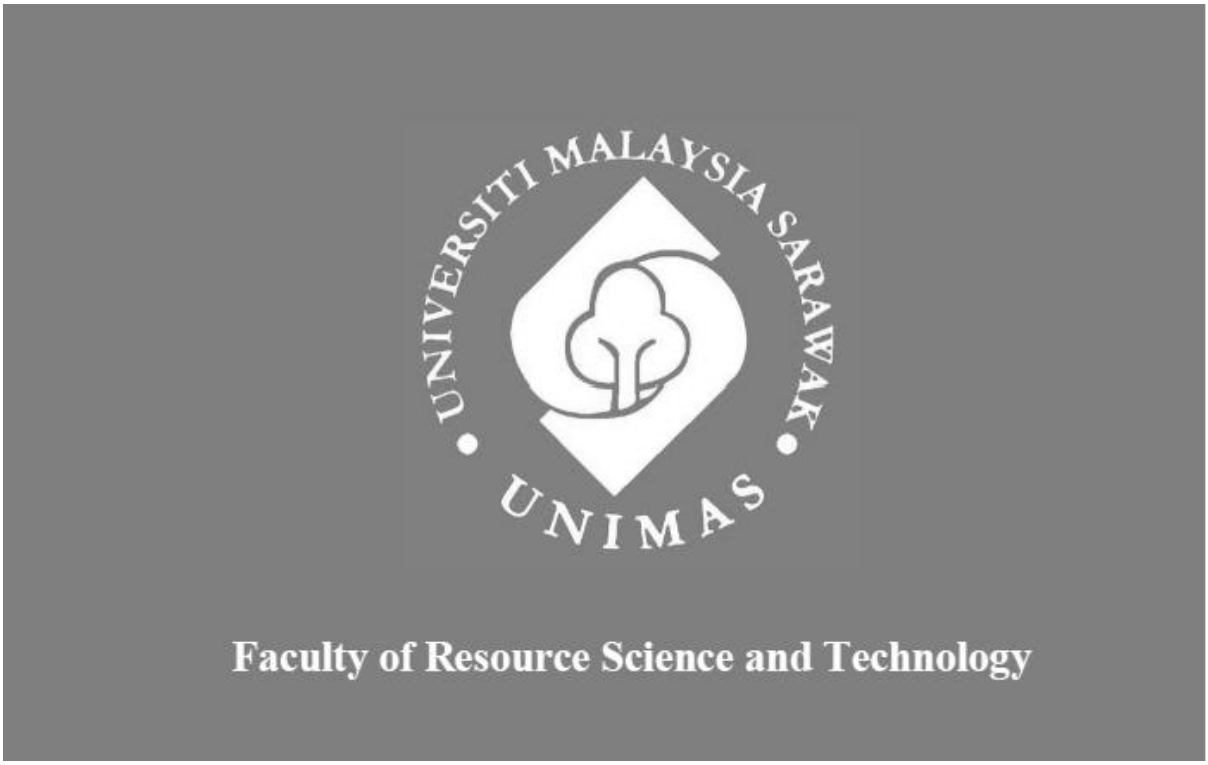

Molecular Phylogeny of *Rasbora* sp (Cypriniformes) Inferred from Cytochrome Oxidase Subunit I (COI)

Ummi Noorhanifah Bt Abdullah 28567

Final Year Project 2 (STF 3012)

This dissertation is submitted in partial fulfillment of the requirements for

The Degree of Bachelor of Science with Honour in

(Animal Resource Sccience and Management)

2013

Molecular Phylogeny of *Rasbora* sp (Cypriniformes) Inferred from Cytochrome Oxidase Subunit I (COI)

UMMI NOORHANIFAH BT ABDULLAH

This project is submitted in partial fulfillment of

The requirements for the Degree of Bachelor of Science with Honours

(Animal Resource Science and Management)

Faculty of Resources Science and Technology

UNIVERSITI MALAYSIA SARAWAK

2013

APPROVAL SHEET

BSc. Candidates: Ummi Noorhanifah Bt Abdullah

Title of Dissertation: Molecular Phylogeny of *Rasbora* sp (Cypriniformes) Inferred from Cytochrome Oxidase Subunit I (COI)

Dr Ramlah Bt Zainuddin

Supervisor

Department Of Zoology

Faculty of Resource Science and Technology

Universiti Malaysia Sarawak (UNIMAS)

DECLARATION

I hereby declare that this finale year project report 2013 is based on my original work except for the quotations and citations which have been dully acknowledged also, declare that it has not been or concurrently submitted for any other degree at UNIMAS or other institutions of higher learning.

