

CHAPTER 3: METHODOLOGY

3.1 Collection of samples

A sample of 176 individuals from six populations of *Channa striata* in Peninsular Malaysia from Pulau Pinang, Negeri Sembilan, Selangor, Johor, Kedah and Terengganu were collected by the Pusat Penyelidikan Perikanan Air Tawar were used in this study. The sampling sites listed in Table 3.1 and Figure 3.1 shows the locations of the sampling sites.

Table 3.1 Sampling sites, location and quantity of samples

No.	State/ Place	Location	Coordinate	Quantity of samples
1.	Kedah	Kubur Panjang, Paddy fields	Latitude: 6.11427 Longitude: 100.54928	30
2.	Pulau Pinang	Permatang Kuang, Paddy fields	Latitude: 5.40665 Longitude: 100.36783	30
3.	Selangor	Tanjung Karang, Small stream	Latitude: 3.44285 Longitude: 101.18340	26
4.	Negeri Sembilan	Sikamat, paddy fields	Latitude: 2.75925 Longitude: 101.94901	30
5.	Johor	Batu Pahat dam	Latitude: 1.85860 Longitude: 102.93777	30
6.	Terengganu	Terengganu river	Latitude: 5.37950 Longitude: 103.08884	30



Figure 3.1 Location of the sampling sites
 Cited from: <http://www.doe.gov.my>

3.2 Isolation of Genomic DNA

DNA extractions were performed using a GF1 DNA extraction kit (Yeastern Biotech Co., Ltd); as prior to the protocol described by the manufacturer. This kit provides a faster, easy, high yield, pure and reliable method for extracting DNA. The extracts of genomic DNA were stored at -20 °C for downstream application.

The tissue samples (each 0.02 g) were cut into small pieces with a clean scalpel. Buffer TL (250 µL) and Proteinase K (20 µL) (20 mg/ml) were added to the sample. The mixture was completely resuspended by vigorous vortex to obtain a homogeneous solution. Lysis Enhancer (12 µL) was added and immediately mixed and then incubate at 65°C for 3 hours in a waterbath and occasionally mixed during incubation to ensure thorough digestion of the sample.

RNase A (20 µL) (DNase-Free, 20mg/ml) was added to the mixture to have RNA-free DNA. Buffer TB (600 µL) was added and thoroughly mixed by pulsed-vortexing until a homogenous solution was obtained. Then, the mixture was incubated for 10 minutes at 65°C. After that, 200 µL of absolute ethanol was added. In order to prevent any uneven precipitation of nucleic acid due to high local ethanol concentration, the mixture was immediately mixed by vigorously pulse-vortexing until a homogeneous solution was obtained.

Approximately 600 µL of sample was transferred into a column assembled in a clean collection tube and centrifuged at 5,000 x g for 1 min. The flow through was discarded and the remaining samples were repeatedly transferred into the column until the sample finished. The column was washed with 750µL Wash Buffer and centrifuged at 5,000 x g for 1 minute. The flow through was discarded and the column was washed again. Then, the column was centrifuged at 10,000 x g for 1 minute to remove all traces of

ethanol. After that, the column was placed into a clean microcentrifuge tube that was provided.

Finally, 150 µL of preheated Elution Buffer was directly added onto the column membrane and left to stand at room temperature for 2 minutes. The tubes were centrifuged at 5,000 x g for 1 minute to elute DNA which is later stored at -20°C.

3.3 RAMs Primers

Several degenerate Randomly Amplified Microsatellite (RAM) primers were used to generate a genomic library enriched for microsatellites through randomly amplified fragments from the *M. rosenbergii* genome. These primers were specifically designed to include a longer tandem repeat motif (n = 10) to allow capturing of longer repeat microsatellites. RAMs primers used were VJ2, BP8, BP11, T79112, and T79110 (Table 3.2). These RAMs primers were used to design microsatellite primers specific to amplify the DNA of *Channa striata*.

