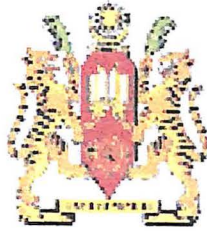


UNIVERSITI SAINS MALAYSIA



PUSAT PENGAJIAN SAINS KESIHATAN
(SCHOOL OF HEALTH SCIENCES, UNIVERSITI SAINS MALAYSIA)

**Sequence Polymorphism of mitochondrial DNA HVS II
region of Malay Population of Peninsular Malaysia**

**Dissertation submitted in partial fulfillment for the Degree of
Bachelor of Science in Forensic Science**

MOHD FAIZUL BIN ABD WAHAB

**School of Health Sciences
Universiti Sains Malaysia
Health Campus
16150, Kubang Kerian, Kelantan
Malaysia**

2008

ACKNOWLEDGEMENT

In the name of Allah, the most gracious, the most merciful. First of all, I would like to thank Allah because at last I had finish doing this thesis. Although there are many problems I face on the process of making this paper, I face it bravely in order to make it become reality.

I am indebted to my research supervisor, Dr. Zafarina Zainuddin, for her guidance and providing me with many information and guided me during research process. There are a number of people who contributed in my work, Mr.S.Paneerechelvan, Puan Norhaslindawaty, Puan Roslina, Puan Hafizah, which I acknowledge the help and assistance rendered.

My special thanks to my parent Mr. Abd Wahab Atan and Mrs Saripah Anuar Hj Majid, and all my friend especially Edinur Hisham, Mohd Yusran, Asrul Fahmi and Siti Fariza Noorani for their help, moral support, assistance and guidance for successful completion of the project.

TABLE OF CONTENTS

Contents	Page
Abstract	1
Introduction	2
Review of literature	10
Objectives of the Study	13
Materials and Methods	14
Results and Discussion	21
Conclusions	45
References	46

LIST OF TABLES AND FIGURES

List of tables;	Page
Table 1	Sequence of primer 16
Table 2	Calculation of dilution for each estimated concentration 19
Table 3	Sequences analysis of polymorphism in HVS II region 33
Table 4	Sequence polymorphism observed in the Malay samples 36
Table 5	HVS-II polymorphism and haplogroup assigned for each sampled by using the RFLP result, HVS-I and HVS-II polymorphism 38
Table 6	Frequency of each haplogroup in Malay population 43

List of figures;	Page	
Figure 1	Structure of mtDNA	5
Figure 2	Graphical mtDNA haplogroup skeleton	9
Figure 3	Structure of EtBr	14
Figure 4	Diagram of Polymerase Chain Reaction	22
Figure 5	Electrophoresis of PCR product in 1% agarose at 90V for 45 minutes	24
Figure 6	Electrophoresis of PCR product in 1% agarose at 90V for 45 minutes	24
Figure 7	Electrophoresis of PCR product in 1% agarose at 90V for 45 minutes	25
Figure 8	Electrophoresis of PCR product in 1% agarose at 90V for 45 minutes	25
Figure 9	Electrophoresis of purified PCR product in 1% agarose at 90V for 45 minutes	27
Figure 10	Electrophoresis of purified PCR product in 1% agarose at 90V for 45 minutes	27
Figure 11	Electrophoresis of purified PCR product in 1% agarose at 90V for 45 minutes	28
Figure 12	Electrophoresis of purified PCR product in 1% agarose at 90V for 45 minutes	28
Figure 13	Electropherogram showing good sequencing result with a clear background	30
Figure 14	Electropherogram showing bad sequencing result	30
Figure 15	Alignment of Anderson Reference sequences with the HVS II sequence for sample M41. The arrow shows the position where tranversion and purine transition had occurred.	32

LIST OF ABBREVIATIONS

bp	: base pairs
DNA	: Deoxyribonucleic Acid
dNTP	: Deoxynucleoside triphosphate
EDTA	: Ethylenediamine tetraacetic acid
EtBr	: Ethidium Bromide
HVS	: Hypervariable Sequence
kb	: kilo base pairs
MgCl	: Magnesium chloride
mtDNA	: mitochondrial DNA
nDNA	: Nuclear DNA
ng	: nanogram
PCR	: Polymerase Chain Reaction
RFLP	: Restriction Fragment Length Polymorphism
RNA	: Ribonucleic Acid
TBE	: Tris Boric EDTA
μL	: microlitre
UV	: Ultra violet

ABSTRACT

Sequence variation in hypervariable II region (HVS-II) in the control region of mitochondrial DNA (mtDNA) was studied in 45 unrelated Malay individuals by PCR amplification and direct sequencing. A 267 bp fragment of HVS-II region was amplified using a set of primer HVC1 and HVD1. Agarose gel electrophoresis was carried to examine the polymerase chain reaction product. The product was purified prior to sequencing using QIAquick PCR Purification kit (QIAGEN). A total of 45 samples were sent for sequencing, but only 40 samples were successfully sequenced.