UMMI NOORHANIFAH BT ABDULLAH

Department Of Zoology

Faculty of Resource Science and Technology

Universiti Malaysia Sarawak (UNIMAS)

ACKNOWLEDGEMENT

First and foremost I would like to extend my highest gratitude to God, with His blessing I finished my final year project entitled 'Molecular Phylogeny of *Rasbora* sp (Cypriniformes) Inferred from Cytochrome Oxidase Subunit 1 (C01)'. Deepest thanks to my parents, Abdullah B Mukhtar and Saloma Bt Hamzah and my family members for their inspiring thoughts and enthusiasm encouraging me from the beginning till the end.

Special thanks to my former Supervisor, Dr Yuzine B Esa for allowing me to choose this project which also fulfill my needs. I would like to give my deepest thanks to my current supervisor, Dr Ramlah Bt Zainuddin on all the guidance and taught she provided to me along the critical period. Thank you to Dr Khairul Adha from Aquatic Department, who is my co-Supervisor for the guidance in species identification.

To all Zoology Department staff especially Mr Nasron Ahmad, Mr Huzal Irwin, Mr Isa and Mr Trevor lots of thanks I would like to give on their sincere help and guidance. Same goes to all the post graduate students that have helped me in order to make this project successful.

Thank you to all my molecular laboratory friends for the accompanying me in the lab. Last but not least, to all Zoology department members those have contributed directly and indirectly. We have completed this project together.

Table of Contents

List of Abbreviation

%	Percentage
	Degree Celcius
µl	Microlitres
bp	Base Pair
CTAB	Cetyl Trimethyl Ammonium Bromide
ddH_20	Dionised Distilled Water
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphate
EDTA	Ethyl Diamine Tetra Acetate
H_20	Water
MEGA	Molecular Evolutionary Genetics Analysis
$MgCl_2$	Magnesium Chloride
Min	Minutes
Ml	Mililitres
mtDNA	Mitochondrial DNA
NaCl	Sodium Chloride
PCR	Polymerase Chain Reaction
Primer COI-a	Forward Primer
Primer COI-e	Reverse Primer
Rpm	Round Per Minute
Sp	Species
TAE buffer	Tris-Acetate-EDTA Buffer
Taq	*Thermus aquaticus*
UV	Ultra Violet
W	Watt

List of Table and Figures

List of Tables

List of Figures

Molecular Phylogeny of *Rasbora* sp (Cypriniformes) Inferred from Cytochrome Oxidase Subunit I (COI)

Ummi Noorhanifah Bt Abdullah 28567

Department of Zoology
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

ABSTRACT

Rasbora sp is member of tribe Rasborini from family Cyprinidae, well known to become significant fish in ornamental trade and molecular tools. Apart from that Rasbora sp give significant role in aquatic diversity. *Rasbora sarawakensis*, *Rasbora einthovenii*, and *Rasbora myersi* are very well-known to become Southern Asia *Rasbora*. In order to understand the genetic variation of these complex genuses, the analysis was done by using mitochondrial Cytochrome Oxidase Subunit I (COI). All of the samples were been caught from their wild habitats. The mitochondrial COI gene segment sequence was amplified through Polymerase Chain Reaction. After alignment, 623 bp sequences were yield. The phylogenetic tree exhibit non-monophyletic relationship due to existence of *Rasbora daniconius.*

