number of individuals genotyped, n_a = observed number of alleles, n_e = effective number of alleles, allelic range = number of repeat units that differ across alleles, A_R = allelic richness, H_o = observed heterozygosity, H_e = expected heterozygosity, $\overline{H}$ = averaged heterozygosity across all 7 loci, $\overline{F}_{IS}$ = population specific inbreeding coefficient indices, G/W Index = Garza – Williamson Index across all loci. Significant $\overline{F}_{IS}$ ($p<0.05$) values following FDR procedure are indicated with (ɵ). For the G/W Index, values less than 0.68 (threshold for recent population decline) are denoted by (†).

Pop.	Locus:	PCT6-6	BP13-6	BP13-14	CS-4	CS-5	T113-11	BP6-2
Johore (Wild)	n	30	30	30	30	29	30	30
	n_a	5	3	9	4	7	3	3
	n_e	3.15	1.39	2.38	1.15	1.70	1.55	2.76
	allelic range	14	4	9	8	14	2	2
	A_R	5	2.97	8.97	3.93	7	3	3
	H_o	0.87	0.33	0.50	0.07	0.28	0.37	0.67
	H_e	0.69	0.29	0.59	0.13	0.42	0.36	0.65
	$\overline{H}$	0.44	(SD: 0.26)		$\overline{F}_{IS}$	0.012		
	G/W Index	0.68	(SD: 0.26)					
Kedah (Wild)	n	30	30	29	30	30	30	30
	n_a	5	3	6	3	4	3	3
	n_e	2.43	2.10	2.48	1.31	2.17	2.20	2.71
	allelic range	9	4	10	3	12	2	2
	A_R	4.97	3	6	2.97	3.97	3	3
	H_o	0.87	0.40	0.45	0.27	0.23	0.40	0.73
	H_e	0.60	0.53	0.61	0.24	0.55	0.56	0.64
	$\overline{H}$	0.48	(SD: 0.24)		$\overline{F}_{IS}$	0.096		
	G/W Index	0.67†	(SD: 0.24)					
Pahang (Wild)	n	30	30	30	30	30	30	30
	n_a	8	3	5	5	3	3	3
	n_e	2.85	1.74	2.52	1.98	1.07	2.53	1.82
	allelic range	11	3	8	16	5	2	2
	A_R	7.90	3	4.97	4.97	2.93	3	3
	H_o	0.90	0.53	0.73	0.50	0.07	0.37	0.47
	H_e	0.66	0.43	0.61	0.50	0.07	0.61	0.46
	$\overline{H}$	0.51	(SD: 0.27)		$\overline{F}_{IS}$	-0.066		
	G/W Index	0.68	(SD: 0.24)					
Kajang (Cultured)	n	49	49	49	49	49	49	49
	n_a	5	6	9	11	10	7	6
	n_e	3.15	1.83	4.47	5.88	4.55	1.48	1.81
	allelic range	10	6	17	16	18	7	15
	A_R	4.59	5.43	8.71	9.59	8.20	5.94	5.26
	H_o	0.88	0.47	0.92	0.82	0.76	0.33	0.45
	H_e	0.69	0.46	0.78	0.84	0.79	0.33	0.45
	$\overline{H}$	0.66	(SD: 0.24)		$\overline{F}_{IS}$	-0.063		
	G/W Index	0.61†	(SD: 0.18)					

Table 4.1. continued.

Pop.	Locus:	PCT6-6	BP13-6	BP13-14	CS-4	CS-5	T113-11	BP6-2
Malacca (Cultured)	n	50	50	50	50	50	50	50
	n_a	3	6	8	13	13	6	5
	n_e	2.04	4.22	2.92	6.09	4.08	2.30	1.75
	allelic range	8	6	16	15	20	8	13
	A_R	2.99	5.97	7.46	11.64	9.61	4.92	4.68
	H_o	0.72	0.84	0.72	0.70	0.74	0.78	0.24
	H_e	0.51	0.77	0.66	0.84	0.76	0.57	0.43
	$\overline{H}$	0.68	(SD: 0.20)		$\overline{F}_{IS}$	-0.040		
	G/W Index	0.59†	(SD: 0.19)					
Rawang (Cultured)	n	50	50	50	50	49	50	50
	n_a	5	7	16	11	13	7	6
	n_e	3.16	3.29	7.68	7.80	4.61	2.00	2.13
	allelic range	8	6	20	13	17	7	5
	A_R	4.51	6.54	13.47	10.06	10.90	6.32	4.96
	H_o	0.98	0.78	0.68	0.62	0.61	0.42	0.48
	H_e	0.69	0.70	0.88	0.88	0.79	0.50	0.53
	$\overline{H}$	0.65	(SD: 0.19)		$\overline{F}_{IS}$	0.080ɵ		
	G/W Index	0.81	(SD: 0.15)					

Table 4.2. Estimates of the presence and frequency of null alleles, and PIC values for each locus.

Locus	Null Present	Oosterhout	Chakraborty	Brookfield 1	Brookfield 2	PIC
PCT6-6	no	-0.1662	-0.1187	-0.1093	0	0.621
BP13-6	no	-0.1801	-0.0821	-0.0394	0	0.569
BP13-14	no	0.0879	0.0741	0.0506	0.0506	0.869
CS-4	yes	0.1322	0.3123	0.0537	0.0537	0.760
CS-5	yes	0.1845	0.1972	0.096	0.1966	0.807
T113-11	no	-0.0348	-0.0146	-0.0078	0	0.481
BP6-2	no	-0.0353	-0.0221	-0.0176	0	0.529
BP6-4	no	-0.0136	-0.0066	-0.0057	0	0.876

All parameters for diversity revealed greater genetic variation in the cultured than wild populations. The average number of alleles per locus (A) for cultured populations ranged from 7.71 to 9.29, a stark contrast to wild populations that ranged from 3.86 to 4.86. Mean effective number of alleles per locus (n_e), ranged from 3.31 to 4.38 in the cultured populations, whilst wild populations exhibited lower values, ranging from 2.01 to 2.20. Across the populations under study, mean allelic richness, a sample size bias

corrected estimator of the number of alleles per locus, ranged from 3.84 (Kedah) to 8.11 (Rawang). A graphical illustration of the mean allelic richness across populations (Figure 4.4) clearly displays the differences between the cultured and wild populations, whereby markedly lower values were observed in the wild populations.