Within the 40 sequences, 39 different haplotypes were observed, which 2 individuals shared the same haplotypes. From the result obtained, 42.5% of the samples were classified under haplogroup M, 12.5% were classified under haplogroup B, 10% were classified under haplogroup F and another 5% were classified under other haplogroup, D. A total 27.5% of the samples were not yet classified.

1.0 INTRODUCTION

1.1 Human mtDNA

Human mitochondrial DNA (mtDNA) is a small enclosed circular molecule, containing 16569 basepairs (bp), the sequence of which has been entirely determined (Buscemi *et al.*, 2000). Mitochondria are the power station for eukaryotes. It produces energy for the cell in the form of ATP. Mitochondria contain two membranes, separated by a space. Inside the space enclosed by the inner membranes is the matrix. This appears moderately dense and one may find strands of DNA, ribosomes, or small granules in the matrix. The mtDNA encodes for 22 tRNAs, 2 rRNAs, and 13 structural proteins for oxidative phosphorylation (Chen *et al.*, 1995).

In 1981, Anderson and his co-researcher had reported the first sequence of human mitochondrial DNA (mtDNA) (Anderson *et al.*, 1981). It is also known as the Anderson Reference Sequence and serves as a reference to which subsequent mtDNA sequence analyses are commonly compared. This sequence has been revised recently with few corrections have been made and named as the Cambridge Reference Sequence (Andrew *et al.*, 1999)

An individual inherits his or her mtDNA only from the mother. Although the father's sperm have mitochondria organelles, they are not usually maintained in the fertilized egg. As a result, all maternally related individuals in a family will share the same mtDNA profile. This pattern is known as maternal inheritance (Andrew *et al.*, 2008)

Mitochondrial DNA (mtDNA) also harbors a series of characteristics, other than matrilineal inheritance such as non-recombination, rapid evolutionary rate, and high population-specific polymorphisms, which make it very useful as a genetic marker for studies on the population genetics and molecular anthropology (Binbin *et al*, 2007). The rate of sequence evolution in mtDNA is 10-20 times higher than in nuclear genome and consequently any two mtDNA may differ by 10-66 nucleotides from each other (Butler and Levin, 1998). Moreover, mtDNA is present in high copy number per cell. There are about 1,016 mtDNA within a typical human cell. For this reason, its analysis is an important tool for forensics identification and may provide useful information when the analysis of nuclear DNA fails (Lima *et al*, 2004).

1.2 Organization of mtDNA

Human mtDNA is divided into two regions the coding region and control region. The coding region is responsible for the production of various biological molecules involve in the process of energy production in the cell. Meanwhile, the control region is responsible for regulation of the mtDNA molecules (Greenberg *et al*, 1983). Basically control region consist of hypervariable region I (HVS-I), ranges from position 16,024 position to 16,365 and hypervariable region II (HVS-II), positioned from 73 to 340 (Wilson *et al*, 1995). Recently, hypervariable region III (HVS-III) was identified (Lutz *et al*, 1997). It is situated between nucleotides position 438 to 574.

The HVS reveal about 3% variability between individuals. Within the HVS, the polymorphic sites are not distributed uniformly but cluster in so-called hot spot (Finnila *et al*, 2001). The sequence in the control region is mainly point mutation SNPs, which do not alter the length of the mtDNA and detected easily only by determining the base sequence. Uncommon single base insertion and deletions have also been found. There are no STRs in mtDNA (Andrew *et al*, 2008)