Keywords : *Rasbora sp*, CO1, phylogenetic tree

ABSTRAK

Rasbora *sp adalah sebahagian daripada kumpulan Rasborini daripada famili Cyprinidae dikenali sebagai ikan yang penting dalam perniagaan perhiasan dan sebagai alat molecular. Banyak ikan Rasbora yang memberi fungsi yang penting dalam diversiti hidupan air.* Rasbora sarawakensis, Rasbora einthovenii, *dan* Rasbora myersi *sangat dikenali sebagai* Rasbora Semenanjung Asia. *Untuk memahami perbezaan genetic dalam genus yang kompleks ini, analisis telah dilakukan menggunakan mitokondria sitokrom oxidas subunit I (COI). Semua sampel ditangkap di tempat asal mereka. Mitokondria CO1 genetik segment didapati melalui Reaksi Rantaian Polimeras. Selepas disusun, 623 bp susunan diperolehi. Analisis filogenetik menghasilkan hubungan tidak monofiletik disebabkan kehadiran Rasbora daniconius.*

Kata kunci : Rasbora sp, CO1, pokok filogenetik

CHAPTER 1.0

BACKGROUND STUDY

Family Cyprinidae is the largest group of freshwater fish in underwater world (He *et al.*, 2007). This family consists of 122 genera and 600 species. Each of the species differ in their morphologies, niches and feeding habits. Cyprinid fishes can be found worldwide except in South America, Australia and Antartica.

One of cyprinid fish is *Rasbora* which is from the order Cyprinoformes. Based on Tang *et al.* (2011), *Rasbora* is located under the subfamily Danioninae, tribe Rasborini. However, modern taxonomy renamed the subfamily Danioninae as Rasborinae. Common name of *Rasbora* in Malay is 'Ikan Seluang'. It is in the same group with carp, barb, gudgeons and golden fishes. It can be found throughout peninsular and East Malaysia. Significant role of *Rasbora* in systematic studies is often utilised as the molecular tools due to its small size. *Rasbora* is also famous as the ornamental fishes that be traded in aquarium.

The classification by Chen *et al.* (1984) assumed that the traditional recognition of subfamily grouping were monophyletic such as Rasborinae (also referred as Danioninae) and Leuciscinae (He *et al.*, 2008). However, recent study showed that *Rasbora* is not monophyletic (Mayden *et al.*, 2007) with *Pectenocypris* and *R. daniconius* clade as the sister group of *Rasbora*, this confirms that *Pectenocypris* is more closely related to *Rasbora* (Tang *et al.*, 2010). Scientists cannot relate the clades of Cyprininae, Leuciscinae

and Rasborinae due to no existence of evidence which can support that one clade is more closely related to one another. Locating *Rasbora* and *Trigonostigma* as the synonym of *Rasbora* and denying the existence *Rasbora daniconius* in the genus might allow to monophyletic *Rasbora* (Tang *et al.*, 2010). However, the experiment is far from complete so the conclusion cannot be done.

In Malaysia, there is lack of molecular studies to verify the classical work on fish taxonomy (Ryan & Yuzine, 2006). The study about *Rasbora* fishes is merely none compared to the huge numbers of fishes in the group; order Cypriniformes. No wide specific phylogenetic relationship has been published recently regarding *Rasbora*.

Objectives

1. To elucidate phylogenetic tree of genus *Rasbora*
2. Testify the monophyly of *Rasbora*

CHAPTER 2.0

LITERATURE REVIEW

2.1) Biology of *Rasbora*

This genus can be found throughout the Indian subcontinent, southern China, and Southeast Asia which includes Sumatra, Java and Borneo. The unique morphology of this family is toothless mouth but has one to three rows of teeth that can be seen at the pharyngeal bones (Mohsin & Ambak, 1983). It has no barbell with small oblique jaws. The other morphology features that can be seen is upwardly projecting symphyseal knob fitting into a concavity in upper lip when mouth is closed (Roberts, 1989). Lateral scale rows 21-44 with predorsal scales 9-17.

Rasbora is small in size and bony. Their spawning habits are varies, including making nests under the heavy objects. The male will guard the eggs and bury the eggs under the gravel (Mohsin & Ambak, 1983). Taxonomically, this genus may be polyphyletic.

One of the species is *Rasbora sarawakensis*. As the name indicate this genus is endemic to Borneo and also can be found at West Kalimantan, Indonesia. They are found in slow-moving forest streams with thick marginal vegetation. These are often shaded by the dense rainforest canopy. The morphology of the species are it has large stout bodied with large, pointed head, distinguished from all other species by intense lateral longitudinal

stripe, darkened dorsal fin and scale marking (Roberts, 1989). *R. sarawakensis* is also has many close-set of breeding tubercules. Its lateral line is complete within the range of 24-26.