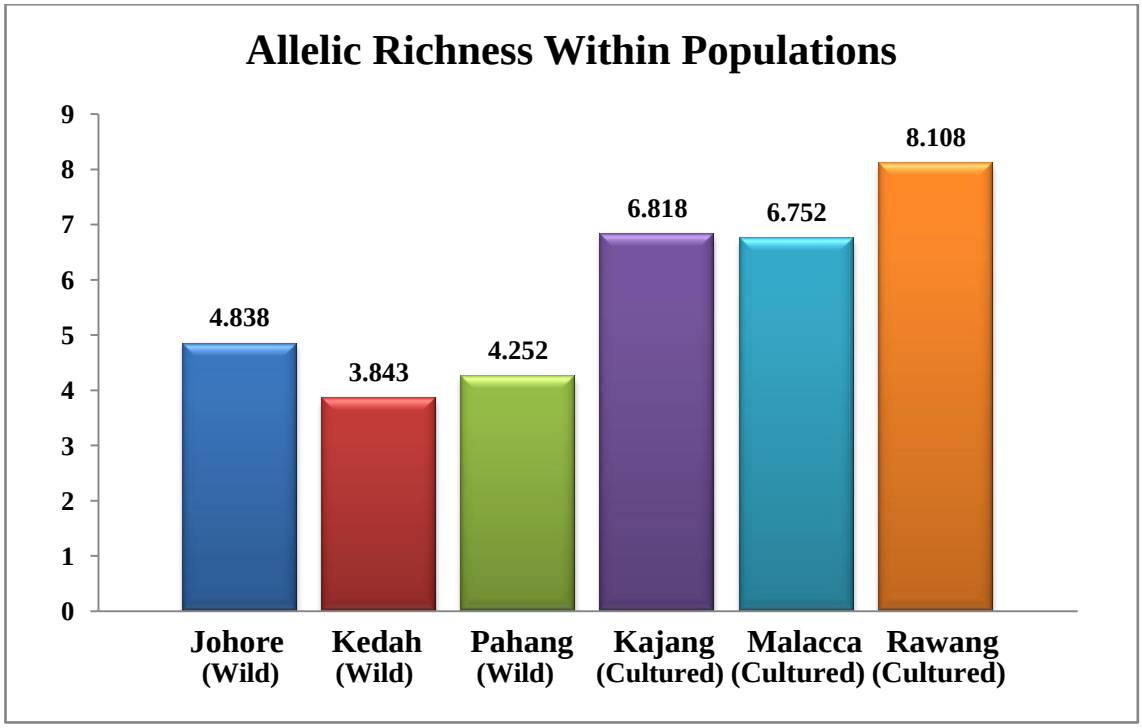


Figure 4.4. Bar chart representing the mean allelic richness values among the cultured and wild *C. striata* populations.

Likewise, a marginally greater degree of diversity as assessed by heterozygosity was observed in cultured compared to wild populations. The observed heterozygosity (H_o) ranged from 0.24 to 0.98 in the cultured populations while in wild populations, H_o ranged from 0.07 to 0.90. The expected heterozygosity (H_e) values for the cultured populations ranged from 0.33 to 0.88 whereas the wild populations exhibited a range of H_e from 0.07 to 0.69. Within the two pooled populations, the range of averaged heterozygosity ($\bar{H}$) was consistently higher in cultured (0.65 – 0.68) than in wild

(0.44 – 0.51) populations. However, several loci (CS-4 and CS-5) showed evidence for the presence of null alleles (Table 4.2) as indicated by homozygote excesses (heterozygote deficiencies) (Table 4.1), whereby consistently lower estimates of H_o were observed compared to H_e in a majority of the cultured and wild populations.

After FDR correction, population specific inbreeding coefficient ($\overline{F}_{IS}$) values estimated (Table 4.1) were marginally low in most populations with the exception of Rawang, which showed significant but moderate $\overline{F}_{IS}$ value of 0.08. One wild and two cultured populations; Kedah, Kajang and Malacca showed evidence for recent reductions in population sizes as indicated by the *M*-statistic (G/W Index) values of 0.67, 0.61 and 0.59 respectively, which are lower than the threshold bottleneck value of 0.68 (Garza and Williamson, 2001).

4.3.3. Genetic diversity among populations

Results of the hierarchical analysis of molecular variance (AMOVA) within and among the six populations, which were categorised *a priori* into two groups, cultured and wild, are tabulated in Table 4.3. The AMOVA analysis revealed significant, although minimal differentiation between pooled cultured and wild populations (F_{CT} = 0.083, p = 0). Similarly, the variation among individuals within populations was substantial albeit moderately valued at F_{IS} = 0.003 (p = 0). Within individual variation accounted for 79.2% of the total genetic variation whilst only 8.32% was partitioned in the variation among the cultured and wild groups.

Table 4.3. Analysis of molecular variance (AMOVA) of pooled cultured (Kajang, Malacca and Rawang) and pooled wild (Johore, Kedah and Pahang) populations of *C. striata* based on 7 microsatellite loci. ** indicate significant departures after FDR correction.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	Fixation indices	P-value
Among groups	1	74.885	0.218	8.32	0.083 (F_{CT})	0.000**
Among populations within groups	4	110.358	0.320	12.24	0.134 (F_{SC})	0.458
Among individuals within populations	233	485.395	0.006	0.23	0.003 (F_{IS})	0.000**
Within individuals	239	495	2.071	79.2	0.208 (F_{IT})	0.101
Total	477	1165.638	2.615			