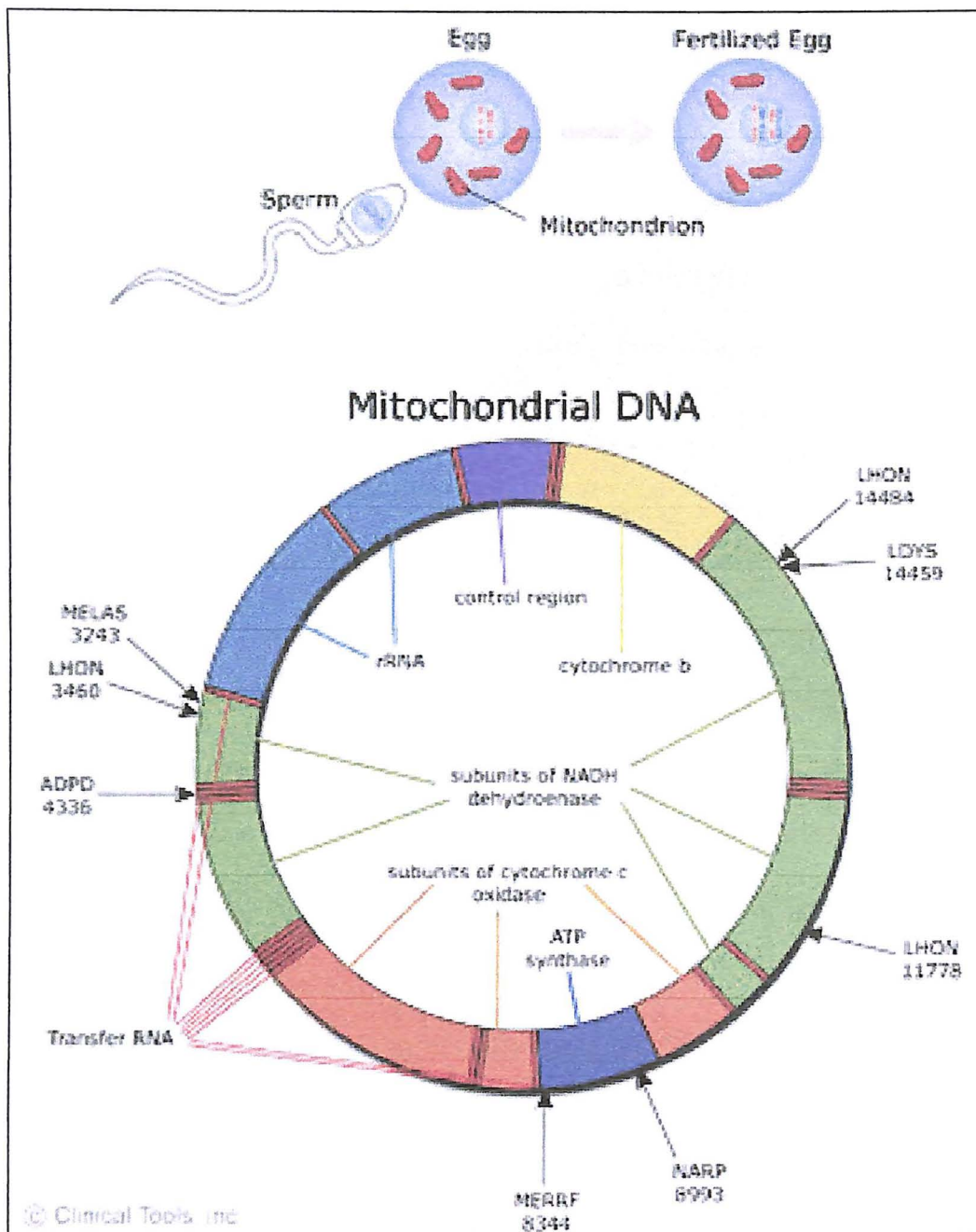


Figure 1: Structure of mtDNA

Source: <http://fig.cox.miami.edu/~cmallery/255/255etc/mito.dna.jpg>

1.3 Advantages of mtDNA in Forensics DNA Analysis

The study of mitochondrial DNA (mtDNA) sequence variation has been a valuable tool for forensic identification. There is a high level of sequence variation in the non-coding control region of mtDNA making these segments suitable for forensic identification. However, reports on heteroplasmic point mutations in these regions in the same individuals were not uncommon. This phenomenon gives rise to the problem of the application of mtDNA in personal identification (Oikawa, *et al*, 2002).

mtDNA advantages in Forensics DNA analysis are due to several features such as the sequences variability of its control region between individual, its efficient polymerase chain reaction (PCR) amplification with limited biological material (Sullivan *et al.*, 1991,1992) and its resistance to extreme environmental condition (Sullivan *et al.*, 1992; Boles *et al.*, 1995).

In addition to the skin, blood, semen and saliva, which are normally used for human identification, mtDNA has also been extracted from teeth, hair shaft, bone fragments and even human feces, all of which often fail to yield forensics result with nuclear DNA marker (Butler and Levin, 1998).

The high copy number of mtDNA in each cell given the forensics scientist a better chance in obtaining mtDNA profile from limited or severely degrade sample. Moreover, the mitochondrial DNA is buried deep within the cell and has a circular structure, which it protects from deterioration (Andrew *et al*, 2008).

1.4 Mitochondrial DNA Analysis in Population Studies

The analysis of Mitochondrial DNA has become an important tool for studying of human population structure and history (Stoneking *et al.*, 1993). The first human population studies that based on mtDNA were performed b RFLP analysis and it has revealed differences between the four great ethnic groups, Caucasian, Amerindian, African and Asian. It is also being used as a genetic marker for reconstructing human evolution (Perièiæ *et al.*, 2005) in understanding modern human evolution and migrations.