Rasbora einthovenii is distributed around Malay Peninsula and Borneo. They differ from the other *Rasbora* with *R. einthovenii* appeared to have solid dark longitudinal stripe from snout-tip to end of middle caudal fin rays, with caudal peduncle lying below midline of body (Roberts, 1989).

Rasbora elegans is endemic to peninsula Malaysia, Singapore and the Greater Sunda Islands of Borneo and Sumatra. The morphology features that can be observed are there are two spots below dorsal and at caudal base with complete lateral line (Kotellat *et al*, 1993). *Rasbora einthovenii* are also distributed around Sumatra, Borneo and Peninsular Malaysia. Differ in morphology, *R. einthovenii* has dark lateral line from snout to end of median caudal rays (Kotellat *et al.*, 1993).

2.2) Mitochondrial DNA

Mitochondria are structures within cells that convert the energy from digested food. Although most DNA is packaged in chromosomes within the nucleus, mitochondria also have a small amount of their own DNA. This genetic material is known as mitochondrial DNA or mtDNA. Differ from nuclear DNA, mtDNA can be found in mitochondria which only inherited the female line of trait.

Previous systematic analyses of cyprinids have been focussly on morphology whereas the most recent study on mitochondrial DNA sequence (mtDNA) (Kong *et al.*, 2008). The limitation of this Cyprinids study upon maternally inherited mtDNA variation (He *et al.*, 2008). The maternal inheritance of genes may lead in dubious species due to hybridization or introgession that may not be reflected in gene tree. Furthermore, the mitochondrial DNA is circular, the independence gene in mitochondrion in representing separate characteristic must be considered (He *et al.*, 2008). When the taxon-specific primer mismatches problem were detected, choosing on the nucleotide composition of new primers will be aided by the very large number of complete mitochondrial genomes available for fishes (Inoue *et al*., 2001).

It was clear that mitochondrial genome of animals is a better target for analysis than the nuclear genome because of its lack of introns, its limited exposure to recombination and its haploid mode of inheritance (Saccone *et al.*, 1999).

