CHAPTER 4: RESULTS

4.1 Isolation and Development of Microsatellite Markers

4.1.1. PCR amplification of RAMs primers

Table 4.1 shows the RAMs primers and their respective annealing temperatures. Figure 4.1 to 4.5 shows the banding profiles for five different RAMs that were tested on five different individuals. Five samples were randomly chosen to get sharpest and brightest bands that will be used for cloning steps. Banding pattern shows multiple bands that indicate different length and microsatellite repeats.

Table 4.1 RAMs primer and its annealing temperature, Tm

RAMs primer	Annealing temperatures (Tm) (°C)
BP11	51.2
T79112	41.3
T79110	54.1
VJ2	55.0
BP8	45.5

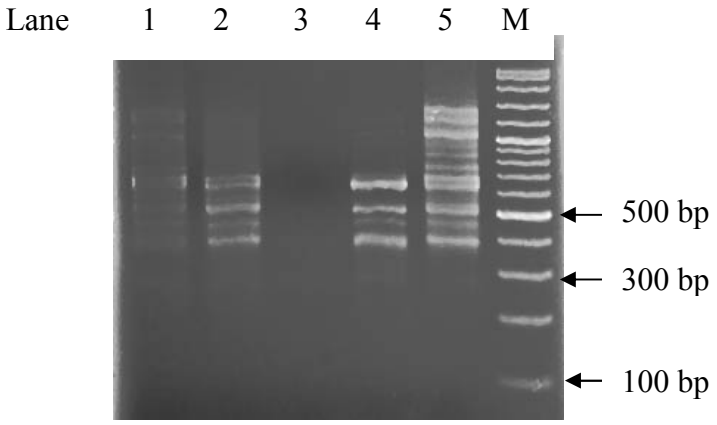

Figure 4.1: PCR products obtained from primer BP11

(Lane 1 to 5: sample 1 to 5; Lane M: 100bp ladder)

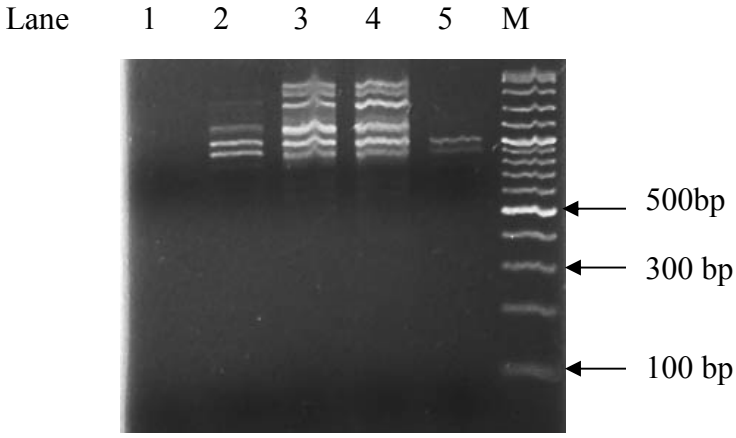

Figure 4.2: PCR products obtained from primer T79112

(Lane 1 to 5: sample 1 to 5; Lane M: 100bp ladder)

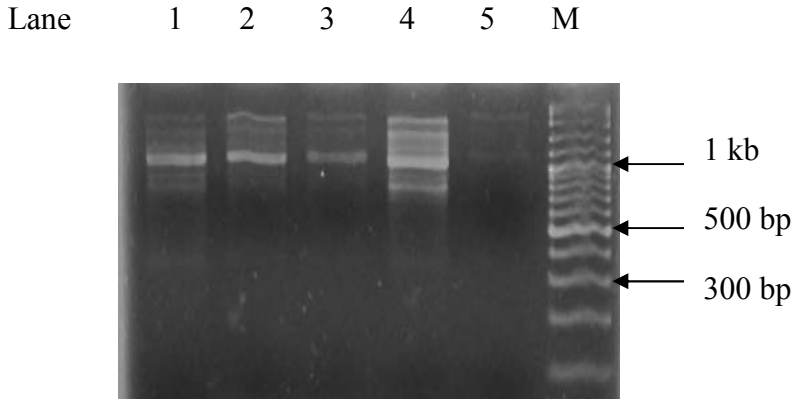

Figure 4.3: PCR products obtained from primer T79110

(Lane 1 to 5: sample 1 to 5; Lane M: 100bp ladder)

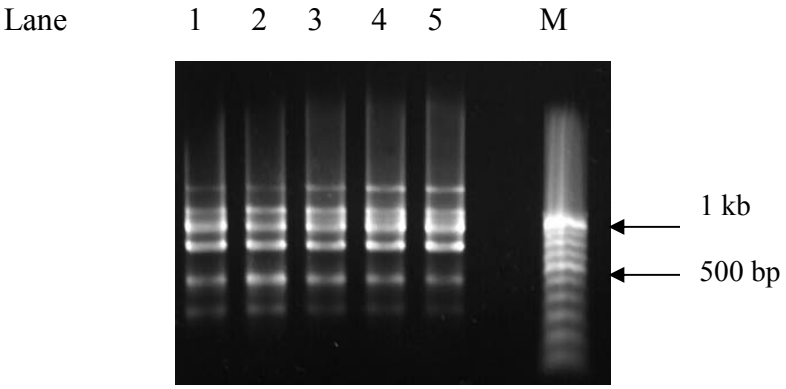

Figure 4.4: PCR products obtained from primer VJ2

(Lane 1 to 5: sample 1 to 5; Lane M: 100bp ladder)

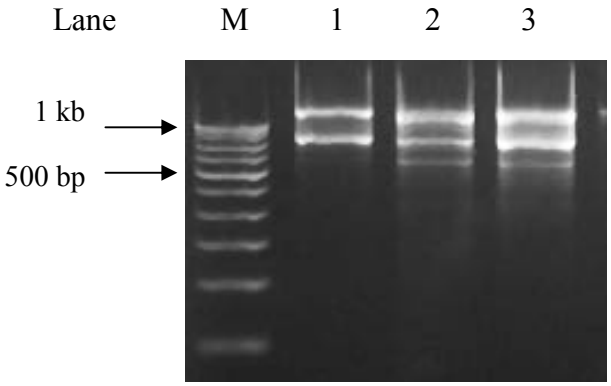

Figure 4.5: PCR products obtained from primer BP8

(Lane 1 to 3: sample 1 to 3; Lane M: 100bp ladder)

4.1.2 Colony PCR

Colony PCR can be use after a transformation to screen colonies for the desired plasmid. M13F and M13R primers were use to generate a PCR product of known size. Thus, any colonies, which give rise to an amplification product of the expected size, are likely to contain the correct DNA sequence. Figure 4.6 shows bacteria colony from blue white screening using M13F and M13R. Bands above 500 bp were chosen for plasmid extraction and further sent for sequencing. These colonies contain DNA insert and M13F and M13R. The sequences of M13F and M13R were about 400bp in length.

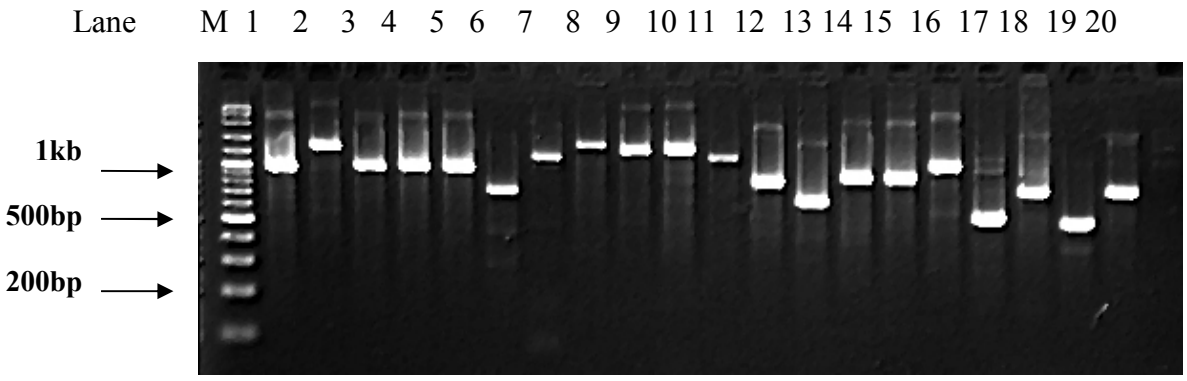

Figure 4.6: Colony PCR

(Lane 1 to 20: bacterial colony 1 to 20; Lane M: 100bp ladder)

4.1.3 Plasmid Extraction

Plasmid extractions have done to check the size of the plasmid that contained cloning vector and DNA insert. After size checking, the sequences sent for DNA sequencing. Figure 4.7 showed plasmid extraction of RAMs primer, BP11. Cloning vector is 2.7kb in length and the bands that show 3 kb and above were chosen for sequencing.

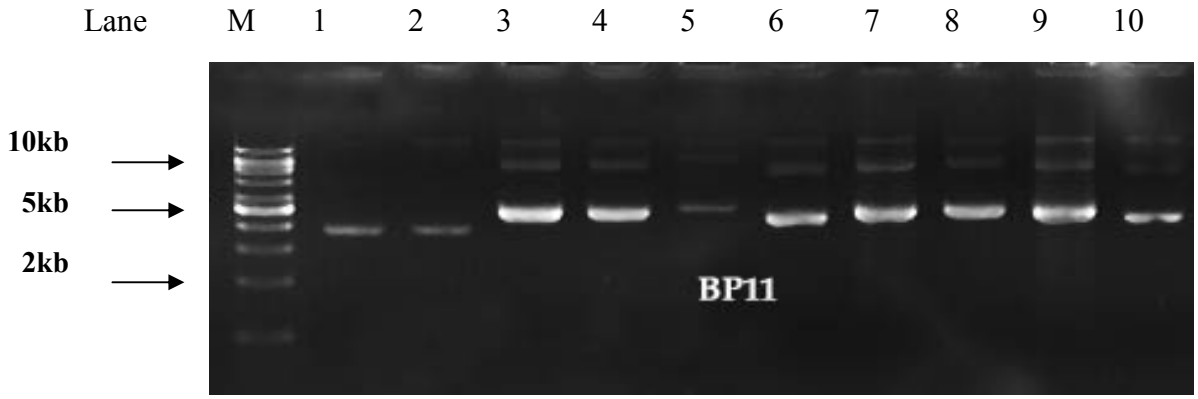

Figure 4.7: Plasmid of 5' anchored PCR clones from BP11 primer

(Lane 1 to 10: colony 1 to 10; Lane M: 1kb ladder)

4.1.4 Sequencing results

Twenty clones from each of the degenerate RAMs primer were randomly chosen for sequencing. Figure 4.8 shows types of microsatellites repeats obtained from sequencing results. Sequences of an isolated fragment of seven polymorphic primers were shown in Appendix C.