One of the objectives of population study is to quantify the amount of inheritable variant present in nature (Konskien, 2002). Primary forces that lead to variation are natural selection, random genetic drift, mutations and gene flow. The study of mtDNA restriction polymorphism has become a source to detect genetic variations between major human ethnic groups (Wilson *et al.*, 1993).

The other aim of population study is for human identification for forensics purpose. In Malaysia, data about Malay population is still limited and not complete. Therefore the study about Malay population is very important to obtain data as references before it can be use as a tool for human identification for forensics purpose (Zafarina *et al.*, 2004)

1.5 mtDNA Haplogroup

In the study of molecular evolution, a haplogroup is a group of similar haplotypes that share a common ancestor with a SNP mutation. Since a haplogroup consists of similar haplotypes, this is what makes it possible to predict a haplogroup from haplotypes. Haplogroups are assigned by alphabets, and its refinement consists of additional number and letter combinations. RFLP analysis of the mtDNA coding region is used to confirm haplogroup.

The first study directed to the identification of mtDNA haplogroups in Europe was carried out by high resolution restriction analysis in individuals of European ancestry living in North America and revealed four European-specific haplogroups (H, I, J and K) (Wallace, 1995). More recently, studies based on analysis of the entire mtDNA in conjunction with sequence data from both hypervariable segments (HVS-I and HVS-II) of the mtDNA control region (Torroni *et al.*, 1996) have not only confirmed the wide distribution of haplogroups H, I, J and K among European populations, but have also revealed the presence of some additional haplogroups. These additional haplogroups were named T, U, V, W and X. Similar to the four haplogroups identified initially, it was observed that haplogroups T, V and W are Caucasoid-specific, while haplogroups U and X are shared between Europeans and Africans (Torroni *et al.*, 1996), respectively. These nine haplogroups, together with a few representatives of the Asian super-haplogroup M, and the African haplogroups L1 and L2 (Chen *et al.*, 1995), were found to encompass virtually all mtDNAs in Europe (Torroni *et al.*, 1996)