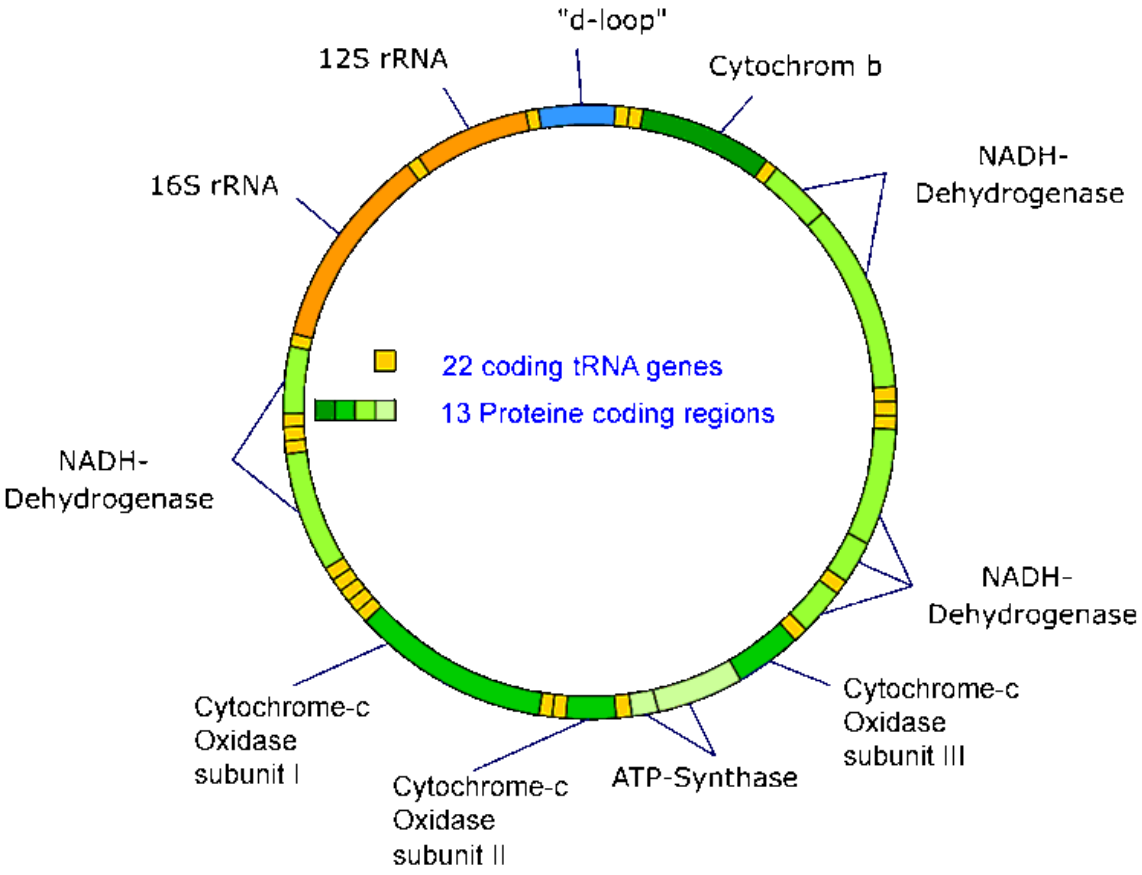

Figure 1: Mitochondrial DNA mapping (Source: Google, 2013)

2.3 COI Primer as barcoding gene

Hebert *et al.* (2003) has stated that mitochondrial gene cytochrome-*c* oxidase I (COI) can serve as the core of a global identification of biological system for animals. The gene can identify almost all of the animals. The advantage of this gene is short enough to enable the sequences to be interpreted quickly. The character of CO1gene is situated in mitochondrial DNA as the energy source of mitochondria.

COI gene ability in reading long sequence was significant regarding two biological condition (Hebert *et al.*, 2003). It was tend to biased into C-G nucleotide in vertebrate animals. COI was better DNA marker than 16S marker (Nicholas *et al.*, 2012). Hebert *et al.* (2003) experiment showed that COI sequences always considered the amino-acid divergence from sequences possessed by other taxonomic groups.

Ironically, the COI-2 primer cocktail that was originally designed for amplification of the mammals barcode region, was also performs very well for fishes (Ivanova *et al*., 2007). Ivanova *et al.* (2007) indicated that the average sequencing success of amplicons for fishes was high for COI-3 (95.2%) and COI-2 (93.0%), but lower for COI-1 (86.0%).

2.4) Phylogenetic Relationship

Some has explored the interrelationships within Cyprinidae in phylogenetic context using osteological characters (He *et al.*, 2008). On the other side, phylogenetic relationships of the Cyprinidae from East Asia was compared by using the mitochondrial cytochrome-*b* (He, 2004). He *et al.* (2008) examine the phylogenetic relationship of Cyprinidae from all hypothesized subfamilies, and test the monophyly of the family using maximum parsimony and Bayesian analysis.

Once a robust phylogenetic tree is established, patterns of change in evolutionary time may be determined by mapping well-defined characters on that tree (Webb & Schilling, 2006). It was clearly shown that *Rasbora* was not a monophyletic genus. *Rasbora daniconius* did not group with the other *Rasbora* sp, hence removing *R. daniconius* from genus *Rasbora* could maintain the monophyly of this genus (Tang *et al*., 2010). There were also relationship between genus *Trigonostigma* and *Boraras* within *Rasbora* which indicated by Tang *et al*. (2010) the *R. pauciperforata* became the sister clade of *Boraras* while *R. spilocerca* and *R. trilineata* became the sister clade of *Trigonostigma* group.

2.5) Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) is rapid amplification of cells techniques. PCR used two synthetic oligos as primers to amplify a nucleotide sequence of interest. These primers are annealing the either end of the targeted nucleotide sequence and are oriented in opposite directions (Ho *et al*., 1989).