4.3.4. Population structure

By and large, the pair-wise measures (F_{ST}, R_{ST} and D_{est}) employed to assess differentiation between populations were congruent and indicated low to extensive genetic differentiation among pair-wise population comparisons after FDR correction. Pair-wise F_{ST} values (Table 4.4) for all combinations of populations demonstrated significant differentiation that ranged from 0.028 to 0.284 whilst pair-wise R_{ST} analysis (Table 4.5) yielded significant values ranging from 0.023 to 0.340. Only one pair-wise R_{ST} comparison was not significantly differentiated, corresponding to low genetic differentiation (0.034) between Malacca and Rawang. In both instances, results from population pair-wise F_{ST} and R_{ST} estimates showed that the largest genetic difference was uncovered between the wild populations Kedah and Pahang, with significant values of 0.284 and 0.340 respectively, whereas the smallest differentiation was observed between the cultured populations Kajang and Rawang with the pair-wise F_{ST} and R_{ST} values of 0.028 and 0.023 respectively. The D_{est} values of genetic differentiation among

populations (Table 4.6) ranged from 0.036 to 0.297, whereby the highest differentiation estimate value was observed between wild population, Kedah and cultured population, Kajang. Nonetheless, D_{est} values showed that the pair-wise differentiation between Kajang and Rawang was again the least of any comparison, concurring with the results generated by F_{ST} and R_{ST} measures. The Mantel test for correlation of parameters, in this case, a series of successive pair-wise comparisons between three genetic differentiation estimators, revealed strong and significantly positive correlation between the data matrices of F_{ST} and R_{ST} (0.838) as well as F_{ST} and D_{est} (0.717). On the contrary, a substantially lower correlation was observed between pair-wise matrices of R_{ST} and D_{est} (correlation coefficient, r = 0.493; p = 0.013).

Table 4.4. Matrix of pair-wise F_{ST} values (above diagonal) and the corresponding significance (below diagonal) among six populations of *C. striata*. ɵ denotes significance (α = 0.05) following FDR adjustment.

	Johore	Kedah	Pahang	Kajang	Malacca	Rawang
Johore		0.236	0.095	0.238	0.235	0.194
Kedah	ɵ		0.284	0.228	0.171	0.164
Pahang	ɵ	ɵ		0.220	0.263	0.187
Kajang	ɵ	ɵ	ɵ		0.146	0.028
Malacca	ɵ	ɵ	ɵ	ɵ		0.098
Rawang	ɵ	ɵ	ɵ	ɵ	ɵ	

Table 4.5. Matrix of pair-wise R_{ST} values (above diagonal) and the corresponding significance (below diagonal) among six populations of *C. striata*. ɵ denotes significance (α = 0.05) following FDR adjustment. Gray cells indicate comparisons that are not statistically significant.

	Johore	Kedah	Pahang	Kajang	Malacca	Rawang
Johore		0.141	0.083	0.208	0.213	0.134
Kedah	ɵ		0.340	0.163	0.094	0.110
Pahang	ɵ	ɵ		0.297	0.333	0.240
Kajang	ɵ	ɵ	ɵ		0.064	0.023
Malacca	ɵ	ɵ	ɵ	ɵ		0.034
Rawang	ɵ	ɵ	ɵ	ɵ		

Table 4.6. Harmonic mean of D_{est} across loci among populations of *C. striata*.

	Johore	Kedah	Pahang	Kajang	Malacca	Rawang
Johore	--	0.137	0.052	0.212	0.238	0.203
Kedah	--	--	0.197	0.297	0.186	0.209
Pahang	--	--	--	0.175	0.277	0.184
Kajang	--	--	--	--	0.205	0.036
Malacca	--	--	--	--	--	0.176
Rawang	--	--	--	--	--	--

Table 4.7. Mantel tests for correlation of three pair-wise genetic differentiation measures (correlation coefficient, r : above diagonal) and the corresponding significance (below diagonal). ɵ denotes significance (α = 0.05) following FDR adjustment.

	F_{ST}	R_{ST}	D_{est}
F_{ST}		0.838	0.717
R_{ST}	ɵ		0.493
D_{est}	ɵ	ɵ	

4.3.5. Cluster analysis and relationship trees

The log probabilities [Ln P(D)] associated with different numbers of genetic clusters (*K*), calculated from Bayesian clustering analysis of 239 individuals showed the highest value at $K = 5$ [Ln P(D) = -4240.1], and the lowest value at $K = 1$ [Ln P(D) = -5409.9] as shown in Figure 4.5 (a). Hence, this indicates that the most appropriate number of genetic groups assigned for the given data set was five [Figure 4.5 (b)].