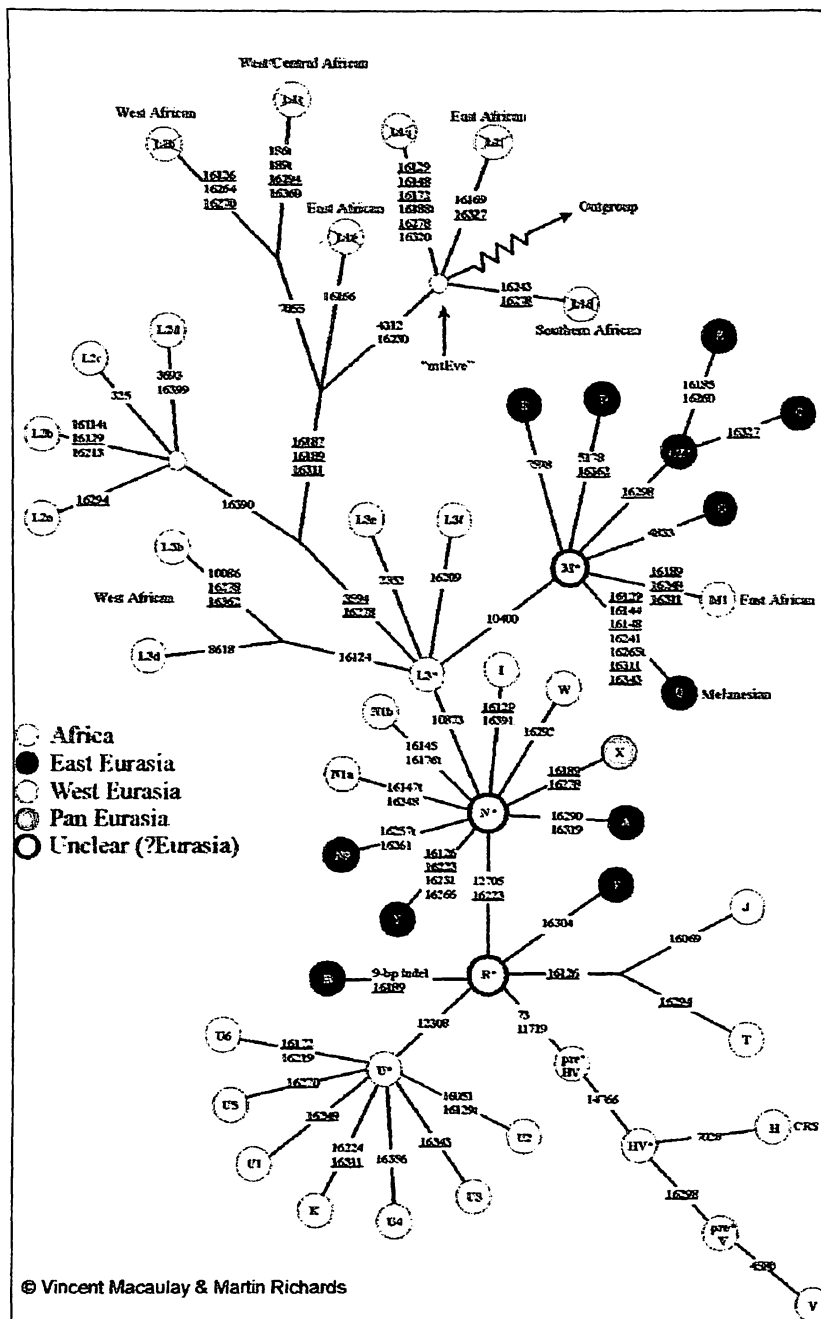


Figure 2: Graphical mtDNA haplogroup skeleton

Source: <http://www.stats.gla.ac.uk/~vincent/images/skeleton07-08-02.jpg>

2.0 REVIEW OF LITERATURE

Variation in human mitochondrial DNA (mtDNA) has been used to infer the origin and migration patterns in human populations. mtDNA analysis has been focused mainly on the hypervariable region I (HVS-I). Nevertheless, although many studies of the hypervariable region II (HVS-II) have been carried out during recent years, the correlation between the first and the second hypervariable regions has not been well established (Salas *et al*, 2000)

Salas and his co-researcher have analyzed 71 individuals from a relatively isolated region at the western edge of continental Europe and they have used available HVS-II sequence information from another 17 European and African populations. They have compared the results with other European and non-European sequences, and with the well-established variation patterns of HVS-I in Europe. The results show high concordance between the two hypervariable regions, not only in variability levels but also in other phylogenetic aspects. The study of the population structure through AMOVA analysis shows a low level of heterogeneity in the European populations.

Research done by Cerspillo and his co-researchers in 2004 had used polymerase chain reaction (PCR) and also direct sequencing to generate a population database from 200 unrelated Caucasian individuals living in Spain. By this technique, they determine the sequence polymorphisms of the mitochondrial DNA (mtDNA) control region, hypervariable regions I and II (HVS-I and HVS-II). A total of 175 different sequences were found as defined by 154 variable positions. The most common sequence

occurred 10 times, this sequence is also the most frequent in other European populations such as Austrian, German and British. The mean pairwise difference for the two regions taken together was 8.25. The study revealed that transitions made up the majority of the deviations (88%), whereas transversions were observed at significantly lower frequency (8%). A statistical estimate of the results for this Caucasian population showed a genetic diversity of 0.9965. The probability of two random individuals showing identical mtDNA haplotypes is 0.84.

Based on Schurr and Wallace (2002), most of haplogroup in Southeast Asia constitute haplogroup B, F and M. Distribution of haplogroup in Southeast Asia is first entered by (macro) haplogroup M from east Africa. Haplogroup F and B also present through out Southeast Asia, but are not genetically diverse as haplogroup M. The Vietnamese and Malaysian aborigines have high frequencies of haplogroup F which also seen in most other Southeast Asian population and haplogroup B was present through out the region.

Haplogroup M which originated from east Africa was later dispersed into East Asia by Indian sub-continent, with a diverse array of haplotype evolving in South Asia since this time (Kivisild *et al.* 1999). As a consequence of this dispersal pattern, haplogroup M has been part of the initial expansions of the modern human groups into Southeast Asia, with most of the mtDNAs present in extent Asian populations being divided into general clusters based on the presence or absence of the +*Dde* II and +*Alu* I sites (Ballinger *et al.* 1992) and the 16223 mutation in the HVS-I (Macaulay *et al.*, 1999)

Haplogroup F is the other mtDNA lineage that is fairly widespread throughout Southeast Asia. Haplogroup F lack the 16223 C to T mutation seen in haplogroup M and possess the 16304 T to C transition in HVS-I region. Haplogroup F appears in a number of Asian population, including Filipinos (Cann *et al.* 1987) and Aboriginal Taiwanese (Melton *et al.* 1995, 1998) and occurs at its highest frequencies in Southeast Asia, especially in the Vietnamese (Ballinger *et al.* 1992). Relatively high frequencies of this haplogroup also occur in the Orang Asli of Malaysia, to whom the Vietnamese show linguistics ties (Bellwood, 1979).