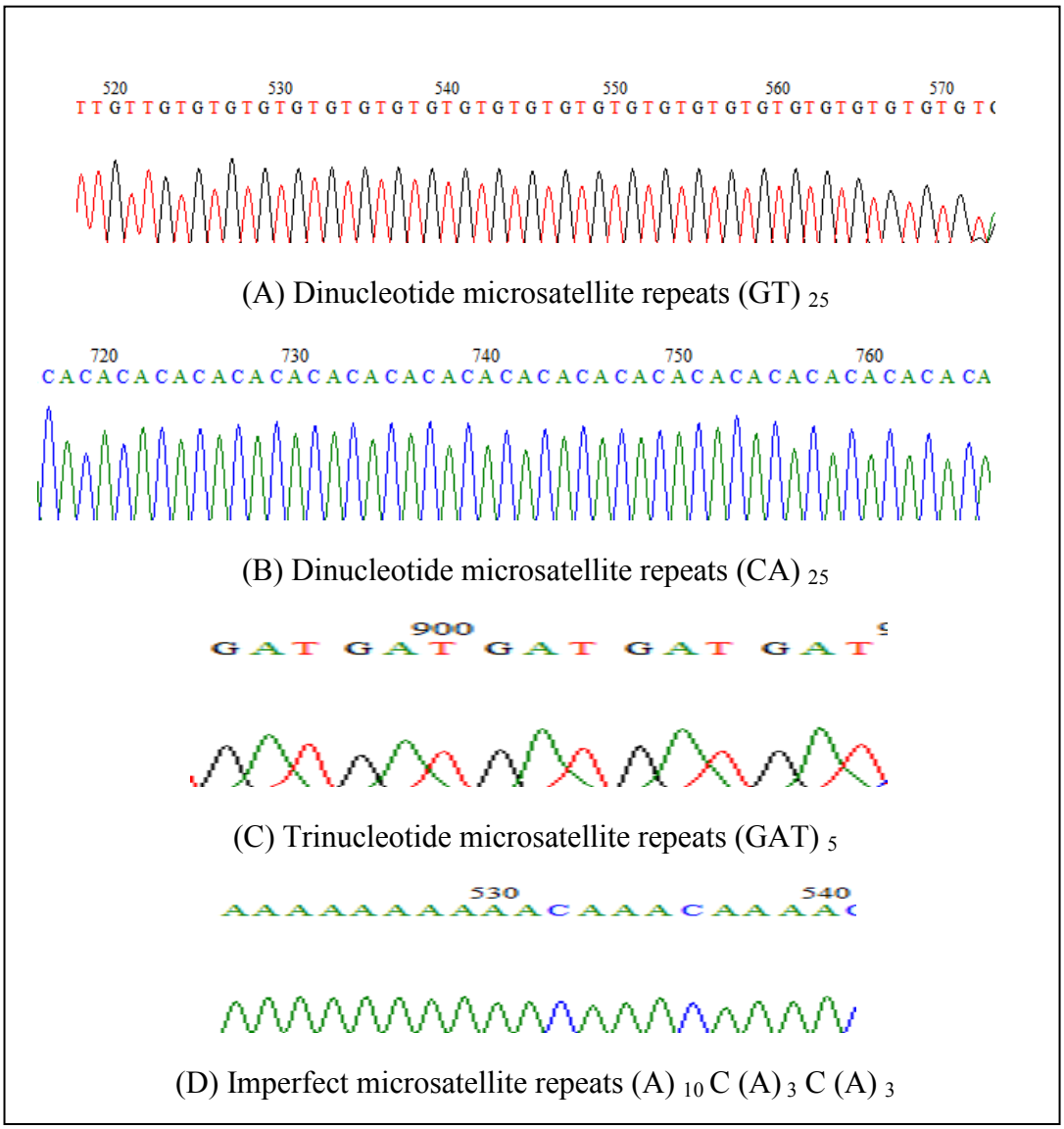

Figure 4.8: Microsatellite repeats

4.1.5. Optimization of Microsatellites Primers

Microsatellite primers were first optimized to determine concentration of $MgCl_2$, annealing temperature, and PCR cycles. Figure 4.9 to 4.14 shows optimization banding profile for seven polymorphic primers. The banding profile is between 150 to 300bp because microsatellite repeats are between these regions. The optimal condition is shown in Table 4.2 and the annealing temperature for seven microsatellite loci is show in Table 4.3.

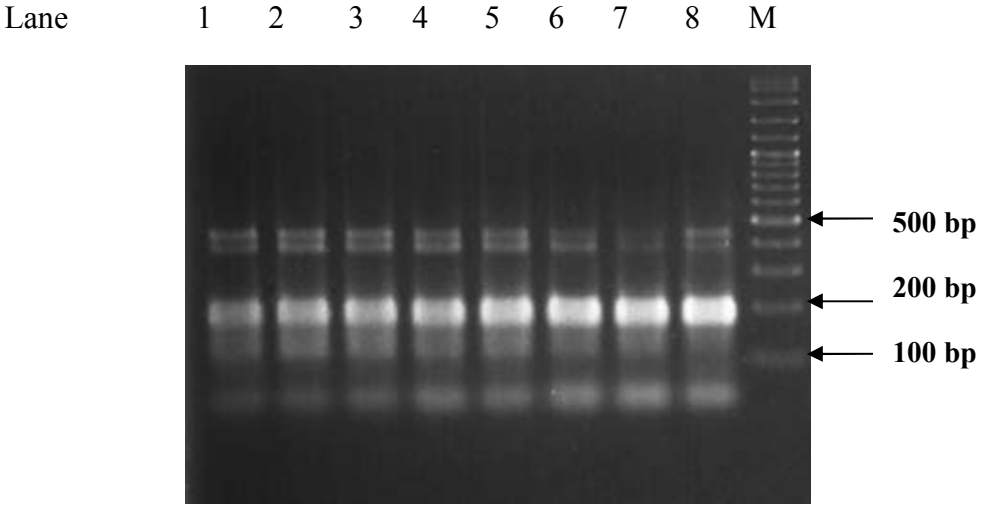

Figure 4.9: Optimization of primer SFO-T112-H4 (optimization shows double bands)

(Lane 1: 48.0°C, Lane 2: 48.5°C, Lane 3: 49.4°C, Lane 4: 50.8°C, Lane 5: 52.5°C, Lane 6: 53.8°C, Lane 7: 54.6°C, Lane 8: 55.0°C; Lane M: 100 bp ladder)

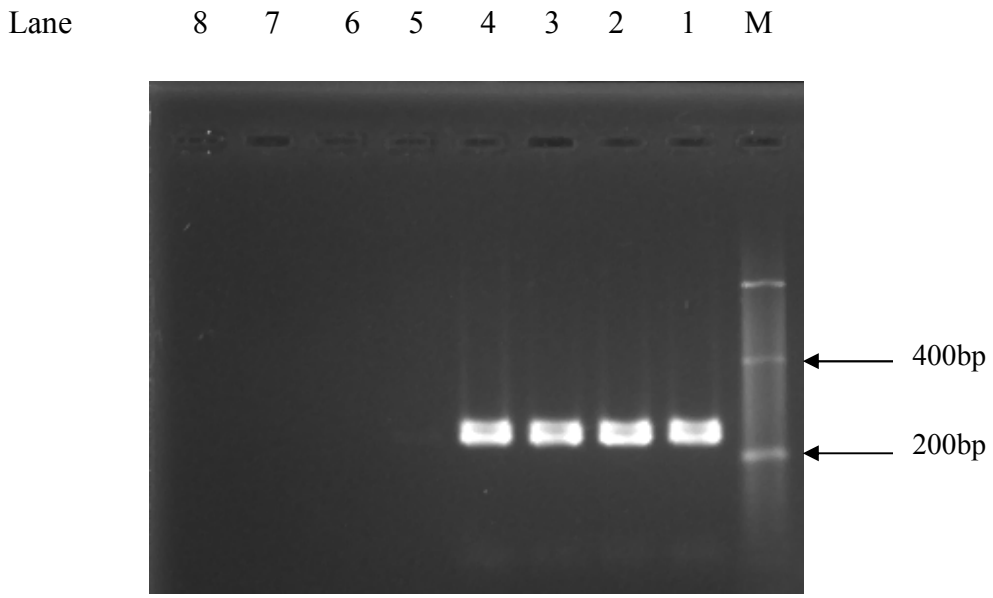

Figure 4.10: Optimization of primer SFO-BP8-H1 (optimization shows double bands)

(Lane 1: 53.0°C, Lane 2: 54.0°C, Lane 3: 55.7°C, Lane 4: 58.2°C, Lane 5: 61.5°C, Lane 6: 64.2°C, Lane 7: 65.9°C, Lane 8: 67.0°C; Lane M: 100 bp ladder)

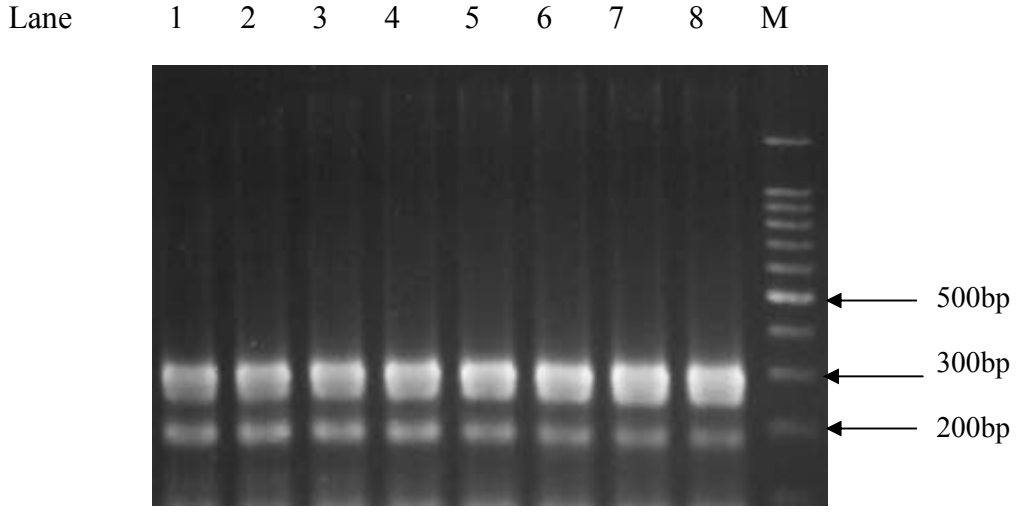

Figure 4.11: Optimization of primer SFO-T112-H6 (optimization shows double bands)

(Lane 1: 49.0°C, Lane 2: 49.5°C, Lane 3: 50.4°C, Lane 4: 51.8°C, Lane 5: 53.5°C, Lane 6: 54.8°C, Lane 7: 55.6°C, Lane 8: 56.0°C; Lane M: 100 bp ladder)

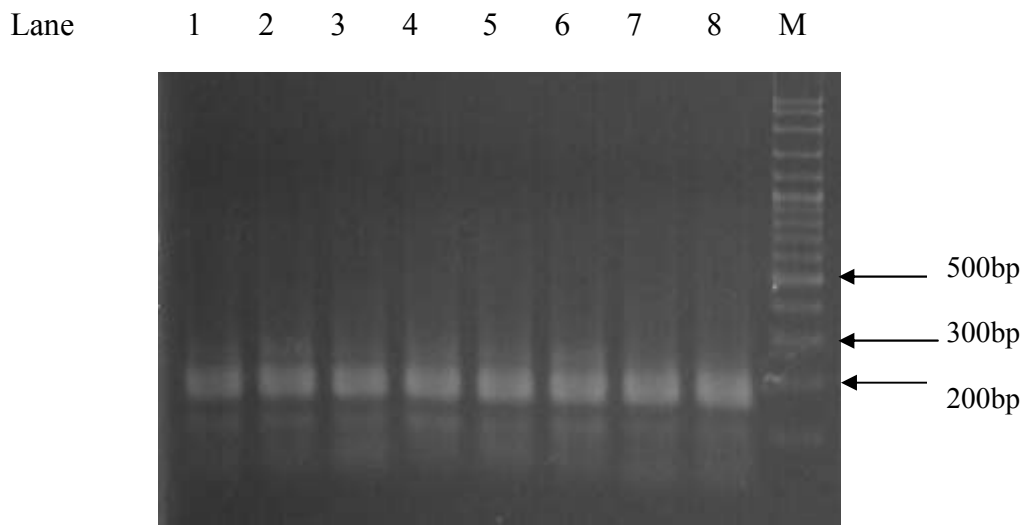

Figure 4.12: Optimization of primer SFO-T112-H6I (optimization shows single band)