Exponential amplification of the target sequence occurs over the course of multiple rounds of denaturation, annealing and 3' extension by DNA polymerase. Denaturation works by using high temperature condition. This broke the double helix into two single strands nucleotide without breaking up the major components. Hence, this allowed the primers to take part. When the double helix has succeeded to be broken off, the temperature decreased which enable the primers to anneal the complementary regions thus building another double strands between primers and complementary sequences. Extension occur when the DNA primers were able to perform double strands, hence the enzymes would read the opposite sequences and lengthen the sequences by adding nucleotide where it can pair.

PCR can be used to introduce additional sequences, such as restriction sites, by incorporating these into the oligo primers (Mullis, 1987). However, this use of PCR as a means of site-directed mulagenesis is limited because all sequence alterations must be introduced within the primer located at the ends of the targeted sequence. Given that restriction sites must be located at the ends of fragment to permit cloning, this approach requires that the sites of mutagenesis be located near these restriction sites. The introduction of mutations at other sites within the amplified gene sequence is therefore not possible.

CHAPTER 3.0

RESEARCH METHODOLOGY

3.1) Samples Preparation

Rabora sp can be found all around Malaysia. All nine samples of the fish's samples have been caught in three different places; Kubah National Park (N 1° 35'76" E 110° 10'85"), Gunung Gading National Park (1°42'11"N 109°49'12"E) and Sg Keratong , Pahang(3°3'0" N 102°58'59" E) (figure 5 and figure 6) and being preserved in 70 % ethanol. The fishes selected were identified as *sarawakensis, myersi*, and *einthovenii*.

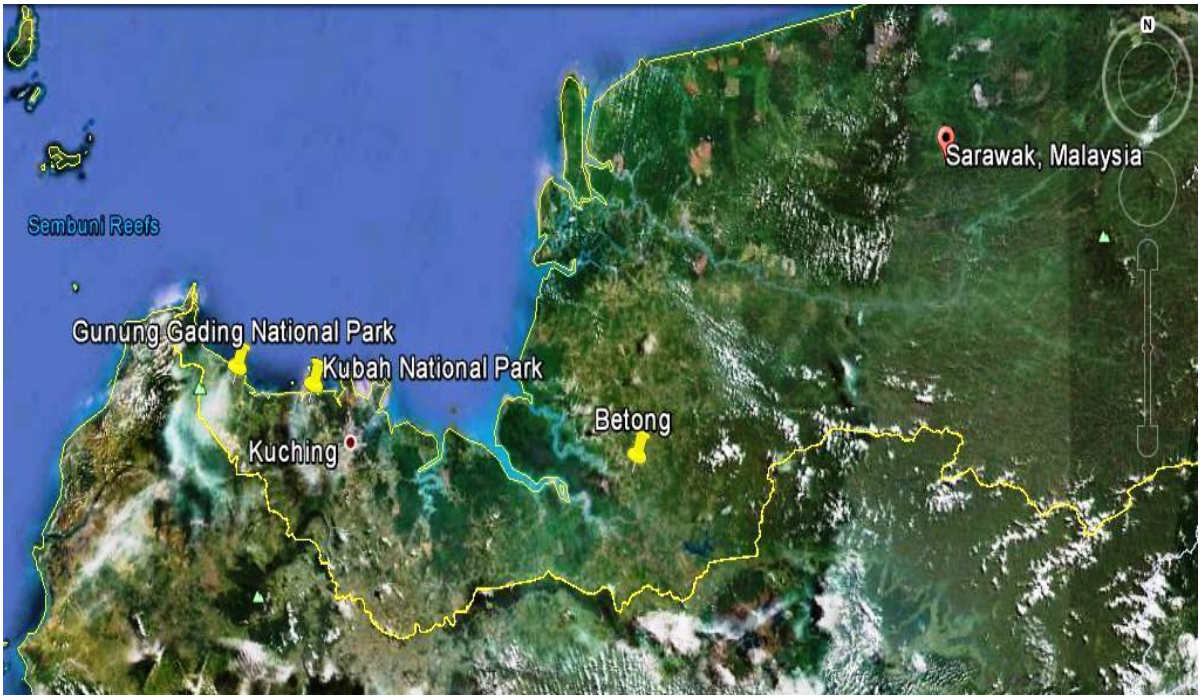

Figure 2 : Sampling location in Sarawak (source: Google Earth)