Table 3.2: List of Random Amplified Microsatellites (RAMs) primers

Primer	Sequence (5' to 3')	Reference
T79112	GCA CAT GCA R(TG) ₇	<i>See Leng Min, 2008</i>
T79110	GAT GCT GAT R(CA) ₇	<i>See Leng Min, 2008</i>
BP11	KKKYHYHY(CAG) ₅	<i>Hoh Boon Peng, 2005</i>
BP8	(K) ₂ (YH) ₃ Y(GTT) ₅	<i>Hoh Boon Peng, 2005</i>
VJ2	NNNKKVRVRV(CTC) ₅	<i>Vijay Kumar, 2003</i>

Where K = G/T, V = G/C/A, R = G/A, N = A/C/G/T, H = A/C/T, Y = T/C,

3.4 PCR Amplification

Polymerase chain reactions were carried out in 10 μL reaction mixtures of 1x PCR buffer (Promega, USA), 2.0 to 3.0 mM of MgCl_2 (Promega, USA), 1.0 unit of Taq polymerase (Promega, USA), 0.25 mM of each dNTPs (Promega, USA), 0.5 mM of each pair primer (Medigene), 20 ng of genomic DNA and deionised water. Amplifications were performed in a BIORAD thermal cycler (Bio-Rad Laboratories, Inc.) with an initial 3 minutes of predenaturation at 96 $^\circ\text{C}$, followed by 40 cycles of denaturation at 94 $^\circ\text{C}$ for 10 seconds, an appropriate annealing temperature for 10 seconds and extension at 72 $^\circ\text{C}$ for 30 seconds. A final extension step at 72 $^\circ\text{C}$ for 7 minutes was included for the attachment of dATP at the 3' terminal that were compatible for direct ligation of products into yT&A cloning vector (Yeastern Biotech Co., Ltd.). Electrophoresis (68V, 150amp, 60min) was then carried out on 2 % agarose gel in 1 X TBE Buffers in order to confirm the presence of the PCR products. Eventually, the resultant products were ready for cloning.

3.5 Cloning

The amplified selected DNA fragments were then cloned into yT&A cloning vector (Yeastern Biotech Co., Ltd.) according to the manufacturer's instructions with minor modification. yT&A cloning vector was ligated with 5 μL of fresh PCR products in 5.0 μL of ligation mixture (Yeastern Biotech Co., Ltd.). First, yT&A cloning vector was centrifuged and ligation buffers were vortex vigorously. The ligation mixtures contained 2 μL of ligation buffer, 2 μL of yT&A cloning vector, 5 μL of fresh PCR product, and 1 μL of DNA ligase in a total volume of 10 μL . The solution was mixed by pipetting and incubated overnight for maximum transformants at 4 $^\circ\text{C}$. The ECOS competent *E.coli* (Yeastern Biotech Co., Ltd.) cells were thawed for about 20 seconds under tap water. After that, the cells were placed on ice. Ligation reaction (6 μL) was added directly to the ECOS

Competent Cells followed by incubation on ice for 5 minutes. The cells were heat shocked at 42 °C for exactly 35 seconds and the transformation mixture was spread and grown on a pre-warmed LB agar media (Difco™ Lennox) containing 50 µg/mL ampicillin (Bio Basic Inc, USA) and 20 µL of 20 mg/mL of X-gal (5Prime). The plates were incubated overnight at 37 °C. After obtaining culture growth on LB agar media, white colonies that were expected to contain PCR products were picked and transferred onto new LB agar plates to enhance the growth for 2 days at 37°C.

3.6 Colony PCR

Colony PCR was done to check successful ligation and transformation. First, 10 µL of ddH₂O was pipette into 2ml tubes. Then, the colonies were touched with sterile stick and make suspensions. The suspensions were heated for 10 minutes at 94°C to lyse the cells. PCR condition is same as provided by manufacturer (Yeastern Biotech Co., Ltd.).

Polymerase chain reaction was carried out in a 20 µL reaction mixture of 1x PCR buffer (Promega, USA), 2.0 to 3.0 mM of MgCl₂ (Promega, USA), 1.0 unit of Taq polymerase (Promega, USA), 0.25 mM of each dNTPs (Promega, USA), 0.5 mM of each pair primer (Medigene), 20 ng of genomic DNA and deionised water. Amplifications were performed in BIORAD thermal cycler (Bio-Rad Laboratories, Inc.) with an initial 2 minutes of predenaturation at 95 °C, followed by 40 cycles of denaturation at 95 °C for 1 minute, 60°C annealing temperature at 60°C for 1 minute and extension at 72 °C for 2 minutes. A final extension step at 72 °C for 10 minutes was included for the attachment of dATP at the 3' terminal that were compatible for direct ligation of products into yTA cloning vector (Yeastern Biotech Co., Ltd.). Electrophoresis was then carried out on 2 % agarose gel in order to confirm the presence of the PCR products.