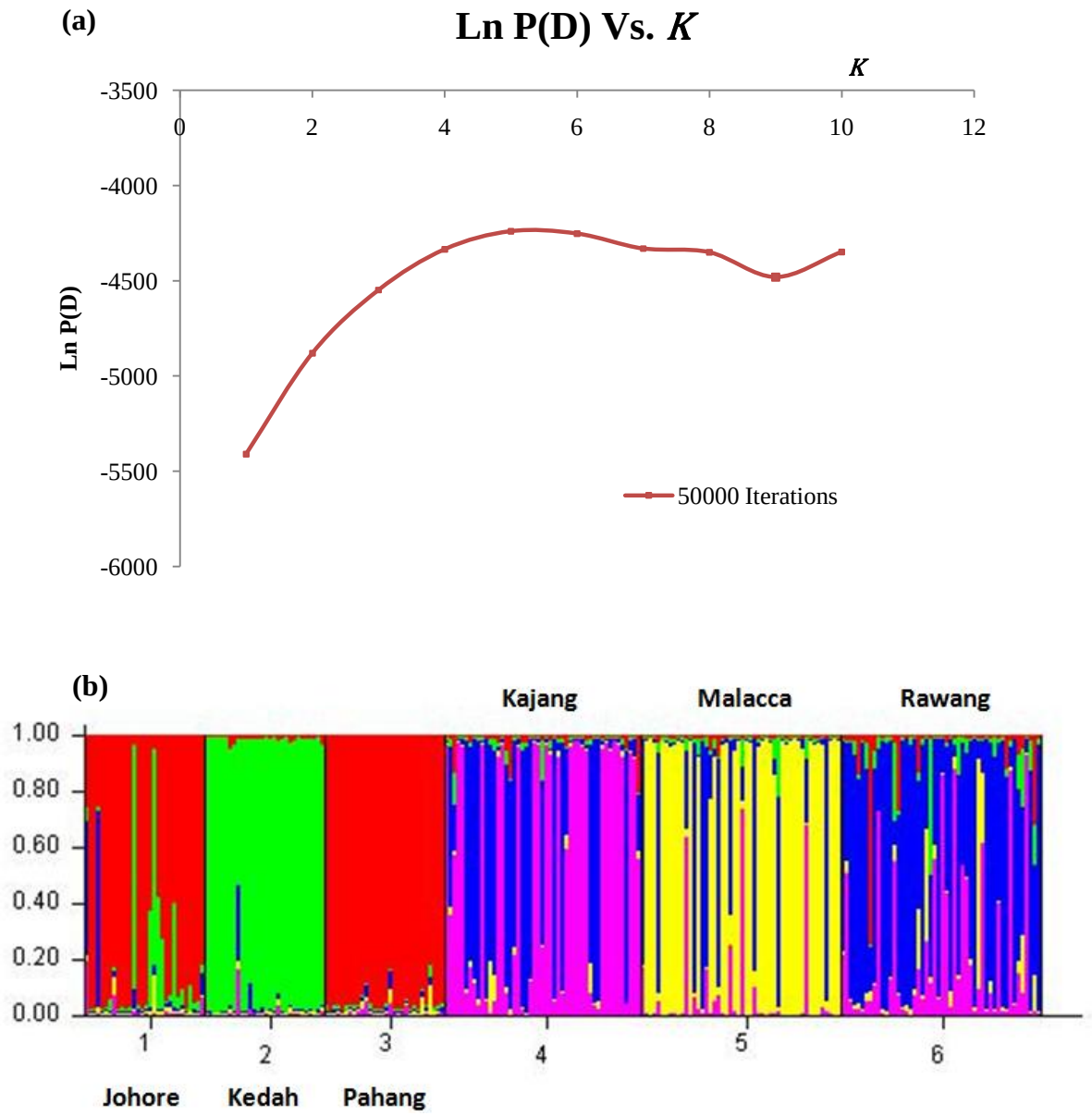


Figure 4.5. (a) Results of Bayesian Cluster analysis of log-likelihood ratios [Ln P(D)] for different *K* values. The peak of Ln P(D) identifies the best number of clusters, in this study five. (b) Results of the cluster analysis generated by the STRUCTURE software. Different colours indicate different genetic clusters. Each column represents an individual whereby the height of the column segments shows the probability of assignment of the individual *C. striata* to the genetic clusters.

For the phylogenetic analysis in this study, Neighbour-Joining (NJ) tree and Unweighted Pair-Group Method of Arithmetic Averages (UPGMA) dendogram were constructed to illustrate the genetic relationships among the cultured and wild populations of *C. striata*. The former (Figure 4.6) was generated on the basis of pair-wise F_{ST} genetic differentiation estimates whereas the latter (Figure 4.7) was based on Nei's 1978 standard genetic distances. Both phylogenetic trees concurred that the six populations were allocated into three major branches. In each figure the wild populations Johore and Pahang cluster together whereas the cultured populations of Kajang and Rawang formed another cluster, reflecting similarity that was observed in F_{ST}, R_{ST} and D_{est} analyses. The three cultured populations group together in the UPGMA dendrogram (albeit with low support), suggesting that some genetic similarity may be present among the cultured samples. This relationship, however, is not evident in the NJ tree, where the three cultured populations clade together with the wild Kedah population.

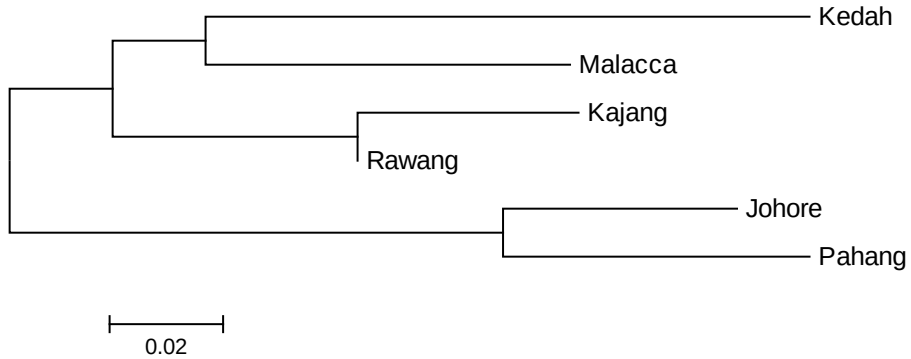


Figure 4.6. The Neighbour-Joining (NJ) tree based on pair-wise F_{ST} estimates among six cultured and wild populations of *C. striata*. Scale bar of F_{ST} value is shown under the tree.

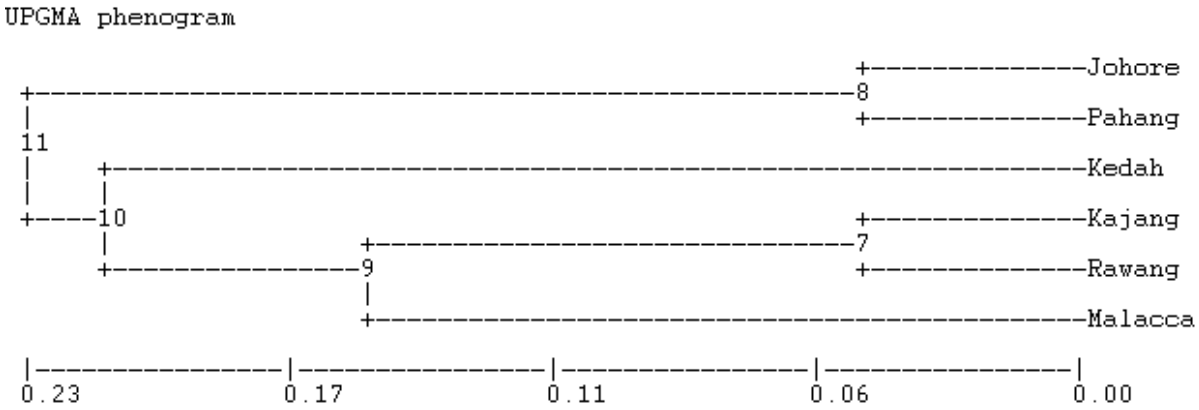


Figure 4.7. Dendogram generated by UPGMA analysis based on Nei's 1978 standard genetic distance displaying phylogenetic relationship among cultured and wild populations of *C. striata*.