(Lane 1: 50.0°C, Lane 2: 50.4°C, Lane 3: 51.2°C, Lane 4: 52.4°C, Lane 5: 53.8°C, Lane 6: 55.0°C, Lane 7: 55.7°C, Lane 8: 56.0°C; Lane M: 100 bp ladder)

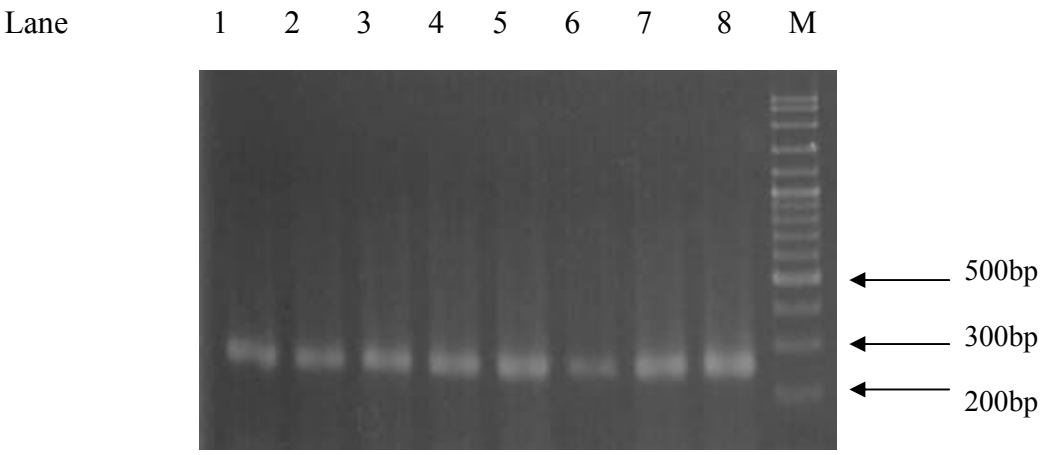

Figure 4.13: Optimization of primer SFO-T112-H6F (optimization shows single band)

(Lane 1: 50.0°C, Lane 2: 50.4°C, Lane 3: 51.2°C, Lane 4: 52.4°C, Lane 5: 53.8°C, Lane 6: 55.0°C, Lane 7: 55.7°C, Lane 8: 56.0°C; Lane M: 100 bp ladder)

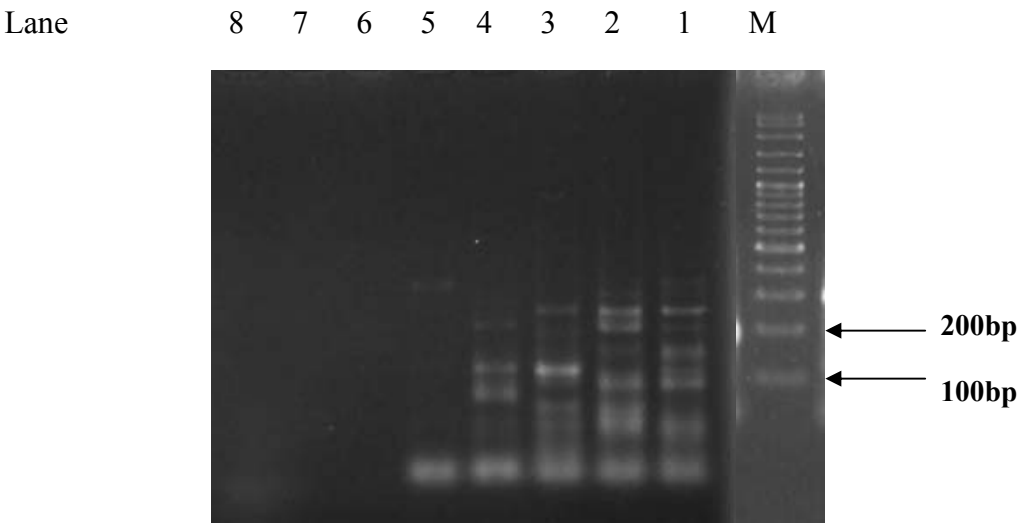

Figure 4.14: Optimization of primer SFO-BP8-H2 (optimization shows multiple bands)

(Lane 1: 53.0°C, Lane 2: 54.0°C, Lane 3: 55.7°C, Lane 4: 58.2°C, Lane 5: 61.5°C, Lane 6: 64.2°C, Lane 7: 65.9°C, Lane 8: 67.0°C; Lane M: 100 bp ladder)

Table 4.2 Optimal PCR conditions for seven microsatellite loci

Solution	**Volume (µL)**
2.5mM of $MgCl_2$	1.5
1 x PCR Buffer	2.0
0.25 mM of each dNTPs:	
dA	0.25
dG	0.25
dC	0.25
dT	0.25
1.5 unit of Taq polymerase	0.2
50.0 pmole of primer (1^{st} Base)	0.5
20 ng of genomic DNA	2.0
deionised water	3.5
Total volume:	**10.0**

Table 4.3 Characterization of seven *Channa striata* microsatellite loci

Locus	GenBank Accession no.	Primer sequences (5'-3')	Repeat sequence	*Tm* (°C)	No. of alleles	$MgCl_2$ (Mm)	Product size (bp)	*Ho*	*He*	*P*
SFO-T112-H4	GQ130215	F:AACAGGAAGTGATATGAGTG R:AAGTTATGGTTCTCCCTATC	$(GT)_9$	55.0	2	2.5	167	0.36667	0.30452	0.5511
SFO-T112-H6	GQ130217	F:GCACAGTCACTGAAGGTA R:CAAACAGCTGTGAGACTG	$(CA)_{10}TG(CA)_3GA(CA)_3$	55.6	2	2.5	278	0.26667	0.23503	1.00000
SFO-T112-H6I	GQ130217	F:GGCAGACTTGTGTAATAGTG R:AACCTAATGCTGGAGAAC	$(GCGT)_4$	56.0	2	2.5	160	0.66667	0.45198	0.0108
SFO-T112-H6F	GQ130217	F:CACAGTAAGTTTCTGAGTGG R:CACACTATTACACAAGTCTGC	$(CA)_{25}$	56.0	2	2.5	256	0.03333	0.03333	1.00000
SFO-BP8-H1	GQ130232	F: GACTTCACGTCACTAACTTT R: TTTCTGCTGATGTTCTCTAC	$(CT)_2CA(CT)_4$	55.7	5	2.5	251	0.23333	0.74407	0.00000
SFO-VJ2-H38	GQ130240	F: CAGGTAGTTGGACGGTATAG R: AGGAACAGCTAGACCAGA	$(GGA)_4$	54.0	7	2.5	160	0.56667	0.72090	0.41838
SFO-BP8-H2	GQ130233	F: CTGCCTTAGTGTCTGTGGTT R:GACGTTCAAGAAACAAGGAA	$(TG)_{12}$	55.7	6	2.5	212	0.50000	0.69209	0.00000

4.1.6. Heterozygosity

Table 4.3 shows the loci, annealing temperatures, primer and repeats sequences, number of alleles, product size, accession numbers, concentration of $MgCl_2$, observed heterozygosities, expected heterozygosities, and probability values. These results were obtained by using Genepop software.

The average observed heterozygosity, expected heterozygosity, and numbers of alleles per locus were 0.37619, 0.45456, and 3.7, respectively. There are two alleles in five loci (SFO-T112-H4, SFO-T112-H6, SFO-T112-H6I, SFO-T112-H6F, SFO-BP8-H1), seven alleles in locus SFO-VJ2-H38 and six alleles in locus SFO-BP8-H2 (Table 4.3).

Three loci show deviation from Hardy Weinberg Equilibrium (HWE) law i.e. SFO-T112-H6I, SFO-BP8-H1, and SFO-BP8-H2.

4.1.7. Polymorphic test on 30 individuals.

All seven polymorphic primer pairs were tested on 30 individuals from three different locations i.e. ten individuals each from Negeri Sembilan, Johor, and Penang. Figure 4.15 show polymorphic test of primer SFO-BP8-H1 in 10 individuals from Negeri Sembilan. Figure 4.16 show polymorphic test of primer SFO-BP8-H1 in 10 individuals from Penang. Figure 4.17 show polymorphic test of primer SFO-BP8-H1 in 10 individuals from Johor. From the gel electrophoresis picture, some individual shows homozygote bands and others were heterozygotes. The product size is between 250 and 300 bp.

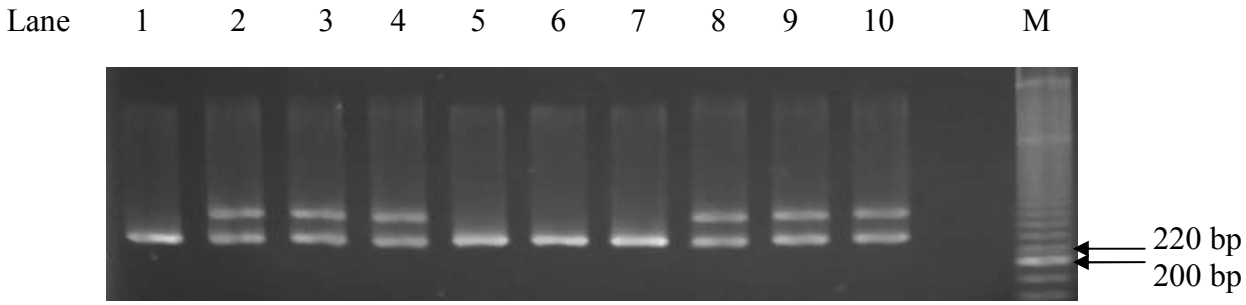

Figure 4.15: Polymorphic banding profile of primer SFO-BP8-H1 in Negeri Sembilan

(Lane 1 to 10: sample 1 to 10 Negeri Sembilan, Lane M: 20bp ladder)

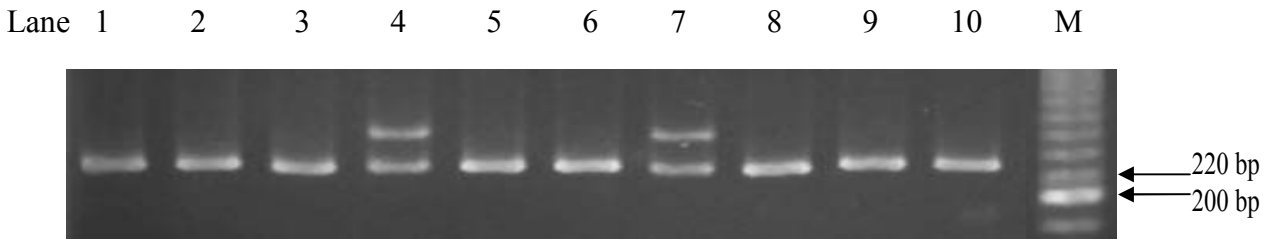

Figure 4.16: Polymorphic test of primer SFO-BP8-H1 in Penang

(Lane 1 to 10: sample 1 to 10 Penang, Lane M: 20bp ladder)

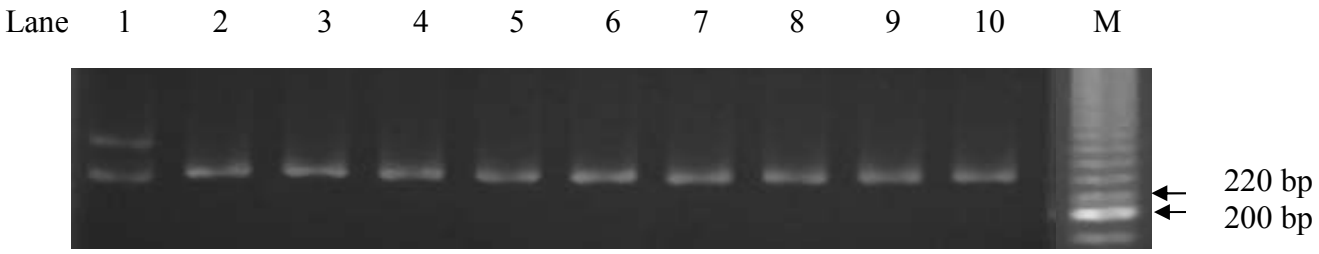

Figure 4.17: Polymorphic test of primer SFO-BP8-H1 in Johor

(Lane 1 to 10: sample 1 to 10 Johor, Lane M: 20bp ladder)

4.1.8. Data Analysis

Data analysis of microsatellite marker was done by using several software, Micro-Checker, Convert, Genepop, Popgene, Arlequin, Ntsys, Phylip, and Structure. Each has different functions and applications.