Table 3.3: PCR condition for colony PCR

Solution	Volume (μL)
1 x PCR Buffer	4.0
MgCl ₂	0.8
dNTPs	0.4
Primer M13(F)	0.4
Primer M13(R)	0.4
Taq DNA Polymerase	0.1
H ₂ O	12.9
Colony DNA	1
Total	20

3.7 Plasmid Extraction

Cultures for plasmid isolation were prepared by inoculating 5 mL LB broth (Difco™ Lennox) containing 50 μ g/mL ampicillin from freshly streaked, well-isolate colonies grown on LB agar using a sterilized stick and shaken at 37 °C for 12-16 hours constantly at 250 rpm. For storage, all recombinant clones were subjected to long term storage by mixing 850 μ L of culture with 150 μ L of glycerol and stored at -80 °C. Ten recombinant clones were randomly selected for plasmid extraction. The plasmid extractions were performed according to the protocol provided by the manufacturer (HiYield Plasmid Mini Kit, Yeastern Labs, 1994).

Bacterial pellets were collected by centrifugation of fresh bacterial cultures at 14,000 x g for 1 minute. The supernatant was removed and 200 μ L of PD1 Buffer (with RNase A added) was added. The pellet was completely resuspended by vigorous vortexing until the lysate was a homogenous solution with no apparent cell clumps visible. Then, 200 μ L of

PD2 buffer was added and gently mixed by inverting the tube 10 times. Clear and non-viscous lysate was obtained by 2 minute incubation at room temperature. PD3 Buffer (300µl) was added and the tube was immediately inverted 10 times. Then, the tubes were centrifuged at full speed (14,000 x g). A PD Column was placed in a Collection Tube. Lysate was then transferred to the PD Column and was centrifuged at 8.0 g for 30 seconds. The flow-through was discarded and the PD Column was placed back in the Collection Tube. The 600 µL of Wash Buffer (ethanol added) was then added to the PD Column followed by centrifugation for 30 seconds at 8.0 g. Flow-through was discarded and centrifugation was carried out for an additional 30 minutes at full speed to dry the PD Column. Dried PD Column was transferred in a clean microcentrifuge tube. The plasmid was eluted and resuspend into a clean microcentrifuge tube with 30 µL of Elution Buffer and stored at -20 °C.

3.8 DNA Sequencing

Cloned PCR products were sequenced in an ABI PRISM 377 DNA sequencer using a BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, USA). The universal primer, M13 (-40), 5' GTTTTCCCAGTCACGAC 3' was used for sequencing.

3.9 Submission of DNA Sequences to GenBank

DNA sequences were first edited to remove vector sequences by using free software VecScreen at www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html before submission to GenBank. Submission was done using the stand-alone software tool Sequin developed by the NCBI and submitted through the internet at www.ncbi.nlm.nih.gov.

3.10 Design of Primers Flanking Microsatellite Regions

The sequences of PCR products were screened for microsatellite sequence and primer pairs flanking each of the microsatellite regions were design by using a free online primer designing software, PRIMER3, http://www.genome.wi.mit.edu/genome_software/other/primer3.html with some considerations as follows:

1. Primers are typically 40 – 60 % GC content with length between 18 to 24 nucleotides.
2. Complimentary sequences of the primer pairs and GC-rich 3' end are avoided.
3. Primers with G or C residues in the 5' and central region are designed.
4. Sequences with the potential to form internal secondary structures are avoided.

3.11 Microsatellite Amplification

3.11.1. Standard PCR

The new microsatellite primers were then optimized for use with *Channa striata*. The criteria of optimization involved annealing temperature, concentration of DNA, quantity of *Taq* polymerase, and concentration of MgCl₂. The PCR mixture contained 1 x PCR buffer, 1.5 unit of *Taq* polymerase, 0.25 mM of each dNTPs, 0.5mM of primer, 20 ng of template DNA and MgCl₂ ranging from 2.5mM in a final volume of 10 µL. The PCR profile for each primer consisted of predenaturation at 96 °C for 3 minutes, 40 cycles of 94 °C denaturation for 10 seconds, 10 seconds at optimum annealing temperature, 30 seconds of DNA extension at 72 °C follow by 7 minutes final extension at 72 °C. The PCR product was then electrophoresed using a 4 % (w/v) Metaphor Agarose gel (LONZA, USA) in 1 X TBE Buffer, followed by ethidium bromide staining (0.1 mg/mL) and visualize over ultra-violet light.

Table 3.4: Standard PCR profiles for optimization of the primers.