CHAPTER 5

DISCUSSION

The Malaysian consumer market has experienced renewed interest in *Channa striata* due to recently discovered therapeutic qualities (Gam et al., 2005; Mat Jais, 2007a; Zuraini et al., 2006) and also due to the nutritional value of the species as an economical protein source. This rapidly expanding market has in turn led to the augmentation of wild fishery catch with farmed *C. striata* produced in commercial aquaculture ventures. However, the consequences of Malaysian farming practices on the genetic diversity of *C. striata* remains largely unexamined. Thus, the present study employed seven polymorphic microsatellite markers to quantify and compare levels of genetic diversity in cultured and wild populations (stocks) within Peninsular Malaysia, and to determine the presence and extent of genetic differentiation present among these study populations.

5.1. Departures from Hardy-Weinberg Equilibrium (HWE)

Microsatellite analysis showed significant deviations from Hardy-Weinberg equilibrium (HWE) in all three cultured populations (Kajang, Malacca and Rawang) along with a wild population (Kedah) in the current investigation. Non-conformance to HWE is a common occurrence in cultured lines where a number of perturbing factors such as small effective population size, founder effects, inbreeding, intentional and/or unintentional selection during domestication process which includes non-random mating, may have acted in discordance with the principle. Interestingly, the population

from Kedah is the only wild population to exhibit departure from HWE. The population is most likely to have experience one or a combination of events involving genetic bottleneck, random genetic drift, mutation or natural selection in the wild.

It is crucial that microsatellite markers are subjected to critical and explicit evaluation as loci that display gross violations of basic assumptions underlying the data analyses, or high error rates may lead to inaccurate and biased genetic estimates (Selkoe and Toonen, 2006). Loci-wise; PCT6-6, CS-5 and BP6-4 showed significant HWE deviations at majority of studied populations following FDR corrections. In addition, Microchecker had identified presence of null alleles at variable microsatellite loci CS-4 and CS-5. This might influence a locus to deviate from HWE. Null alleles may result in an excess of homozygotes when null allele heterozygotes are erroneously scored as homozygotes for the variant allele. In this case of multilocus analyses as employed in this study, the effects of null alleles are normally confined to the locus. Prior to this, Microchecker analysis did not identify any large allele size dropouts or scoring error to the rest of the microsatellite data other than the null alleles. An interesting observation was noted in loci BP13-14 and CS-4 which only detected non-conformity to HWE in all three cultured populations and none of that in the wild.

5.2. Genetic diversity within *C. striata* populations

Genetic diversity is a prerequisite in successful aquaculture programmes where information on the genetic relationships among cultured and wild populations are central to the development of sustainable breeding strategies for a competitively commercialised culture. Genetic variation of *C. striata* in this study was characterised by moderate allele diversity (6.29 alleles/population per locus) compared to the averages for freshwater fish (A = 9.1 ± 6.6 averaged across 13 species, De Woody and Avise, 2000). However, the mean heterozygosity (0.57) observed in this microsatellite analysis were marginally higher than the unrelated *C. striata* population investigation of Kedah (0.52) (Jamaluddin et al., 2011) and markedly inflated compared to the populations in Thailand (0.039) (Hara et al., 1998).

5.2.1. Comparison between cultured and wild populations

In the present study, cultured populations demonstrated substantially greater genetic diversity than wild populations of *C. striata*. Analysis showed that cultured population of *C. striata* contained a mean of 8.24 alleles/population per locus, nearly double the amount found in the wild populations at 4.33 alleles/population per locus. This is somewhat surprising as it restructures the notion that cultured populations will almost be certainly accompanied by genetic variation depression as observed in number of cultured species [black tiger shrimp (Xu et al., 2001), Atlantic cod (Pampoulie et al., 2006), Japanese flounder (Sekino et al., 2002), silver-lipped pearl oysters (Lind et al., 2009)]. However, there have been reports on the lack of genetic differences found between cultured and wild populations in aquaculture species such as giant freshwater

prawn (Chareontawee et al., 2007), mud carp (Yang et al., 2008) and striped catfish (Ha et al., 2009).

Current investigation has revealed that the measure of genetic diversity based on heterozygosity (H_o, H_e, and $\bar{H}$) were all significantly elevated in the cultured populations of *C. striata* relative to the wild. The highest averaged heterozygosity among all populations was observed in Malacca (cultured) at 0.68 ± 0.20 while Johore (wild) showed the least heterogeneity at 0.44 ± 0.26. The relatively high H_e values of cultured populations implied that the growth and progression of the current population relied on a large number of breeding individuals (effective population size, N_e). In addition, high values H_o observed suggests that populations were founded with ample N_e. In contrast, heterozygote deficit were detected in the wild population of Kedah, and cultured population of Rawang at major loci tested. The absence of homozygote excess in the other two wild population (Johore and Pahang) affirms that the samples collected were good representatives of the wild stocks.

Nevertheless, heterozygosity in particular, does not adequately represent the range of genetic variation (Beardmore et al., 1997) and, hence allele-related diversity statistics (n_a, n_e, allelic range and A_R) were included in the analyses to gain a thorough comprehension of the level of variation among populations. Results from the allele-based diversity indices mirrors the findings discovered in the initial measure of variation (heterozygosity) with the exception of Malacca, whereby Rawang instead, was the population possessing the greatest genetic diversity with A_R value at 8.108. Beardmore and colleagues' finding's (1997) were re-asserted in the present study with an interesting observation of Malacca is more diverse (mean heterozygosity-wise) at 0.68 ± 0.20 compared to its cultured counterpart Rawang (0.65 ± 0.19), although

Rawang has a much greater number of alleles and A_R value (65, 8.108) than that seen in Malacca (54, 6.752).