Haplogroup B is commonly found in Vietnamese, Malaysian and Bornean population (Ballinger *et al.*, 1992). It is define by nucleotide transition at 16189, 16217 and 16519 in HVS-I region (Passarino *et al.*, 1993; Schurr *et al.*, 2000). Haplogroup B has a much wider distribution in Asia and the Pacific (Schurr *et al.*, 2000). Moreover it also has been observed in Aboriginal Taiwanese and Taiwanese Han population (Ballinger *et al.*, 1992; Melton *et al.*, 1998, 1995)

3.0 OBJECTIVES OF THE STUDY

The main objective of this project is to analyze the sequence of hypervariable region II (HVS II) of Malay population in Peninsular Malaysia. This information help in determination of haplogroup for the Malay population, in corroboration with HVS-I sequence data and RFLP analysis of the coding region.

mtDNA data obtained in this study can be included into the current database for application in forensics DNA analysis as well in tracing the Malay lineage.

4.0 MATERIALS AND METHODS

4.1 Materials

4.1.1 Chemicals

TBE Buffer

An amount of 53.9g of Tris base and 3.27g of EDTA is dissolved in 400 ml distilled water. Then, 23g of Boric Acid is added to the solution. The pH of the solution was adjusted to pH 8 with remaining 4.5 gram boric Acid and the final volume was make up to 500 ml and autoclaved.

Ethidium Bromide (EtBr)

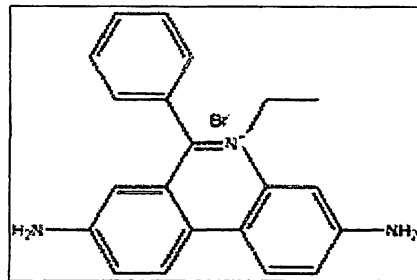


Figure 3: Structure of EtBr

Source: www.wikipedia.com/ethidiumbromide/

Ethidium bromide is an intercalating agent commonly used as a nucleic acid stain in molecular biology laboratories for techniques such as agarose gel electrophoresis. When exposed to ultraviolet light, it will fluoresce with a red-orange color, intensifying almost 20-fold after binding to DNA. Ethidium bromide may be a very

strong mutagen, and may possibly be a carcinogen or teratogen, although this has never been definitively proven.(www.wikipedia.com/ethidiumbromide/)

DNA Ladder

A DNA ladder is a solution of DNA molecules of different lengths used in agarose gel electrophoresis. It is applied to an agarose gel as a reference to estimate the size of unknown DNA molecules. In addition it can be used to approximate the mass of a band by comparison to a special mass ladder.

1 kb DNA ladder was used to determine the size of the PCR product and in agarose gel electrophoresis. It was purchased from Promega, USA

4.1.2 ReddyMix™ PCR Master Mix

ReddyMix™ PCR Master Mix was purchased from ABgene, UK. It contained all components that are required in PCR which is; *Taq* DNA polymerase, deoxynucleotide triphosphate (dNTPs), reaction buffer, Magnesium chloride, red dye and precipitant.

4.1.3 Thermal Cycler

The Thermal cycler (also known as a thermocycler, PCR machine or DNA amplifier) is a laboratory apparatus used for PCR. The device has a thermal block with holes where tubes with the PCR reaction mixtures can be inserted. The cycler then raises and lowers the temperature of the block in discrete, pre-programmed steps. Modern

thermal cyclers are often equipped with a hot bonnet, a heated plate that presses against the lids of the reaction tubes. This prevents condensation of water from the reaction mixtures to the insides of the lids and makes it unnecessary to use PCR oil. Some thermal cyclers are equipped with multiple blocks allowing several different PCR reactions to be carried out simultaneously. Also some apparatus have a gradient function, which allows different temperatures in different parts of the block. This is particularly useful when testing suitable annealing temperatures for primers.

4.1.4 PCR Purification Kit

QIAquick PCR Purification Kit (QIAGEN) was used for purification of PCR product. It has two methods either using collection tube or second is vacuum. The collection tube method was used for this study. This kit consist of QIAquick spin column, PBI buffer, PE buffer, EB buffer, collection tubes (2mL) and loading dye.

4.1.5 PCR Primer

Both forward and reverse primers used were synthesized by invitrogen. The sequence for each primer is listed below:

Primers	Sequences	Estimated melting temperature
HVC1 (L048)	5'-CTC ACG GGA GCT CTC CAT GC-3'	$T_m = 2(7) + 4(13)$ $= 66^{\circ}\text{C}$
HVD1 (H408)	5-CTG TTA AAA GTG CAT ACC GCC A-3'	$T_m = 2(12) + 4(10)$ $= 64^{\circ}\text{C}$

Table 1: Sequence of primer

4.2 Method

4.2.1 Sample collection

Sample used was collected by previous student for their final project. A total of 45 samples were collected from Malay population of Peninsular Malaysia. To avoid sample from individual with mixed marriage background, each subject shortly interviewed. Buccal swab sample was collected by rubbing their inner cheek using sterile cotton swab.