From the results obtained, most loci show no evidence for stuttering bands, allele drop out, and null allele except for loci SFO-BP8-H2 that shows evidence of null allele (Table 4.4).

Table 4.4 Stuttering bands, alleles drop out, and null alleles for polymorphic loci.

Locus	Stuttering Bands	Allele Drop Out	Null Alleles
SFO-T112-H4	NO	NO	NO
SFO-T112-H6	NO	NO	NO
SFO-T112-H6I	NO	NO	NO
SFO-T112-H6F	NO	NO	NO
SFO-BP8-H1	NO	NO	NO
SFO-VJ2-H38	NO	NO	NO
SFO-BP8-H2	NO	NO	YES

Genepop program calculates values of F_{IS} (Weir and Cockerham, 1984). This parameter measures the reduction in heterozygosity due to non-random mating within the subpopulation and thus helps to detect departures from Hardy-Weinberg Equilibrium by measuring the amount of heterozygote deficiency or excess observed in the sample. Values significantly greater than zero indicate an excess of homozygotes possibly resulting from inbreeding, population admixture or failure to detect heterozygotes. Conversely, negative F_{IS} indicates an excess of heterozygotes and outbreeding. Table 4.5 shows F_{IS} values for 30 individuals of *Channa striata*. Loci SFO-T112-H4, SFO-T112-H6, SFO-T112-H6I, and SFO-BP8-H1 show an excess of heterozygotes and outbreeding while loci SFO-T112-H6F, SFO-VJ2-H38 and SFO-BP8-H2 show an excess of homozygotes that indicate inbreeding, population admixture or failure to detect heterozygotes.

Table 4.5: F_{IS} values for 30 individuals.

Locus	F_{IS}
SFO-T112-H4	-0.2083
SFO-T112-H6	-0.1373
SFO-T112-H6I	-0.4872
SFO-T112-H6F	0.0000
SFO-BP8-H1	-0.0943
SFO-VJ2-H38	0.2168
SFO-BP8-H2	0.2810

F_{IS}: computed as in Weir & Cockerham (1984)

The frequency of alleles in a population can be use to predict the frequencies of the corresponding genotypes. It can be calculated using Genepop software. The total of allele frequency for each locus is equal to one. From results, the highest allele frequency is 0.983 (SFO-T112-H6F) and the lowest is 0.017 (SFO-VJ2-H38) (Table 4.6).

Table 4.6: Allele frequency

Locus	Allele Frequency						
	Allele A	Allele B	Allele C	Allele D	Allele E	Allele F	Allele G
SFO-T112-H4	0.817	0.183					
SFO-T112-H6	0.133	0.867					
SFO-T112H6I	0.667	0.333					
SFO-T112H6F	0.983	0.017					
SFO-BP8-H1	0.900	0.100					
SFO-VJ2-H38	0.233	0.450	0.150	0.017	0.017	0.033	0.100
SFO-BP8-H2	0.067	0.133	0.500	0.200	0.083	0.017	

4.2. Cross species amplification of microsatellite primers in *Channa striata*

Figure 4.18 and 4.29 show polymorphic test of cross species microsatellite loci from common carp (MFW5) in *Channa striata*. These figures show single and double bands at 176-280bp and there were 7 alleles in six populations. Table 4.7 shows optimized annealing temperature and product size obtained from five cross-amplified primers. Touchdown PCR was used to amplify PCR product in *Channa striata.*

Table 4.7 Optimized temperatures and product size of each primer used

Primer	Sequence (5' to 3')	Annealing temperature, Tm/ ºC	Product Size(bp)
MFW1	F: GTCCAGACTGTCATCAGGAG	58, 60	168-210
	R: GAGGGTGTACACTGAGTCACGC		
MFW2	F: CACACCGGGCTACTGCAGAG	58, 60	167-181
	R: GTGCAGTGCAGGCAGTTTGC		
MFW5	F: GAGATGCCTGGGGAAGTCAC	58, 60	176-276
	R: AAAGAGAGCGGGGTAAAGGAG		
MFW7	F: TACTTTGCTCAGGACGGATGC	58, 60	176-274
	R: ATCACCTGCACATGGCCACTC		
MFW15	F: CTCCTGTTTTGTTTTGTGAAA	58, 60	151-259
	R: GTTCACAAGGTCATTTCCAGC		

MFW1, 2, 5, 7, 15*: Crooijmans, 1997*

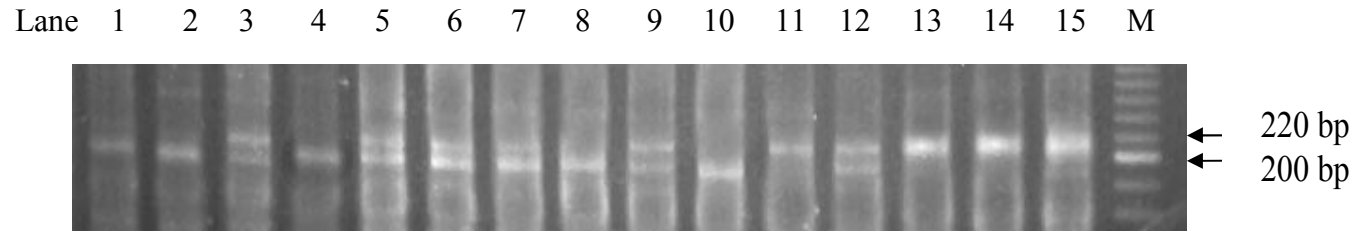

Figure 4.18 Polymorphic tests on MFW5

(Lane 1 to 15: sample 1 to 15 Negeri Sembilan, Lane M: 20bp ladder)

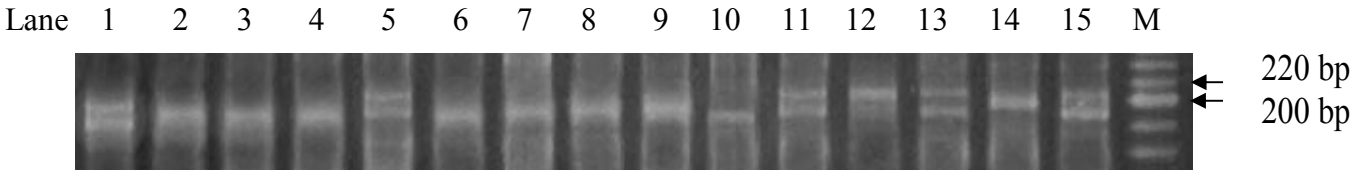

Figure 4.19 Polymorphic tests on MFW5

(Lane 1 to 15: sample 16 to 30 Negeri Sembilan, Lane M: 20bp ladder)

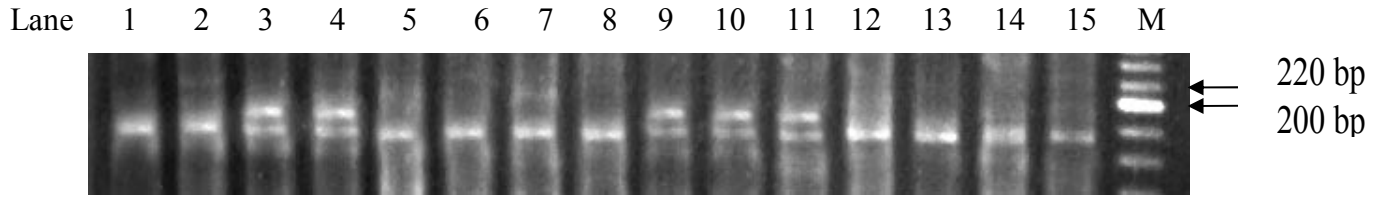

Figure 4.20 Polymorphic tests on MFW5

(Lane 1 to 15: sample 1 to 15 Johor, Lane M: 20bp ladder)

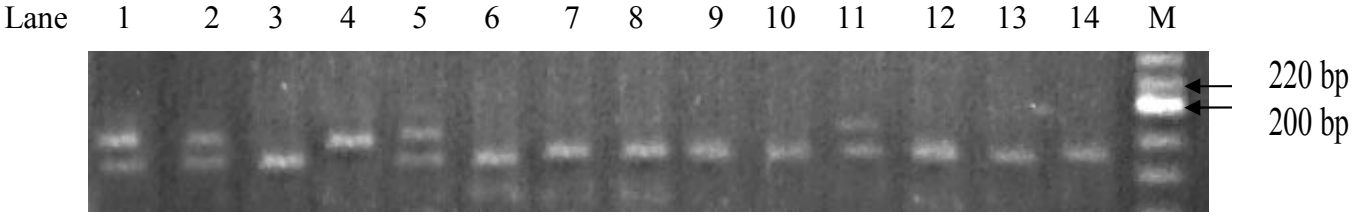

Figure 4.21 Polymorphic tests on MFW5

(Lane 1 to 14: sample 16 to 29 Johor, Lane M: 20bp ladder)

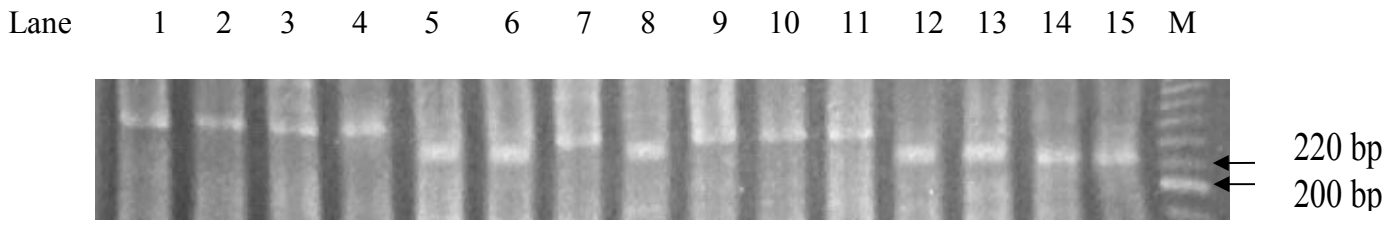

Figure 4.22 Polymorphic tests on MFW5

(Lane 1 to 15: sample 1 to 15 Penang, Lane M: 20bp ladder)