A comparison of private alleles (24 alleles) among pooled wild and cultured populations of *C. striata* showed that most of these alleles belonged to the cultured populations (21), with only a fraction (3) of them found in the wild populations. Site-specific (private) alleles method is an oblique estimator of average migration rates (Nm) (Slatkin, 1985). Gene flow is vital to a population because low gene flow can lead to local inbreeding and in turn, inbreeding depression whereas high or increased gene flow can limit local adaptation and cause outbreeding depression. A linear relationship exists between Nm and average frequency of private alleles. If Nm is low, numerous private alleles arises through mutation, of which new alleles will remain site-specific depending on the migration rates. The higher the Nm, lower the proportion of private alleles. If Nm is high, site-specific alleles will be common (Allendorf and Luikart, 2007). Care should be taken when interpreting results as the reliability of the private allele method have not been thoroughly investigated for loci with potential homoplasmy (microsatellites under the stepwise mutation model). Although in this case, the increase in private alleles in cultured populations is probably due to artificial admixing of the broodstocks used by the farmers, resulting in the multitude of diversity observed in the study.

5.3. Inbreeding

The genetic variation of a cultured species is very much influenced by aquaculture practices which in turn, driven by both the local and global demands. As such, the upheave struggle faced by farmers and industry-alike is to prevent the loss of genetic diversity over successive generation, mainly due to a small founder population and ultimately small effective population size. In a closed and strictly controlled environment, domestication selection, inbreeding, non-random mating, differential survival of progeny, admixture mixing of genetically distant strains or combination of these events entails genetic changes that may not favour sustainability of the culture lines for the stressful nature of commercial production.

According to the current results, the Rawang population showed significant but low indication of inbreeding. Among the cultured populations of *C. striata* in this study, Rawang was the only large-scale commercial farm which solely reared this fish species in 16 ponds that spans over 4.3 hectares. The farm generates up to 18 tonnes metric of *C. striata* per month and is the main *C. striata* supplier for the local food production of canned Haruan (*C. striata*) soup. The broodstocks, according to the farm owner was sourced locally and from the neighbouring country of Thailand. The farm owns hatcheries and nurseries that are capable of producing 150, 000 – 200, 000 *C. striata* seedlings for every spawning cycle. Hence, it is not unexpected that the risk of inbreeding rises in tandem with the magnitude of the farming operations and that insufficient management practices may be acting negatively on the stocks. However, it should be noted that inbreeding techniques are generally and traditionally applied in commercial ventures in order to boost productivity and profits by creating genetically superior stocks. This may be the case for the Rawang farm. On the other hand, there is

also a chance that the samples collected do not accurately reflect the total genetic content of the Rawang population because of the sheer scale of operations, thus the resulting inflated homozygosity which led to the erroneous inference of inbreeding.

5.4. Population bottleneck

The Garza-Williamson statistics indicate recent reductions of population size (M < 0.68) in the Kedah (wild), Kajang (cultured) and Malacca (cultured). The population bottleneck together with homozygote excess discovered in the wild population of Kedah, would certainly coincide with the departure of Kedah from HWE. On the other hand, the Garza-Williamson statistic seems to be biased towards populations that highly diverse, shown from the cultured populations of Kajang and Malacca. Based on observations gathered from the microsatellite analyses, in a population which contains high genetic variation (in this case, Kajang and Malacca) with a low sample size and a wide distribution range of alleles, the allelic range R, would increase by a greater proportion than the increase in number of alleles, k. A lower estimate might arise because of the Garza-Williamson statistic could be impervious to the uneven spacing of alleles due to high variation when the samples are not representative of the whole population, which causes underestimation of the Garza-Williamson index. Therefore, the statistics obtained from Garza-Williamson index must be interpreted with caution when dealing with other than wild populations and should be substantiated using alternative methods.

5.5. Potential genetic influence of escaped culture *C. striata* on the wild populations

There are also growing concerns raised on the possible impact of cultured *C. striata* that had escaped or were released (intentionally or unintentionally) into the wild. Escapees can adversely affect local ecosystems through interbreeding between genetically distinct strains, and consequently cause arbitrary yet predominantly detrimental alteration to the genetic variation of the native populations (local gene pool). This in turn will lead to reduced fitness (survival and recruitment) of wild populations (Bekkevold et al., 2006). The introgression of cultured lines into the wild population will increase negative interactions within the local ecology including possible spread of pathogens and parasites, eventually changing the landscape via predation and competition for natural resources, as well as the transmission of diseases.

The present result implied that the current cultured stocks of *C. striata* in Malaysia may have significant potential impact to the local genetic diversity. However, since genetic variation of cultured populations are substantially higher than the wild, loss of genetic diversity of the wild populations at this point is highly implausible. Rather, the emphasis would be on the loss of private alleles and genetic identity of the natural (original) population in the event of introgression between the cultured runaways and wild population. To sum up, it would be interesting to observe the influence of farmed strains on the wild gene pools of the native Malaysian *C. striata* in the persistence of conservational genetics. Hence, the need to establish another study to investigate *sine qua non* whether there are any incidences of cultured *C. striata* escaping and introgressing in the wild. A comparison between cultured *C. striata* from farms with wild populations sampled at streams or freshwater bodies in close proximity

to the farms could be carried out to assess whether there have been significant changes in the diversity of the wild population.