4.2.2 Sterilization

All glassware, pipette tips and microcentrifuge tubes were sterilized by autoclaving for 5 minutes at 20 psi. Buffer and dionised water was also autoclaved.

4.2.3 Separate working areas

Pre and post PCR work must be carried out in separate area. This is very important in order to avoid cross contamination. Each working area was cleaned prior and after work with 70% ethanol and each working area is allocated with different set of pipettes and other consumables.

4.2.4 DNA extraction

There are 3 methods commonly used for DNA extraction, organic extraction, chelex extraction and using FTA card. DNA template used in this study was extracted by previous student using phenol/chloroform technique (organic method).

4.2.5 Polymerase Chain Reaction (PCR)

Each PCR reaction consist of 22 μl of ReddyMixTM PCR Master Mix, 1 μl of forward primer, 1 μl of reverse primer, 0.5 μl of ddH₂O and 0.5 μl of DNA template.

The PCR process was started by initial incubation at 95°C for 3 mins followed by 30 cycles of denaturation at 95°C for 30 sec, annealing at 60°C for 30 sec and extension at 72°C for 30 sec. A final extension, at 72°C was carried for 5 mins. The PCR products were kept at 4°C in a separate fridge from the pre-PCR components to avoid contamination.

4.2.6 Agarose Gel Electrophoresis

A total of 1g of agarose powder was added into 100 ml of 0.5% TBE Buffer to make a 1% agarose gel. The mixture was heated in microwave oven until all agarose powder completely dissolved.

After the gel had completely solidified the gel was then transferred into an electrophoresis tank and submerged by adding approximately 300 ml of 10X TBE buffer. A total of 5 μl of the PCR product and 2 μl of DNA ladder were loaded into the gel and electrophoresed at 90 V for 45 mins.

4.2.7 Purification of PCR Product

A total of 5 volume of PBI buffer was added to 1 volume of PCR product and was mixed gently. The QIAquick column was placed in a provided 2 ml collection tube.

The sample was then applied to the QIA quick column and was centrifuged at 13,000 rpm for 60 sec. The flow through was discarded and the QIAquick column was placed back into the same tube. A total of 0.75ml PE buffer was then added and centrifuged again at 13,000 rpm for 60 sec. The flow through collected was again discarded and the QIAquick column was placed back to the same tube. The sample was centrifuged again in a 2ml collection tube at 13,000 rpm for 60 sec. Then each QIAquick column was placed in a clean 1.5 microcentrifuge tube. A total of 50 μ l of EB buffer was added and was centrifuged at 13,000 rpm for 60 sec to elute PCR product.

4.2.8 Agarose Gel Electrophoresis

Agarose gel electrophoresis was carried out for quantification purpose only. An amount of 2 μ l of purified PCR product was mixed with 1 μ l of loading dye before loading into each well. Electrophoresis was done at 90V for 45 mins. The gel was soaked with EtBr for 20 mins and the bands were observed under UV light.

4.2.9 Dilution

After the concentration was estimated, the sample was diluted to 30 ng in 50 μ l. The calculation is as below:

Estimated concentration	Calculation	Amount of ddH ₂ O added
50 ng	$M_1V_1 = M_2V_2$ $50(V_1) = 30(50)$ $V_1 = 30 \mu\text{l}$	50 μ l – 30 μ L = <u>20 μL</u>
60 ng	$M_1V_1 = M_2V_2$ $60(V_1) = 30(50)$	50 μ l – 25 μ L = <u>25 μL</u>

	$V_1 = 25 \mu\text{l}$	
75ng	$M_1V_1 = M_2V_2$ $75(V_1) = 30(50)$ $V_1 = 20 \mu\text{l}$	$50 \mu\text{l} - 20 \mu\text{L}$ $= \underline{30 \mu\text{L}}$
80ng	$M_1V_1 = M_2V_2$ $80(V_1) = 30(50)$ $V_1 = 18.8 \mu\text{l}$	$50 \mu\text{l} - 18.8 \mu\text{L}$ $= \underline{31.2 \mu\text{L}}$

Table 2: Calculation of dilution for each estimated concentration

Example

Sample 1 (estimated concentration = 50 ng)

Using $M_1V_1 = M_2V_2$

Estimated concentration (M_1) = 50 ng

Volume of sample 1 (V_2) = unknown

Concentration desired (M_2) = 30 ng

Volume desired (V_2) = 50 μl

Therefore,

$$50 \text{ ng } (V_2) = 30 \text{ ng } (50\mu\text{l})$$

$$V_2 = 30 \mu\text{l from sample 1}$$

Since the final volume required is 50 μl a total of 20 μl of ddH₂O was add.