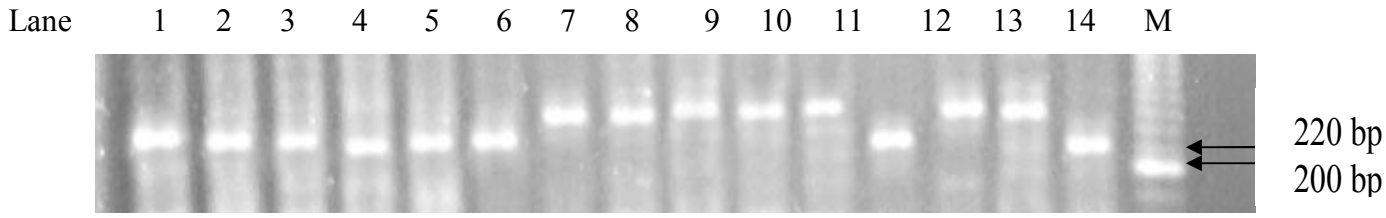

Figure 4.23 Polymorphic tests on MFW5

(Lane 1 to 15: sample 16 to 30 Penang, Lane M: 20bp ladder)

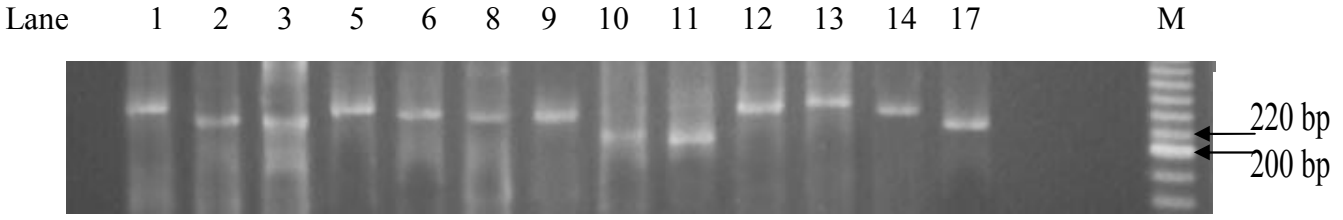

Figure 4.24 Polymorphic tests on MFW5

(Lane 1 to 17: sample 1 to 17 Selangor, Lane M: 20bp ladder)

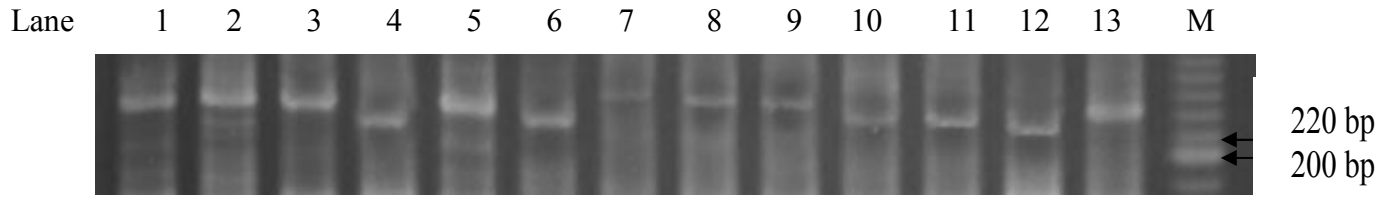

Figure 4.25 Polymorphic tests on MFW5

(Lane 1 to 13: sample 18 to 30 Selangor, Lane M: 20bp ladder)

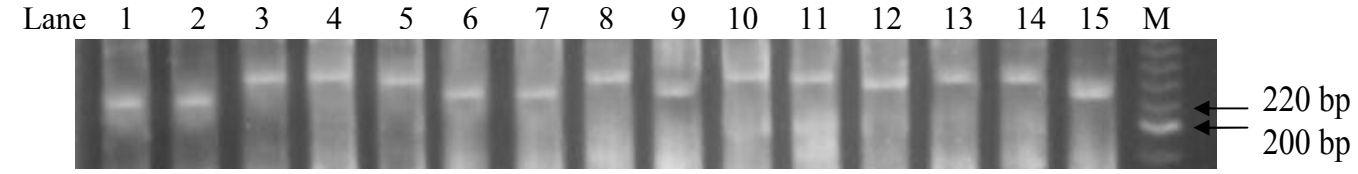

Figure 4.26 Polymorphic tests on MFW5

(Lane 1 to 15: sample 1 to 15 Terengganu, Lane M: 20bp ladder)

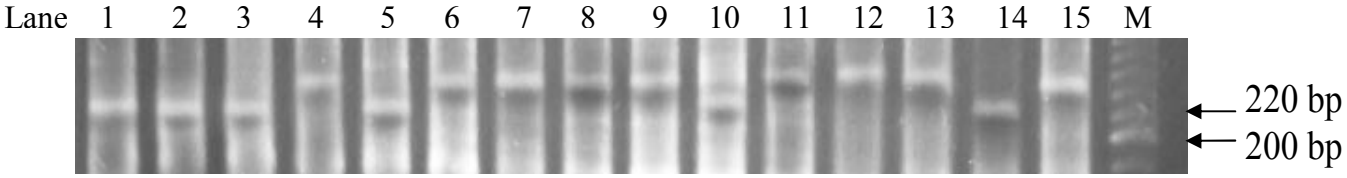

Figure 4.27 Polymorphic tests on MFW5

(Lane 1 to 15: sample 16 to 30 Terengganu, Lane M: 20bp ladder)

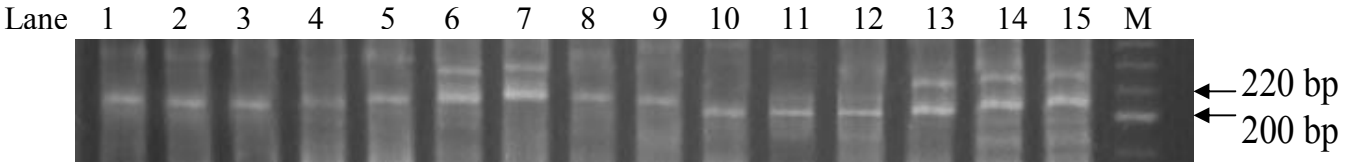

Figure 4.28 Polymorphic tests on MFW5

(Lane 1 to 15: sample 1 to 15 Kedah, Lane M: 20bp ladder)

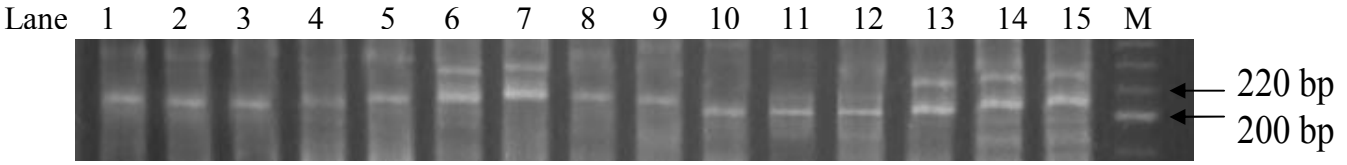

Figure 4.29 Polymorphic tests on MFW5

(Lane 1 to 15: sample 16 to 30 Kedah, Lane M: 20bp ladder)

4.3. Population genetic studies in six populations of *Channa striata*

Five isolated microsatellite primers (SFO-T112-H4, SFO-T112-H6, SFO-T112-H6I, SFO-T112-H6F, SFO-BP8-H1) and five cross-amplified primers (MFW 1, MFW 2, MFW 5, MFW 7, MFW 15) were used to construct population structure for six populations of *Channa striata* (Table 4.8). These primers were tested on 176 individuals from six populations originating from all over Malaysia. Figure 4.30 to 4.34 show amplified microsatellite pair primer SFO-T112-H4 in different populations. Product size for this primer is 167. The lower band shows the expected size and the upper show different length of microsatellite repeats.

The amplification of two loci (SFO-BP8-H2 and SFO-VJ2-H38) out of twelve microsatellite primer pairs used did not produced distinct bands after many trials under different PCR conditions. It was difficult to identify and score them as their bands were not clear. Therefore these loci were ignored in the population study.

Table 4.8 Polymorphisms at the microsatellite loci amplified for six populations of *Channa striata*

Locus	Size of PCR product	Number of alleles	Annealing temperature (°C)	$MgCl_2$	Buffer	Status
SFO-T112-H4	167-197bp	2	55.0	1.5 µl	3.0 µl	P
SFO-T112-H6	268-278bp	2	55.6	1.5 µl	3.0 µl	P
SFO-T112-H6I	160-196bp	2	56.0	1.5 µl	3.0 µl	P
SFO-T112-H6F	256-296bp	2	56.0	1.5 µl	3.0 µl	P
SFO-BP8-H1	250-280bp	4	55.7	1.5 µl	3.0 µl	P
MFW 5	176-280bp	7	58,60	1.5 µl	3.0 µl	P
MFW 1	168-188bp	6	58,60	1.5 µl	3.0 µl	P
MFW 7	176-274bp	7	58,60	1.5 µl	3.0 µl	P
MFW 15	151-259bp	8	58,60	1.5 µl	3.0 µl	P
MFW 2	167-181bp	2	58,60	1.5 µl	3.0 µl	P

P: Polymorphic

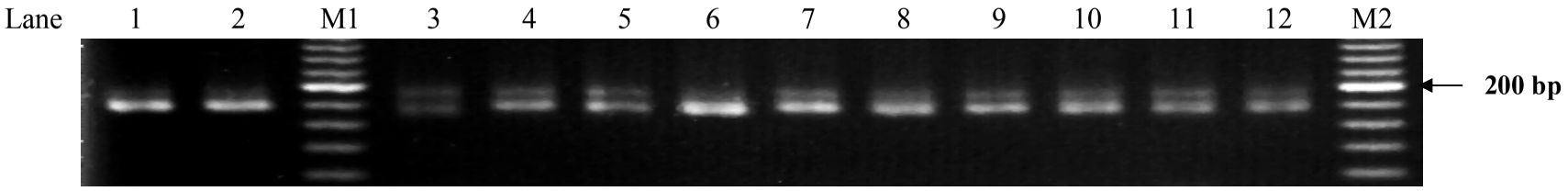

Figure 4.30: Polymorphic test on primer SFO-T112-H4

(Lane 1 to 12: sample 1 to 12 Penang, Lane M1 and M2: 20bp ladder)

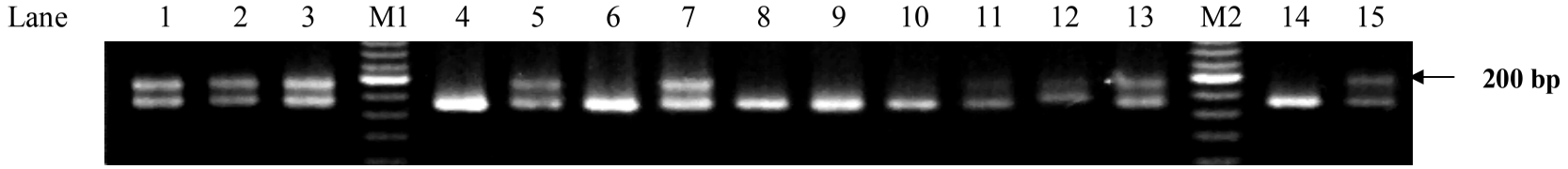

Figure 4.31: Polymorphic test on primer SFO-T112-H4

(Lane 1 to 3: sample 1 to 3 Negeri Sembilan, Lane 4 to 15: sample 4 to 15 Johor,

Lane M1 and M2: 20bp ladder)