5.5. Genetic differentiation between cultured and wild populations

Genetic differentiation between cultured and wild population (stocks) has been observed in many other studies involving freshwater and marine species [Artic charr (Lundrigan et al., 2005), Atlantic cod (Pampoulie et al., 2006), black tiger shrimp (Xu et al., 2001), channel catfish (Simmons et al., 2006), gilthead sea bream (Alarcon et al., 2004), mud carp (Yang et al., 2008) and New Zealand greenshell mussel (Apte et al., 2003)]. AMOVA analysis detected significant albeit low group association among pooled cultured and wild groups in the study. In addition, 79.20% of the total variation were partitioned to within individuals components whereas only 8.32% of variation was found among the groups (cultured versus wild) assigned. This implies that there is little variation between the cultured and wild groups as majority of the variation exists on the individual level without any regards to its population.

In the present study, pair-wise comparisons of F_{ST}, R_{ST} and D_{est} between population revealed significant differentiation ranging from minimal to distinctly differentiated. These differentiation statistics seem to agree that the population with least genetic divergences were between the cultured populations of Kajang and Rawang. According to the management of the Kajang farm, the broodstocks were obtained from Klang. Since Rawang, Kajang and Klang were all situated in the state of Selangor and that the distances between any of these locations did not exceed 50 kilometres, it may be possible that the farm in Rawang may have sourced breeding fishes within the

vicinity of Klang. This would describe relatedness between Kajang and Rawang although further studies need to be conducted to compare these cultured populations with the wild population sampled near these farms.

On the other hand, the wild populations of Kedah and Pahang were extensively dissimilar genetic-wise, according to the pair-wise estimators of F_{ST} and R_{ST}. This is mainly due to the geographical barrier of the Titiwangsa Mountain Range that separates the west coast of peninsular Malaysia from the east, whereby Pahang is situated on the east coast of the peninsular central whilst Kedah is located at the northern region of the west coast, with part of its border neighbouring Thailand. Ambak et al. (2006) and Mat Jais et al. (2009) reported similar geographic structuring of *C. striata* that occurred between hydro-geographic factors (Titiwangsa Mountain Range) and population genetic differentiation based on RAPD and RFLP data respectively. However, D_{est} estimates indicate that the highest genetic delineation exists between the population of Kajang (cultured) and Kedah (wild). Despite this anomaly, Mantel tests revealed significant correlation between the differentiation estimators, hence it would be reasonable to assume that the results obtained were in sync and valid.

5.6. Genetic similarity and dissimilarity suggested by cluster analysis and phylogenetic trees

The Bayesian-approached tests implemented in STRUCTURE has allotted five populations to the studied populations by assigning individuals to spatial population based on their genotypes. The analysis noted similarities between Johore and Pahang and hence, were clustered as a single population. Likewise, results obtained from the NJ tree construed from pair-wise F_{ST} estimates and UPGMA dendogram based on Nei's 1978 genetic distance suggested noticeable genetic similarity between wild populations of Johore and Pahang, consistent with the patterns from cluster analysis which may reflect the origin of the population. Since Johore and Pahang is connected by the Endau river which flows along the borders of both states and into the South China Sea, therefore, the possibility of Johore and Pahang population sharing a common ancestry seems completely reasonable even though *C. striata* are not migratory species. In addition, similarity was also detected in the trees between the cultured populations of Kajang and Rawang. This mirrors the findings discussed earlier about the relatedness of both populations. Nevertheless, current findings can only conclude the most probable origin from the set of wild populations sampled, and without extensive sampling exact origins cannot be established.

UPGMA dendogram showed quite the opposite where greatest dissimilarity was between wild population of Johore and cultured population of Malacca even though the state of Malacca is located adjacent to Johore. The breeder from this particular farm (Malacca) was adamant about revealing any information pertaining to the origin of the cultured stocks and would only affirm that re-introduction of wild *C. striata* (sourced locally from many different parts of the country) into his hatchery stocks were

conducted periodically. Moreover, the wild population of Johore was sampled in a dam. Thus, this may likely explain the lack of similarity genetic-wise, given that both populations were genetically and artificially isolated although both states are geographically adjoined to the other.

5.7. General Discussion

On account of its medicinal benefits, *C. striata* locally known as Haruan has become one of the fastest growing aquacultural commodity in Malaysia. High demand of this species has lead to gradual depreciation of the natural population due to over-exploitation, placing increased pressure on farmed aquatic resources to fulfil the demands. Large–scale commercial farming of *C. striata* is still at its infancy stage in Malaysia compared to other established cultures (freshwater and river catfishes, red tilapias, giant freshwater prawns & river carps, Department of Fisheries, 2010), as most *C. striata* production is dependent on breeding stocks derived from the wild. Genetic variation is imperative in cultured stocks in order to obtain genetically diverse lines for accommodating varying aquaculture environments which include the rising emergence of disease outbreaks. However, wild stocks are also of key importance as in-depth knowledge on the levels and patterns of genetic diversity contained in wild stocks can serve as a valuable resource in the improvement of culture stocks. Therefore, understanding of the distribution of genetic variation within cultured and wild populations is vital to ensure adequate management of broodstocks for the development of sustainable and economically viable industries.

Findings from the current study indicated that the cultured populations of *C. striata* analysed contained high gene and allelic diversity than the wild populations. This could be attributed to the combinations of genetically diverse individuals to generate a broodstock with greater variation by the operating farms. The regular practice of re-introducing wild *C. striata* into the existing broodstock by certain operators may also contribute to the high variation observed among the cultured population. In addition, the elevated diversity observed in the cultured populations seem to suggest short domestication history since the populations showed signs of sufficient N_e to warrant the lack of genetic drift (random changing of gene frequencies).

The high degree of genetic variability demonstrated in cultured populations has lead to questions on whether this level of diversity could be maintained for a sustained response from long-term selection for commercially important traits (Davis and Hetzel, 2000). This is because, major concerns of aquaculture practices does not only revolved around gathering but also maintaining as much of the naturally occurring variation as possible within the cultured populations (Lind et al., 2009). Domestication of populations over generations, by and large are associated with the decline in genetic variation due to founders effect, non-random mating, genetic drift, selection (based on commercial traits) and inbreeding. Therefore, retention of the genetic diversity of the cultured population is a requisite to the long-term success of breeding programs.