Steps	Temperature (°C)	Times(minute)
Initial Denaturation	96°C	3:00
Denaturation	94°C	0:10
Annealing	53°C - 67°C	0:10
Elongation	72°C	0:30
Final Elongation	72°C	7:00
Storage	4°C	∞

3.11.2 Touchdown PCR

Touchdown PCR is a modification of conventional PCR that will avoid nonspecific amplification. The annealing temperature is higher than the target optimum in early PCR cycles. The annealing temperature is decreased by 1°C every cycle until a specified annealing temperature is reached. The touchdown temperature is then used for the remaining number of cycles. This is to increase the correct product over any non-specific product.

A series of touchdown PCR were used for the other five microsatellite primers reported in common carp (Crooijmans *et al.*, 1997) which were MFW1, MFW2, MFW5, MFW7, and MFW15. The PCR reactions mixtures were same as described for standard PCR but the PCR profile is different (Table 3.5). The sequences of these primers which amplified polymorphic bands in all six populations were shown in Table 4.7.

Table 3.5: Touchdown PCR profile

Step	Temperature (°C)	Time
1. Predenaturation	95	5 min
2. Denaturation	94	40 sec
3. Annealing	58	35 sec
4. Extension	72	35 sec
5. Go to step 2 for 20 times		
6. Denaturation	94	40 sec
7. Annealing	60	35 sec
8. Extension	72	35sec
9. go to step 6 for 20 times		
10. Final Extension	72	7 min

3.12 Blast Analysis of Microsatellite Markers

Expressed sequence tags (ESTs) of each of the new microsatellite markers were determined by comparing the nucleotide sequences to nucleotide sequences in the NCBI database in order to identify sequences that share regions of homology using the BLASTn (Basic Local Alignment Search Tool) algorithm available at <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>.

3.13 Data Analysis

3.13.1. Microchecker software

Micro-Checker is a Windows application that checks for microsatellite null alleles and scoring errors. It also provides null allele estimates, and adjusts allele and genotypes frequencies.

3.13.2. Convert (Version 1.31) software

This program facilitates the conversion of diploid genotypic data files into formats that can be directly read by a number of commonly used population genetic computer programs: gda, genepop, arlequin, popgene, microsat, phylip and stucture. , convert can be used to produce a table of allele frequencies in a convenient format, allowing the visual comparison of allele frequencies across populations. The input file for convert follows a 'standard' format that can be easily obtained via an excel file containing the genotypic data (Glaubitz, 2004).

3.13.3. GenePop (Version 4.0) software (Rousset, 1997)

The primary purpose of the original DOS based software is to compute exact tests for the following:

- Hardy-Weinberg equilibrium
- Population differentiation, and
- Genotypic disequilibrium among pairs of loci

In addition, the software estimates classical population parameters, including:

- Allele Frequencies and other basic information,

- F_{st} and other correlations,
- Maximum likelihood estimates in the presence of a null allele,
- N_m estimates by the private allele method.

3.13.4. PopGene (Version 1.31) software

PopGene computes both comprehensive genetic statistics (*e.g.*, allele frequency, gene diversity, genetic distance, G -statistics, F -statistics) and complex genetic statistics (*e.g.*, gene flow, neutrality tests, linkage disequilibria, multi-locus structure).

3.13.5. Phylip (Version 3.67) software

PHYLIP (the *PHY*Logeny Inference Package) is a package of programs for inferring phylogenies (evolutionary trees). This program includes parsimony, distance matrix, and likelihood methods, including bootstrapping and consensus trees. Data types that can be handled include molecular sequences, gene frequencies, restriction sites and fragments, distance matrices, and discrete characters (Felsenstein, 2004).

3.13.6. Arlequin (Version 3.1) software

The goal of Arlequin is to provide the average user in population genetics with quite a large set of basic methods and statistical tests, in order to extract information on genetic and demographic features of a collection of population samples.

Analysis of Molecular Variance (AMOVA) is a method of estimating population differentiation directly from molecular data and testing hypotheses about such differentiation. A variety of molecular data – molecular marker data (for example, RFLP or AFLP), direct sequence data, or phylogenetic trees based on such molecular data – may be analyzed using this method (Excoffier, *et al.*, 1992).

3.13.7 Structure (Version 2.2) software

The program structure implements a model-based clustering method for inferring population structure using genotype data consisting of unlinked markers. The method was introduced in a paper by Pritchard, Stephens and Donnelly (2000) and extended in sequels by Falush, Stephens and Pritchard (2003, 2007). Applications of this method include demonstrating the presence of population structure, identifying distinct genetic populations, assigning individuals to populations, and identifying migrants and admixed individuals.