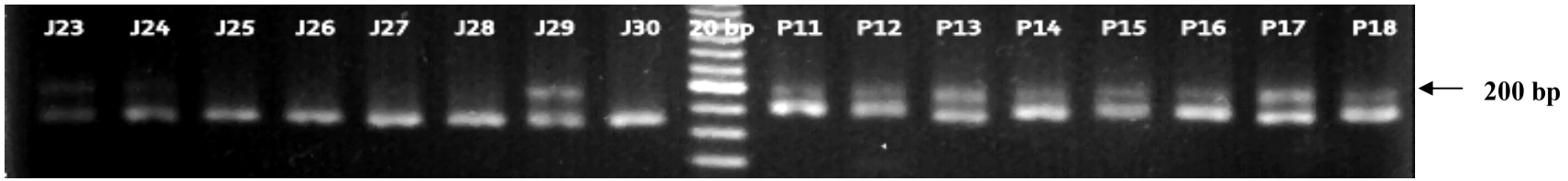

Figure 4.32: Polymorphic test on primer SFO-T112-H4

(Lane 23 to 30: sample 23 to 30 Johor, Lane 11 to 18: sample 11 to 18, Lane M1: 20bp ladder)

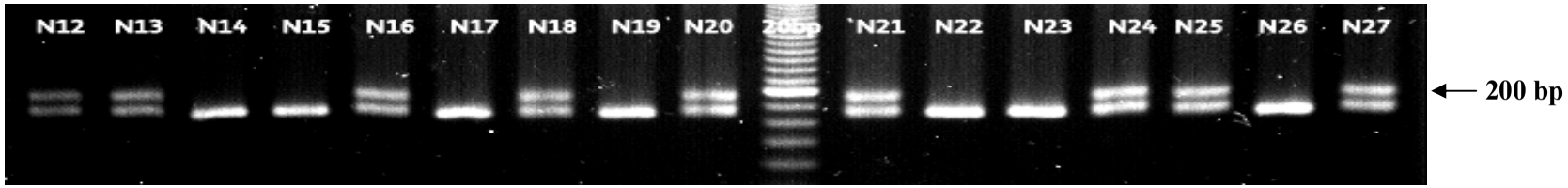

Figure 4.33: Polymorphic test on primer SFO-T112-H4 Negeri Sembilan

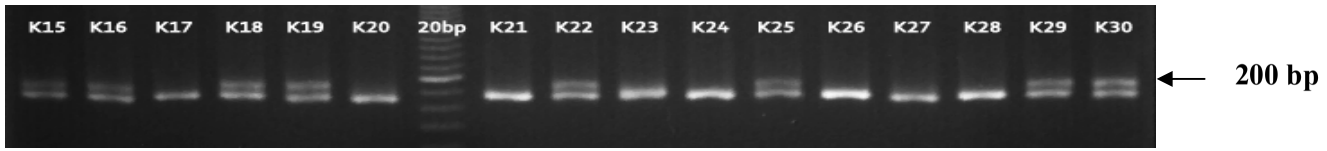

Figure 4.34: Polymorphic test on primer SFO-T112-H4 Kedah

4.3.1. Genotype frequencies

Different banding pattern will show the DNA polymorphism. Table 4.9 shows genotype frequencies of ten microsatellite loci in six populations of *Channa striata.* In primer SFO-T112-H6, it shows that all population has homozygotes bands except Negeri Sembilan population. However, this marker still consider as polymorphic marker. Primer SFO-T112-H4 and SFO-T112-H6F show homozygotes bands in Terengganu population while MFW 15 shows homozygote bands in Johor population.

From the analysis of MicroChecker, MFW 5 and MW15 show homozygote excess (homozygote expected is higher than the homozygote observed) and have null alleles. Therefore these alleles were excluded for further analysis.

Table 4.9 Genotype frequencies of 10 microsatellite loci in six populations of *Channa striata*

Locus/	Phenotype					
Population	Negeri Sembilan	Johor	Penang	Selangor	Terengganu	Kedah
SFO-T112-H4	167/167 (10) 167/197 (20)	167/167 (17) 167/197 (13)	167/167 (19) 167/197 (11)	167/167 (22) 167/197 (4)	167/167 (30)	167/167 (17) 167/197 (13)
SFO-T112-H6	268/278 (17) 278/278 (13)	278/278 (30)	278/278 (30)	278/278 (26)	278/278 (30)	278/278 (30)
SFO-T112-H6I	160/160 (16) 160/196 (14)	160/160 (11) 160/196 (12) 196/196 (7)	160/160 (19) 160/196 (11)	160/160 (11) 160/196 (15)	160/160 (16) 160/196 (14)	160/160 (10) 160/196 (20)
SFO-T112-H6F	256/256 (13) 256/296 (17)	256/256 (25) 256/296 (5)	256/256 (20) 256/296 (10)	256/256 (24) 256/296 (2)	256/256 (30)	256/256 (28) 256/296 (2)
SFO-BP8-H1	250/250 (16) 250/270 (12) 250/280 (2)	250/250 (27) 250/280 (3)	250/250 (17) 250/280 (13)	250/250 (10) 250/300 (16)	250/250 (26) 250/280 (4)	250/250 (21) 250/280 (9)

Table 4.9 Genotype frequencies of 10 microsatellite loci in six populations of *Channa striata* (continue)

MFW 5	192/192 (12) 192/210 (8) 210/210 (10)	176/176 (20) 176/192 (9) 192/192 (1)	226/226 (10) 246/246 (20)	226/226 (12) 246/246 (14)	246/246 (13) 246/276 (17)	210/210 (15) 210/226 (14) 226/226 (1)
MFW 1	168/168 (15) 168/182 (15)	176/176 (14) 176/188 (15) 180/188 (1)	172/172 (22) 172/182 (8)	172/172 (11) 172/188 (15)	182/182 (10) 182/210 (20)	168/168 (17) 168/182 (13)
MFW 7	176/176 (8) 176/194 (10) 194/194 (12)	176/188 (8) 188/188 (22)	188/194 (17) 194/194 (13)	246/246 (21) 246/268 (5)	194/194 (22) 194/204 (8)	246/246 (26) 246/274 (4)
MFW 15	241/259 (1) 259/259 (29)	151/151 (30)	151/151 (7) 171/171 (8) 171/181 (7) 181/181 (8)	231/251 (7) 251/251 (19)	167/167 (20) 167/181 (10)	171/181 (8) 181/181 (22)
MFW 2	167/167 (13) 167/181 (17)	167/167 (19) 167/181 (11)	167/167 (19) 167/181 (11)	167/167 (23) 167/181 (3)	167/167 (19) 167/181 (10) 181/181 (1)	

Table 4.10 shows the polymorphisms observed at the microsatellite loci amplified for the six populations of *Channa striata*. These results obtained from the genotype frequencies for ten loci in six populations. If only one allele for a locus in a population, it is consider monomorphic, while more than one allele considered as polymorphic for the population.

In primer SFO-T112-H6, it shows only one allele in all population except Negeri Sembilan population. However, this marker still consider as polymorphic marker. Primer SFO-T112-H4 and SFO-T112-H6F show one allele in Terengganu population while MFW 15 shows one allele in Johor population.

Two loci show monomorphic band in Johor population (SFO-T112-H6 and MFW15). Locus SFO-T112-H6 show monomorphic band in 3 populations (Penang, Selangor, and Kedah). Three loci show monomorphic bands in Terengganu population (SFO-T112-H6, and SFO-T112-H6F) (Table 4.10).

All ten loci show polymorphic in Negeri Sembilan population, 9 polymorphic loci in Penang, Selangor, and Kedah, 8 polymorphic loci in Johor population and the lowest number of polymorphic loci is in Terengganu population (Table 4.10).

Table 4.10 shows the polymorphisms observed at the microsatellite loci amplified for the six populations of *Channa striata*.

No	Locus	N9	Johor	Penang	Selangor	Terengganu	Kedah	Overall
1	SFO-T112-H4	P	P	P	P	M	P	P
2	SFO-T112-H6	P	M	M	M	M	M	P
3	SFO-T112-H6I	P	P	P	P	P	P	P
4	SFO-T112-H6F	P	P	P	P	M	P	P
5	SFO-BP8-H1	P	P	P	P	P	P	P
6	MFW5	P	P	P	P	P	P	P
7	MFW 1	P	P	P	P	P	P	P
8	MFW 7	P	P	P	P	P	P	P
9	MFW15	P	M	P	P	P	P	P
10	MFW 2	P	P	P	P	P	P	P
	Total number of monomorphic loci	0	2	1	1	3	1	
	Total number of polymorphic loci	10	8	9	9	7	9	

P: Polymorphic, M: Monomorphic

4.3.2. Number of alleles

The number of alleles per locus ranged from two to eight with an average of 4.20 (Table 4.11). F_{IS} values greater than zero indicate that the population is inbreeding.

Table 4.11 Number of observed and expected alleles and the value of F_{IS} for all loci.

Locus	No of observed allele	No of expected allele	F_{IS} value
SFO-T112-H4	2.000	1.4016	-0.2996
SFO-T112-H6	2.000	1.1012	-0.3806
SFO-T112-H6I	2.000	1.6857	-0.2266
SFO-T112-H6F	2.000	1.2249	-0.2269
SFO-BP8-H1	4.000	1.4208	-0.2492
MFW 5	6.000	5.0507	-0.1068
MFW 1	7.000	4.8835	-0.3267
MFW 7	7.000	3.7240	-0.0432
MFW 15	8.000	5.7051	0.2281
MFW 2	2.000	1.4309	-0.2053
Mean	4.200	2.7628	

4.3.3. Heterozygosity

The estimated values of the observed and expected heterozygosity for all populations are show in Table 4.12, the highest mean observed heterozygosity found in the Negeri Sembilan population with a value of 0.4433 while Johor population had the lowest value of 0.2567.

Negative F_{IS} values indicate an excess of heterozygosity in the population. Three populations show an excess of heterozygosity i.e. Selangor, Terengganu, and Kedah. F_{IS} value greater than zero indicate a deficit of heterozygosity. Population Johor, Negeri Sembilan and Penang show a deficit of heterozygosity due to inbreeding (Table 4.12).

Six loci for the six populations conformed to Hardy Weinberg expectation ($P>0.05$) with the exception of loci T112-H6I shows deviation from Hardy Weinberg Equilibrium (HWE) law ($P<0.05$) in Kedah population, MFW5 shows deviation from HWE in Negeri Sembilan and Penang populations, MFW1 shows deviation from HWE in Terengganu population, and MFW15 shows deviation from HWE in Penang population.

Significant deviation from Hardy Weinberg for loci MFW5 and MFW15 might be cause by the presence of null allele while in loci T112-H6I and MFW1 there were maybe presence of migration, mutation, non-random mating, selection, and small population size.

The highest mean observed number of alleles found in the Negeri Sembilan population with a value of 2.10, the lowest mean observed number of allele found in Terengganu population with a value 1.70 (Table 4.12).

Effective number of allele is the estimation of the reciprocal of homozygosity (Hartl and Clark, 1989). The highest mean effective allele number was 1.659 in the Negeri Sembilan population and the lowest was 1.313 in the Terengganu population (Table 4.12). The summary of the allele sizes and allele frequencies for the polymorphic microsatellite loci show in appendix 2.