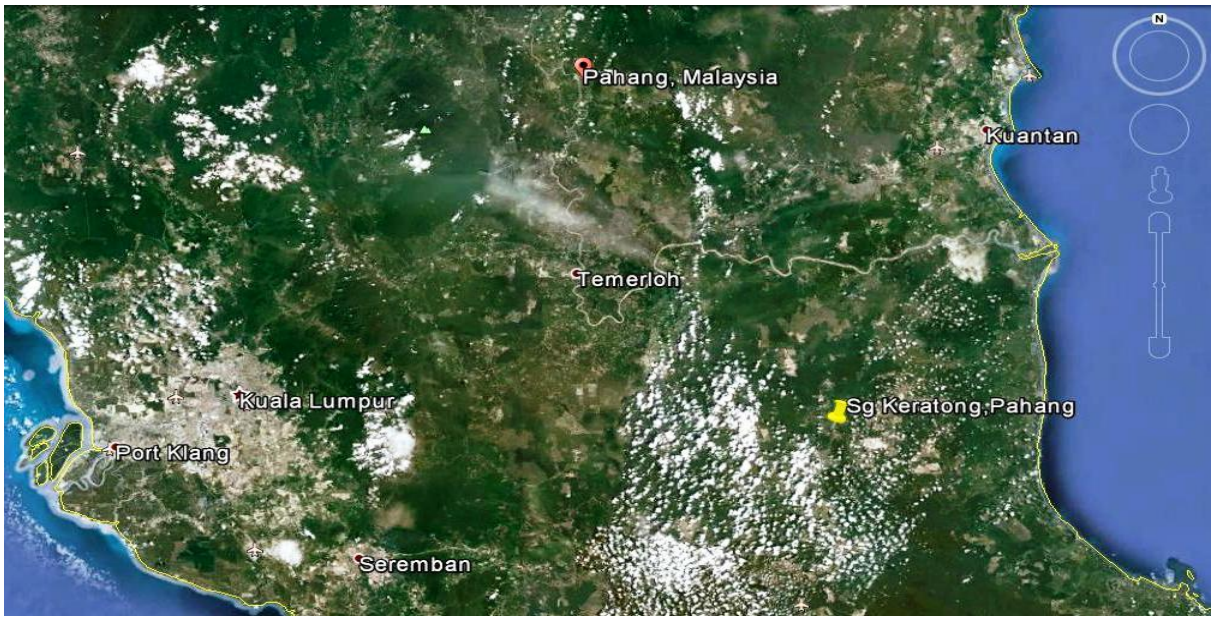

Figure 3: Sampling location in Pahang (Source: Google Earth)

3.2) DNA Extraction Samples

The extraction was done by following CTAB protocol. Approximately 0.5g of flesh from the samples (*Rasbora* sp) were taken to be grinded. The tissue was put into 1.5ml eppendorf tube with 700µl CTAB buffer and 10µl Proteinase k. The eppendorf tube was shaked well. The tissue was incubated until it was dissolved. The incubation was monitored every 10 minutes. After the tissue has dissolved, 600µl of chloroform-isoamyl alcohol in the ratio of 24:1 was added and be shake. Then, the samples were centrifuged at 13,000 rpm for 20 minutes.

Once the centrifugation has complete, two layers appeared. The samples were taken out and only the upper layer of the supernatant was pipette out into new labelled tube. About 600µl cold absolute ethanol was added, was mixed well and spanned at 13,000 rpm for 20 minutes. The ethanol is discarded by pouring it. About 600 µl cold 70% ethanol and 25 µl of NaCl were added then be spanned for the third time. The ethanol was then discarded and ensured the DNA pellet was still present. The pellet is left to dry at room temperature. The pellet was redisolved in 30µl of dionised distilled water (ddH_20). The tube be placed overnight then be kept in the freezer with -20℃.

3.3) Gel Electrophoresis

A lot of 0.3 g agarose powder was mixed with 30 ml of TAE buffer which was the electrophoresis buffer to form 1% agarose gel. The mixture is heated in a microwave oven until it is completely melted around 1 minute. After the mixture has completely melted, the colour became translucent and 1µl ethidium bromide was added to the gel with final concentration of 0.5 ug/ml. This was to facilitate the visualization of DNA after electrophoresis. Then, it is poured into a casting tray containing a sample comb and allowed to solidify at room temperature. The buffer was poured into the gel and covered the whole gel. This is important to allow the charge of electric current to go through the gel.