4.2.10 Sequencing

All purified PCR product was sent to Tech Dragon Limited for sequencing service along with the forward and reverse primer used in amplification (HVC1 and HVD1).

5.0 RESULT AND DISCUSSION

5.1 Sample collection

Sample was collected by previous final year student. A total of 45 samples were collected from unrelated individuals. A short interviews and briefing about this project was carried out to ensure that subject has no mix-marriage background for at least 3 generations.

5.2 Polymerase Chain Reaction (PCR)

The polymerase chain reaction is an extremely powerful method that allows selective amplification *in vitro* of specific target DNA sequences from large, heterogeneous source such as genomic DNA. In order to be able to select a particular target from a mixture of sequences, specific sequence information must be available. This information is used to synthesize short oligonucleotides also known as primer that initially binds to the complimentary sequences in a molecular hybridization reaction.

Generally, at the start of PCR process, DNA templates must be in single-stranded form, which is achieved by heat-denaturation. The reaction mixture also contains the DNA building block, the four deoxynucleoside triphosphates dATP, dCTP, dTTP, and dGTP. DNA polymerase is then added, which synthesizes a complementary strand by connecting the deoxynucleoside triphosphates according to the sequence information given by the template strand. The specificity of the PCR reaction derives from the fact that DNA polymerase is unable to start a completely new strand; it can only extend the primer sequence. Subsequently, the newly formed double strand DNA molecules are

heat-denatured again, and can themselves serves as template in a new cycle of DNA synthesis.

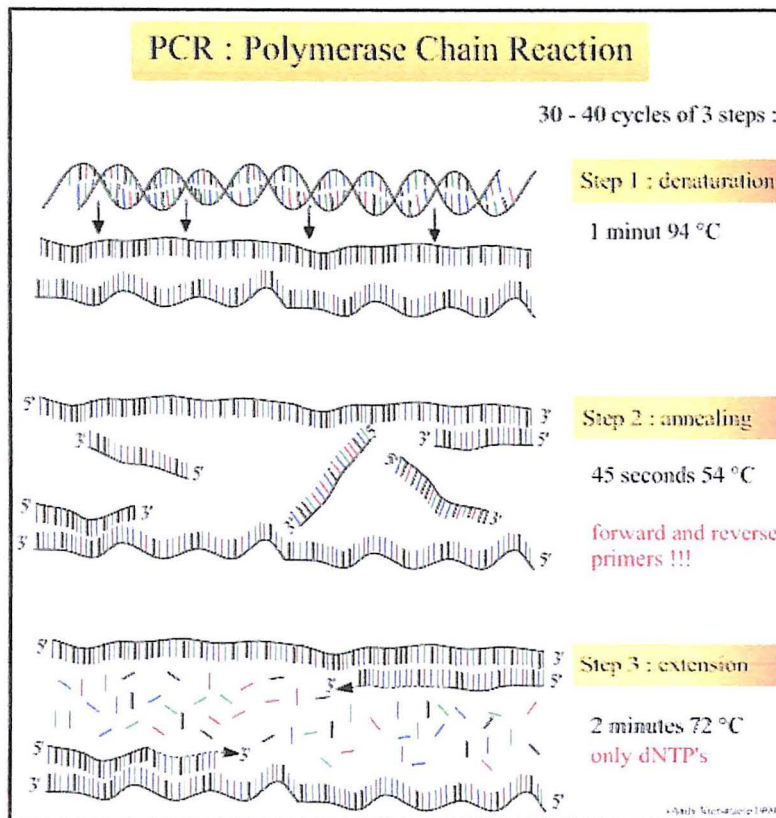


Figure 4: Diagram of Polymerase Chain Reaction

Source: <http://users.ugent.be/~avierstr/principles/pcr.html>

There are three basic steps in PCR reaction; denaturing, annealing and elongation/extension (figure 4). The first temperature called denaturation stage, of 94°C that breaks the hydrogen bonds of the double helix, so the DNA will become single stranded, exposing the bases in order that they can bind. Reducing the temperature to 60°C allows hydrogen bonds to reform, permitting the primers to base

pair to the complimentary sequences at the ends of the target region. At 72°C the heat-stable *Taq* polymerase enzyme extends the primers, synthesizing new DNA strands.

The product of this amplification process can be seen after agarose gel electrophoresis. The samples were loaded into agarose gel, electrophoresed at 90V for 45 mins and visualized under UV light.

5.3 PCR result