Table 4.12: Level of heterozygosity of six populations of *Channa striata*

Locus		N9	Johor	Penang	Selangor	Terengganu	Kedah
SFO-T112-H4	**Ho**	0.67	0.43	0.37	0.15	-	0.43
	He	0.45	0.35	0.30	0.14	-	0.35
	F_{IS}	-0.49	-0.26	-0.21	-0.06	-	-0.26
	P	0.01	0.29	0.55	1.00	-	0.29
	na	2.00	2.00	2.00	2.00	1.00	2.00
	ne	1.80	1.51	1.43	1.17	1.00	1.51
SFO-T112-H6	**Ho**	0.57	-	-	-	-	-
	He	0.41	-	-	-	-	-
	F_{IS}	-0.38	-	-	-	-	-
	P	0.07	-	-	-	-	-
	na	2.00	1.00	1.00	2.00	1.00	1.00
	ne	1.68	1.00	1.00	1.00	1.00	1.00

Table 4.12: Level of heterozygosity of six populations of *Channa striata* (continue)

SFO-T112 H6I	**Ho**	0.47	0.40	0.37	0.58	0.47	0.67
	He	0.36	0.50	0.30	0.42	0.36	0.45
	F_{IS}	-0.29	0.20	-0.21	-0.39	-0.29	-0.49
	P	0.30	0.46	0.55	0.07	0.29	0.01
	na	2.00	2.00	2.00	2.00	2.00	2.00
	ne	1.56	1.97	1.43	1.70	1.56	1.80
SFO-T112-H6F	**Ho**	0.57	0.17	0.33	0.08	-	0.07
	He	0.41	0.16	0.28	0.08	-	0.07
	F_{IS}	-0.38	-0.07	-0.18	-0.02	-	-0.02
	P	0.07	1.00	0.56	1.00	-	1.00
	na	2.00	2.00	2.00	2.00	1.00	2.00
	ne	1.68	1.18	1.39	1.08	1.00	1.07

Table 4.12: Level of heterozygosity of six populations of *Channa striata* (continue)

SFO-BP8-H1	**Ho**	0.47	0.10	0.43	0.62	0.13	0.30
	He	0.38	0.10	0.35	0.43	0.13	0.26
	F_{IS}	-0.24	-0.04	-0.26	-0.43	-0.05	-0.16
	P	0.42	1.00	0.29	0.06	1.00	1.00
	na	3.00	2.00	2.00	2.00	2.00	2.00
	ne	1.59	1.11	1.51	1.74	1.14	1.34
MFW 5	**Ho**	0.27	0.30	0.67	0.54	0.43	0.47
	He	0.51	0.30	0.45	0.40	0.35	0.40
	F_{IS}	0.48	0.02	-0.49	-0.35	-0.26	-0.18
	P	0.01	1.00	0.01	0.13	0.29	0.64
	na	2.00	2.00	2.00	2.00	2.00	2.00
	ne	1.99	1.43	1.80	1.65	1.51	1.64

Table 4.12: Level of heterozygosity of six populations of *Channa striata* (continue)

MFW 1	**Ho**	0.50	0.53	0.27	0.58	0.67	0.43
	He	0.38	0.42	0.24	0.42	0.45	0.35
	F_{IS}	-0.32	-0.27	-0.14	-0.39	-0.49	-0.26
	P	0.14	0.05	1.00	0.07	0.01	0.29
	na	2.00	3.00	2.00	2.00	2.00	2.00
	ne	1.60	1.71	1.30	1.70	1.80	1.51
MFW 7	**Ho**	0.33	0.27	0.57	0.19	0.27	0.13
	He	0.50	0.24	0.41	0.18	0.24	0.13
	F_{IS}	0.34	-0.14	-0.38	-0.09	-0.14	-0.05
	P	0.13	1.00	0.07	1.00	1.00	1.00
	na	2.00	2.00	2.00	2.00	2.00	2.00
	ne	1.97	1.30	1.68	1.21	1.30	1.14

Table 4.12: Level of heterozygosity of six populations of *Channa striata* (continue)

MFW 15	**Ho**	0.03	-	0.23	0.27	0.33	0.27
	He	0.03	-	0.66	0.24	0.28	0.24
	F_{IS}	0.00	-	0.65	-0.14	-0.18	-0.14
	P	1.00	-	0.00	1.00	0.56	1.00
	na	2.00	1.00	3.00	2.00	2.00	2.00
	ne	1.03	1.00	2.87	1.30	1.39	1.30
MFW 2	**Ho**	0.57	0.37	0.37	0.12	0.37	0.33
	He	0.41	0.30	0.30	0.11	0.30	0.33
	F_{IS}	-0.38	-0.21	-0.21	-0.04	-0.21	-0.02
	P	0.07	0.55	0.55	1.00	0.55	1.00
	na	2.00	2.00	2.00	2.00	2.00	2.00
	ne	1.68	1.43	1.43	1.12	1.43	1.47

Ho: Observed heterozygosity, He: Expected heterozygosity, F_{is}: within-population inbreeding coefficients, P: Probability value, na: observed number of allele, ne: effective number of allele

4.3.4. Null alleles

The checking of the microsatellite data for null alleles was performed using the Micro-checker (version 2.2.3) software. The result revealed that two loci, MFW5 and MFW15 show evidence for the presence of null alleles in Negeri Sembilan and Penang population respectively (Table 4.13 and 4.14). Therefore, MFW5 and MFW15 were excluded from further analysis. All loci show no evidence of stuttering bands and allele drop out.

After removing loci that show evidence of null allele, it shown that the percentage of consensus tree UPGMA constructed using Phylip software is higher (Figure 4.36). In AMOVA analysis, the percentage of genetic differentiation within population also increases (Table 4.19). However, there were no changes in genetic distance (Table 4.17).

Table 4.13: Locus MFW5 shows null allele in Negeri Sembilan population.

Locus	Stuttering Bands	Allele Drop Out	Null Alleles
SFO-T112-H4	NO	NO	NO
SFO-T112-H6	NO	NO	NO
SFO-T112-H6I	NO	NO	NO
SFO-T112-H6F	NO	NO	NO
SFO-BP8-H1	NO	NO	NO
MFW5	NO	NO	YES
MFW1	NO	NO	NO
MFW7	NO	NO	NO
MFW15	NO	NO	NO
MFW2	NO	NO	NO

Table 4.14: Locus MFW15 shows null allele in Penang population.

Locus	Stuttering Bands	Allele Drop Out	Null Alleles
SFO-T112-H4	NO	NO	NO
SFO-T112-H6	NO	NO	NO
SFO-T112-H6I	NO	NO	NO
SFO-T112-H6F	NO	NO	NO
SFO-BP8-H1	NO	NO	NO
MFW5	NO	NO	NO
MFW1	NO	NO	NO
MFW7	NO	NO	NO
MFW15	NO	NO	YES
MFW2	NO	NO	NO

4.3.5. Linkage Disequilibrium (LD)

Two loci are said to be in linkage disequilibrium if their respective alleles do not associate independently in the studied population. Based on analysis using Popgene version 1.31 and Genepop version 1.2.software, no linkage disequilibrium was found in each pair of loci in all population except two loci SFO-T112-H6 and SFO-T112-H6I ($P<0.05$). Data analysis to determine linkage disequilibrium based on P-value for each locus pair across all populations is shown in the Table 4.15.

Table 4.15: P-value for locus pair in all populations

Locus pair		Chi2	df	P-Value
T112-H4	& T112-H6	0	2	1
T112-H4	& T112-H6I	10.615	10	0.388289
T112-H6	& T112-H6I	6.795407	2	0.03345
T112-H4	& T112-H6F	8.940859	10	0.537725
T112-H6	& T112-H6F	2.510462	2	0.28501
T112-H6I	& T112-H6F	20.42124	10	0.025511
T112-H4	& BP8-H1	15.98522	10	0.100056
T112-H6	& BP8-H1	1.494869	2	0.47358
T112-H6I	& BP8-H1	8.407371	12	0.752541
T112-H6F	& BP8-H1	11.92749	10	0.289939
T112-H4	& MFW5	6.503598	10	0.771329
T112-H6	& MFW5	0.414491	2	0.81282
T112-H6I	& MFW5	15.93095	12	0.194419
T112-H6F	& MFW5	8.137283	10	0.615429
BP8-H1	& MFW5	14.33195	12	0.280015
T112-H4	& MFW1	8.203093	10	0.609007
T112-H6	& MFW1	1.525938	2	0.46628
T112-H6I	& MFW1	7.49957	12	0.822914
T112-H6F	& MFW1	6.570632	10	0.765262
BP8-H1	& MFW1	6.990172	12	0.858262
MFW5	& MFW1	12.98079	12	0.370438
T112-H4	& MFW7	4.647779	10	0.913432
T112-H6	& MFW7	1.27342	2	0.52903
T112-H6I	& MFW7	12.47553	12	0.408283
T112-H6F	& MFW7	13.45274	10	0.199448
BP8-H1	& MFW7	4.378971	12	0.975587
MFW5	& MFW7	13.74294	12	0.317424

MFW1	& MFW7	1.912015	12	0.999529
T112-H4	& MFW15	13.32413	8	0.101173
T112-H6	& MFW15	1.6763	2	0.43251
T112-H6I	& MFW15	13.1381	10	0.216052
T112-H6F	& MFW15	5.394176	8	0.714734
BP8-H1	& MFW15	5.962868	10	0.818373
MFW5	& MFW15	2.021285	10	0.996174
MFW1	& MFW15	8.681114	10	0.562611
MFW7	& MFW15	13.0401	10	0.221439
T112-H4	& MFW2	3.154511	10	0.97755
T112-H6	& MFW2	3.999487	2	0.13537
T112-H6I	& MFW2	7.602203	12	0.815394
T112-H6F	& MFW2	1.351616	10	0.999328
BP8-H1	& MFW2	5.342124	12	0.945565
MFW5	& MFW2	11.2087	12	0.511123
MFW1	& MFW2	6.497359	12	0.888968
MFW7	& MFW2	8.25103	12	0.765218
MFW15	& MFW2	12.19423	10	0.272268

4.3.6. Genetic distance

The genetic identity between Penang and Terengganu was 0.8152 and the genetic distance between these two populations was 0.2043 (Table 4.16). This genetic distance suggested the Penang and Terengganu were the most related populations. The highest genetic distance was between the Negeri Sembilan and Selangor populations (Table 4.16).

Clustering analysis using consensus tree of the populations was constructed based on Nei's (1978) unbiased genetic distance using the software Phylip (version 3.67) (Figure 4.35). The consensus tree divided these six populations of the *Channa striata* into 2 major clusters. Populations Kedah, Negeri Sembilan, Johor, Selangor, and Terengganu belonged to one group while the Penang population clustered by itself. The first major cluster subdivided into four groups. The populations from Kedah and Negeri Sembilan belonged to a single group and the Johor, Selangor, and Terengganu belonged to three different groups.