The loading buffer was mixed with the DNA sample in the ratio of 1:1. 2 µl of DNA sample was used. The solutions were mixed by pipetting. Once the solution has well-mixed, be pipetted by one-time pipetting. The samples that containing DNA mixed with loading buffer are then be pipetted into the sample wells vertically. The power leads were placed on the apparatus, and a current as much as 100 watt was applied. It flowed from negative to positive electrode. In order to visualize DNA, the gel was placed on an UV lightbox; UV transilluminator.

3.4) Polymerase Chain Reaction (PCR)

The master mix was mixed briefly and spins down by pulsing in a microcentrifuge to bring all the reactions to the bottom of the tube. 25µl mixture was pipete into each tube including 2.0 µl of template DNA to each reaction except for control tube which needed to be added 2.0 µl of ddH_2O. It was been spin down by pulsing in a microcentrifuge. The amplification will be carried out by using 25 µL reaction volume containing 2 µL DNA,1xPCR buffer, 0.2 mM dNTPs, 1.5mM MgCl2, 1.25 pmol of each primer and 0.05 U Taq polymerase (Ryan & Yuzine, 2006). Two primers from barcoding gene of COI have been chosen for this project (table 1).

Table 1: CO1 Primer Sequence (Ivanova *et al*., 2007)

Locus	Primer sequence (5' to 3')
VF1d_t1	F: TGTAAAACGACGGCCAGTTCTCAACCAACCACAARGAYATYGG
VR1d_t1	R: CAGGAAACAGCTATGACTAGACTTCTGGGTGGCCRAARAAYCA

Table 2: Master mix

Bil	COMPONENT	1 REACTION/µL
1	ddH_2O	13.3
2	1 x reaction buffer	5.0
3	dNTP mix (0.2mM)	0.5
4	Primer COI-2(10mM)	1.0
5	Primer COI-2(10mM)	1.0
6	MgCl	1.5
7	Template DNA	2.0
8	*Taq* polymerase	0.2
	Total Volume of Master mix	**25.0**

Table 3 : Amplification in PCR machine

Step	Temperature	Time
Pre-denaturation	$94^{O}C$	5 min
Denaturation	$94^{O}C$	45 seconds
Annealing	$47.5^{O}C$	45 seconds
Extension	$72^{O}C$	45 seconds
Final extension	$72^{O}C$	7 min
Cycle : 35		

3.5 DNA Purification and DNA Sequencing

The GeneJET™ PCR Purification Kit was designed for rapid and efficient purification of DNA from PCR and other enzymatic reaction mixtures The kit can be used for purification of DNA fragments from 25 bp to 20 kb. Each GeneJET purification column has a total binding capacity of up to 25 μg of DNA.

Binding Buffer was added in ratio of 1:1 to the PCR mixture and be mixed thoroughly. All of the mixture was transferred into GeneJET purification column. Centrifuge for 30-60 s. Discard the flow-through. 700 μL of Wash Buffer that has been diluted with the ethanol) was added to the GeneJET purification column. Centrifuge for 30-60 s. The empty GeneJET purification column was centrifuged for an additional 1 min to completely remove any residual wash buffer. The GeneJET purification column was transferred into a clean 1.5 mL microcentrifuge tube. Another 25 μL of Elution Buffer was added to the center of the GeneJET purification column membrane and centrifuged for 1 min. The GeneJET purification column was discarded and the purified DNA was stored at -20°C.

The mechanism of DNA sequencing is run through DNA facility. It operates on Applied Biosystems 3730xl DNA Analyzer. The DNA Analyzer is an automated system used for detecting fluorescently labelled dye terminators that have been added to the end of DNA extension products during sequence cycling. DNA is drawn electro kinetically into the capillaries then electrophoresed through a polymer matrix. The fluorescent is excited by a laser dyes as they pass through a detector. Each of the terminator (A, G, C, or T) has its own dye which fluoresces at a different wavelength allowing the detector to distinguish each of the dyes in the same capillary that exhibit as different colour peaks on the chromatogram.