Recommendations for *C. striata* culture industry:

- Enrichment of the domesticated *C. striata* broodstocks through a specific genetic improvement programme similar to Genetically Improved Farmed Tilapia (GIFT), to select for desirable traits (morphology, growth, survival rate), disease resistant strains as well as superior tolerance to sub-optimal conditions. Aquaculture of tilapia has benefitted tremendously from the GIFT where rapid genetic improvement of cultured tilapia through selective breeding were successful in augmenting species production, enhancing the profitability of fish farming, increasing welfare of the local people and decreasing the price of the fish (Gupta et al., 2000). Methods employed in the GIFT tilapia have been practiced in the genetic enhancement of other species viz., silver barb (Bangladesh and Thailand), rohu (Bangladesh and India), common carp (Vietnam), mrigal (Vietnam) and blunt snout bream (China) (Siriwardena, 2007). Malaysia should pursue this opportunity to develop a superior strain of *C. striata* well adapted for the aquaculture production industry.

- To create a strategic network among *C. striata* farmers, state fisheries department and scientific institutions in order to accelerate research in developing a commercially viable and disease resistant culture strain. Under this alliance, pooled resources and knowledge, technical assistance, regular community consultations, discussions and feedback as well as collaborations with researchers will contribute to enhancing management of the species genetically and performance wise. Proper channels could be set up to ensure that the supply of the fish swiftly fulfils the demand of the market.

Further research to be conducted includes:

- A preliminary set of tests should be conducted in order to detect any potential hybrid introgression events that may have occurred to the wild populations in Malaysia due to cultured *C. striata* escapees. At the time of writing, there have not been any data published on this particular issue. Therefore, it is difficult to predict the potential genetic impact of *C. striata* escapees on the wild due to the lack of reliable data. Subsequently, baseline information obtained from these type of analyses may reveal estimates of proportions of escapees in wild populations, as well as dynamics and magnitudes of introgressions that can assist in genetic conservation measures and also in broodstock management.

- To test the theory of whether the high levels of variation in cultured populations was due to short domestication history and sufficient effective population size, genetic monitoring should be conducted on a regular basis on the successive generations of the same commercial farms so that the progression of differences in genetic variation could be comprehensively examined.

- Research into potential heterosis or outbreeding depression among the cultured yet diverse populations of *C. striata* as crosses between strains may provide a significant improvement in production efficiency (Davis and Hetzel, 2000). Efficiency will be based on the performance of the stocks under a controlled environment as well as the value of its commercial traits.

- Multiplex PCR genotyping should be considered as it is less time-consuming and labour-intensive than single locus PCR. However, a preliminary trial run of the multiplex system should be compared against the conventional approach to determine the extent of accuracy of multiplex genotyping.

CONCLUSION

Data analyses from seven polymorphic microsatellite markers (PCT6-6, BP13-6, BP13-14, CS-4, CS-5, T113-11, BP6-2 and BP6-2) revealed genetic variation and differentiation between the wild (Johore, Kedah and Pahang) and cultured (Kajang, Malacca and Rawang) *Channa striata* (*C. striata*) populations in this study. Genetic parameters (heterozygosity and allelic diversity) estimated from microsatellite data indicated that the cultured populations of *C. striata* have substantially higher genetic variation compared to the wild populations. Further analysis indicated that majority of the genetic diversity found were among individuals as opposed to within and, between the pooled cultured and pooled wild groups.

In the current study, wild stocks were adequately represented in each of the wild populations, apart from Kedah, which appears to have deviated from HWE. Further investigations detected recent population bottleneck along with over-representation of homozygotes in Kedah. On the other hand, although not unusual, all of the cultured populations exhibited non-conformance to HWE. High values of heterozygosity-based indices suggested that a generous number of breeding individuals accounted for, and created the present cultured populations with the exception of Rawang, whereby a significant yet low level of inbreeding was detected.

Significant levels of genetic differentiation, ranging from low to high was observed among, and between, the cultured and wild populations. Analysis (AMOVA) showed minimal however, significant differentiation between pooled cultured and pooled wild group. Differentiation estimators based on the microsatellite data revealed a positive correlation between population genetic delineation and spatial distribution

(which included hydro-geographic factors) of the wild *C. striata* populations. This was reflected between the wild populations of Johore and Kedah which were the least differentiated, possibly since both states are adjoined to each another in which a river (Sungai Endau) mutually connects them together. On the contrary, physical barrier in the form of extensive mountain ranges (Titiwangsa) have increased the genetic dissimilarity between wild populations located in the east (Pahang), and west (Kedah) coast of peninsular Malaysia. As for the cultured populations however, no apparent pattern was discernible given that the broodstock history was not disclosed by the farmers and that the culture stocks were not necessarily sourced close to the farms.

All in all, the present study indicates that there is a large amount of genetic variation distributed among cultured populations of *C. striata* in a stark contrast to the wild populations. Hitherto, it appears that the domestication of *C. striata* have not led to depreciation in levels of genetic diversity, implying that the genetic "health" of the current culture stocks studied are satisfactory and have yet to experience any genetic erosion. This finding is of importance as loss of genetic variation affects the fitness of a species which probably induces a decline in the potential adaptability of a population under intense aquaculture environment. The resulting shortfalls due to lack of adaptation includes deterioration of economically significant traits such as reduction of growth, size, survival rate, disease resistance and the rise of abnormality rates in the cultured stocks. Also, a comparison study like this is vital as it identifies and regulates the changes in the genetic dynamics of a culture in order to prevent outbreeding depression that occurs when interspecific crosses from genetically divergent lines produces progenies with inferior survival capacity.

Since a relatively high degree of genetic diversity has accompanied the cultured populations of *C. striata* in this study, the key management issue then shifts to whether it is plausible to retain as much of the naturally sourced variation as possible within the cultured stocks. To be economically lucrative and viable, the *C. striata* industry should therefore ensure sustenance of variation in the ensuing generations of culture with sound genetic management practices. Hence, a following study should be conducted in the farms surveyed, to monitor the fluctuations of genetic diversity in the successive generations so as to gain a thorough insight on the current management practices. Potential directions for future research were also discussed in the study.