Table 4.16: The values of genetic distance between six populations of *Channa striata*

Pop ID	N.9	Johor	Penang	Selangor	Terengganu	Kedah
N.9	****	0.5888	0.6702	0.5679	0.6802	0.7172
Johor	0.5297	****	0.6876	0.6116	0.6202	0.6052
Penang	0.4002	0.3745	****	0.8007	0.8152	0.6910
Selangor	0.5659	0.4916	0.2223	****	0.6825	0.7216
Terengganu	0.3854	0.4777	0.2043	0.3820	****	0.6663
Kedah	0.3324	0.5023	0.3697	0.3263	0.4060	****

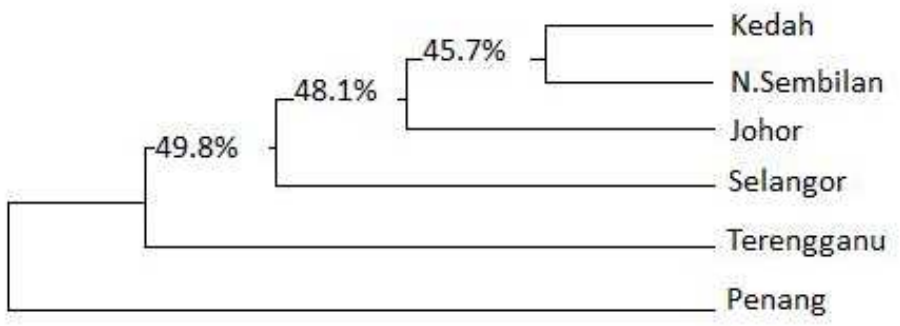

Figure 4.35: Consensus tree generated using 10 microsatellite loci

After removing null alleles from loci MFW5 and MFW15, population Penang and Terengganu still gives the highest genetic distance. The genetic identity between Penang and Terengganu was 0.8874 and the genetic distance between these two populations was 0.1195 (Table 4.17). This genetic distance suggested the Penang and Terengganu were the most related populations. The highest genetic distance was between the Negeri Sembilan and Selangor populations (Table 4.17). This is not support the geographical location.

Clustering analysis of UPGMA (unweighted pair-group method using the arithmetic average) consensus tree of the populations was constructed based on Nei's (1972) unbiased genetic distance using the software Phylip (version 3.67), with bootstrap value 1000 (Figure 4.36). UPGMA is clustering method that use average linkage method and minimizes the inter-group distance by taking the average pairwise distance among all individuals of the sample. After removing loci that show evidence of null allele, the percentage of the branch is higher. The consensus tree divided these six populations of the *Channa striata* into 2 major clusters. Populations Kedah, Negeri Sembilan, Selangor, Johor and Terengganu belonged to one group while the Penang population clustered separately. It

is following the geographical location. The first major cluster subdivided into four groups. The populations from Kedah and Selangor belonged to a single group and the Johor, Negeri Sembilan, and Terengganu belonged to another group. All population show more than 50% clustering.

Table 4.17: The values of genetic distance between six populations of *Channa striata* (after removing null allele)

Pop ID	N.9	Johor	Penang	Selangor	Terengganu	Kedah
N.9	****	0.7456	0.8240	0.7179	0.8569	0.8450
Johor	0.2935	****	0.7975	0.7655	0.7736	0.7593
Penang	0.1935	0.2262	****	0.8539	0.8874	0.7571
Selangor	0.3315	0.2672	0.1580	****	0.7715	0.8819
Terengganu	0.1544	0.2567	0.1195	0.2595	****	0.7887
Kedah	0.1684	0.2754	0.2783	0.1257	0.2374	****

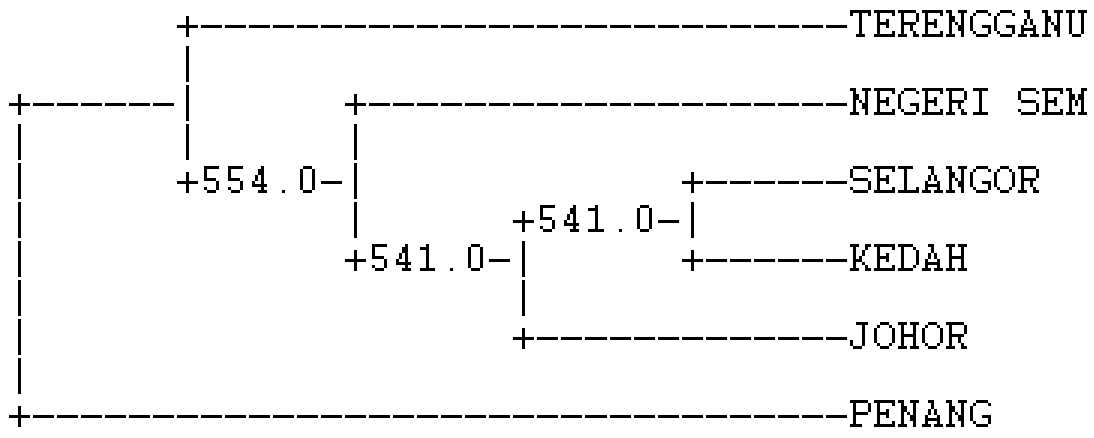

Figure 4.36: Consensus tree generated using 8 microsatellite loci

4.3.7. Analysis of Molecular Variance (AMOVA)

An Analysis of Molecular Variance (AMOVA) done with the ARLEQUIN package (Schneider *et al.*, 2000) for measuring variance within populations, and among populations, applying the estimator of Weir and Cockerham (1984).

The results of the analyses of molecular variance (AMOVA) given in Table 4.18. Based on 10 polymorphic markers investigated, 45.80% of the variations were among populations and 54.20% of the variations were within populations. An AMOVA revealed that the greatest amount of variance occurs within populations (Table 4.18).

Table 4.18: Amova design and results

Source of variation	Sum of squares	Variance components	Percentage of variation
Among populations	350.633	1.17230 Va	45.80
Within populations	480.006	1.38730 Vc	54.20
Total	830.639	2.55960	
		Fixation Index	FST: 0.45800

The results of the analyses of molecular variance (AMOVA) given in Table 4.19. Based on 8 polymorphic markers investigated, 35.27% of the variations were among populations and 64.73% of the variations were within populations. An AMOVA revealed that the greatest amount of variance occurs within populations (Table 4.19). From the F_{ST} value, the difference is significant.

Table 4.19: Amova design and results (after removing null allele)

Source of variation	Sum of squares	Variance components	Percentage of variation
Among populations	175.584	0.58072 Va	35.27
Within populations	368.751	1.06576 Vb	64.73
Total	544.335	1.64647	
		Fixation Index	FST: 0.35270

4.3.8. Pairwise F_{ST}

Table 4.20 shows the population pairwise F_{ST}, which estimated distances, according to the pairwise differences method. The highest divergences value was between the Johor and Terengganu populations (0.56780) and the lowest one was between Penang and Selangor populations (0.32795).

Table 4.20: Population pairwise F_{ST} distance

	N.9	Johor	Penang	Selangor	Terengganu	Kedah
N.9	0.00000					
Johor	0.47771	0.00000				
Penang	0.37003	0.44073	0.00000			
Selangor	0.48229	0.55192	0.32795	0.00000		
Terengganu	0.43164	0.56780	0.33257	0.52029	0.00000	
Kedah	0.37544	0.54689	0.42626	0.45467	0.52227	0.00000

Table 4.21 shows the population pairwise F_{ST}, which estimated distances, according to the pairwise differences method. These results were obtained after removing null allele loci. The highest divergences value was between the Selangor and Terengganu populations (0.47159) and the lowest one was between Negeri Sembilan and Kedah populations (0.24596).

Table 4.21: Population pairwise F_{ST} distance (after removing null allele)

	N.9	Johor	Penang	Selangor	Terengganu	Kedah
N.9	0.00000					
Johor	0.33653	0.00000				
Penang	0.25241	0.35634	0.00000			
Selangor	0.37461	0.42285	0.30228	0.00000		
Terengganu	0.26118	0.44231	0.27303	0.47159	0.00000	
Kedah	0.24596	0.41907	0.41083	0.27737	0.43752	0.00000

4.3.9. Structure analysis

According to degree of admixture (alpha value), K= 3 gives the nearest value to zero. K is the true number of clusters. Figure 4.37 shows a bar plot represent population structure of *Channa striata* (K=3). Cluster 1 show clustering of population Johor and Selangor, Cluster 2 show clustering of population Negeri Sembilan and Kedah, while population Penang and Terengganu in other cluster. In Penang population it is shown that hybridization occur from population Johor and Selangor. This is maybe because of geneflow, migration and the environmental adaptation.

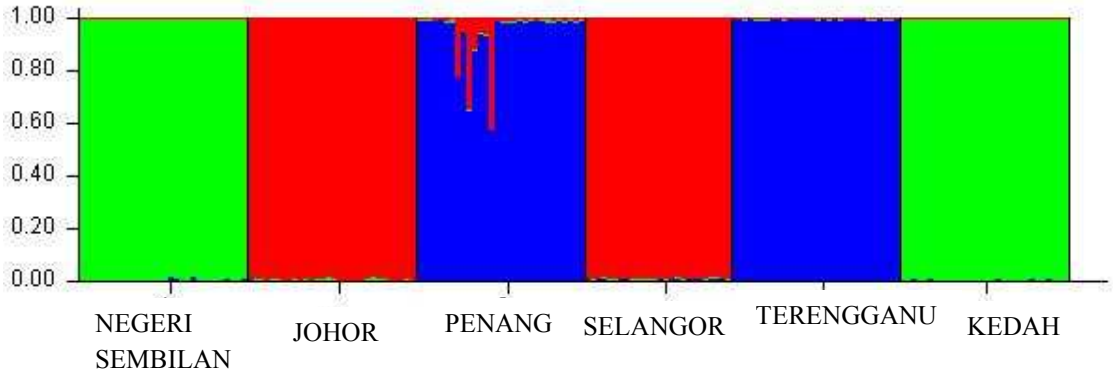

Figure 4.37: Bar plot represent population structure of *Channa striata* (K=3)

After removing null allele loci, the degree of admixture (alpha value), K= 6 gives the nearest value to zero. K is the true number of clusters. Figure 4.38 shows a bar plot represent population structure of *Channa striata* (K=6). Each population clustering separately and there was hybridization occur in each population. Negeri Sembilan and Penang show that hybridization from Terengganu population may occur; this is maybe because of transportation of fish by fisheries department. The Pleistocene drainage basins that linked sites on the Sunda Shelf that are today geographically isolated may also have caused gene flow among some populations of *Channa striata* in the past.

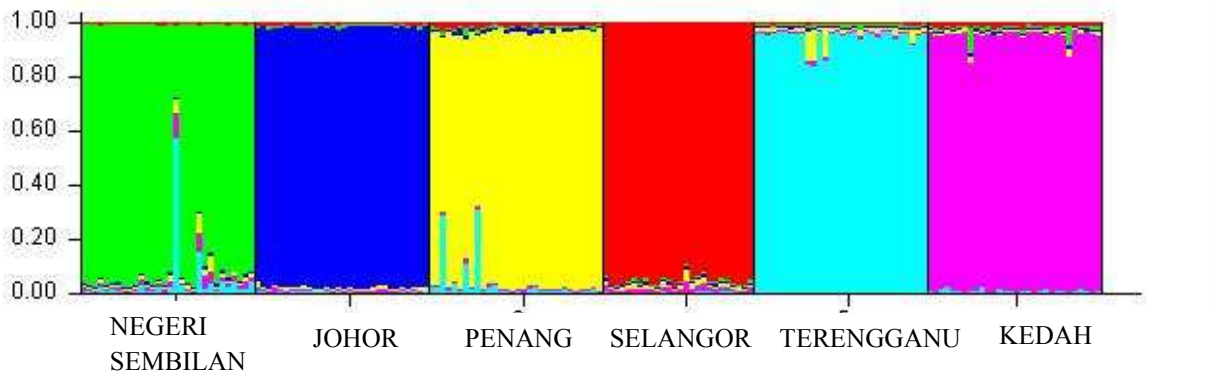

Figure 4.38: Bar plot represent population structure of *Channa striata* (K=6)