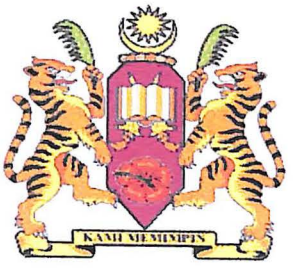
UNIVERSITI SAINS MALAYSIA



UNIVERSITI SAINS MALAYSIA

PUSAT PENGAJIAN SAINS KESIHATAN

(SCHOOL OF HEALTH SCIENCES, UNIVERSITI SAINS MALAYSIA)

Sequence Polymorphism of Mitochondrial DNA HVS-II Region of Chinese Population of Peninsular Malaysia

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

ASRUL FAHMI BIN ABDUL MUTALIB

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

2008

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious and Most Merciful.

First of all, I would like to express my deepest gratitude to Allah SWT, because of His grace and blessings, I am able to finish this research project.

My sincere thanks also go to my supervisor, Dr. Zafarina Zainuddin for her kind, valuable help, advices, opinions and support which guided me a lot in completing this research project.

Besides, I would like to thank to all forensic lab staff especially to Mrs. Roshaslindawaty, Mrs. Rosniah Yusoff and Mrs. Hafizah Harun for their help for me to carry out the lab work.

I also would like to thank to all my friends, especially Mohd. Faizul Abd. Wahab and Mohd Redhuan Mohd Noor for their support that accompanying me from the start till the end of the project.

Last but not least, I would like to thank my beloved family for giving me morale support and also to anyone who has given their advice and encourage me. Thank you

LIST OF CONTENTS

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

DNA	Deoxyribonucleic acid
mtDNA	Mitochondrial DNA
HVS-1	Hypervariable segment I
HVS-II	Hypervariable segment II
HVS-III	Hypervariable segment III
RFLP	Restriction fragment length polymorphisms
CR	Control region
CRS	Cambridge Reference Sequence
tRNA	Transfer ribonucleic acid
rRNA	Ribosomal ribonucleic acid
bp	Base pair
kb	Kilobase pair
np	Nucleotide position

ABSTRACT

The mitochondrial DNA hypervariable segment II (HVS-II) sequence variations from various populations are less published compared to hypervariable segment I (HVS-I). The primary objective of this study is to determine the HVS-II sequence polymorphisms of the Chinese population and then use the data acquired to classify the mtDNA haplogroup. The study was carried out by using 37 samples from Chinese population of Peninsular Malaysia. The HVS-II of the mtDNA control region was amplified and sequenced. Polymorphisms were reported by aligning each sequence to the Cambridge Reference Sequence (CRS). A total of 57 polymorphisms were detected, which encompass 27 transitions and 20 transversion. By using the HVS-II sequence data, 22 samples were classified into superhaplogroup M (M7b1, M8a, G2a, D5a, D4 and C) and 3 samples were classified into superhaplogroup N (R9a and N9a). The unclassified samples need further analysis for their classification.

1. INTRODUCTION

1.1 Human Mitochondrial DNA

1.1.1 Organisation of mitochondrial DNA

Mitochondria are subcellular organelles that contain an extrachromosomal genome separate and distinct from the nuclear genome. A mitochondrion contains between 2 to10 copies of mitochondrial DNA (mtDNA), and there can be as many as 1000 mitochondria per somatic cell. The human mtDNA is a histone-free, double-stranded circular DNA molecule (Anderson *et al.*, 1981). The complete nucleotide sequence of the human mtDNA was determined in 1981 (Anderson *et al.*, 1981) and reanalysis of the sequence was carried out in 1999 by Andrews and his co-researchers, where few corrections have been made and is known as Cambridge Reference Sequence (Andrews *et al.*, 1999). Human mtDNA is 16,569 base pair (bp) in length and present in 1,000 to 10,000 copies per cell (Robin, 1988). The total number of bases however varies slightly depending on the presence of insertions or deletions and the lengths of certain highly variable tandem repeats. Nucleotide positions in the mtDNA genome are numbered according to the Anderson References Sequence (Anderson *et al.*, 1981)

The mtDNA molecule has two strands, a guanine-rich strand or heavy strand (H) and cytosine-rich strand which also known as light strand (L), and it codes for 13 proteins, 22 tRNAs and 2 rRNAs. Most genes are transcribed from the H-strand, but one subunit of Complex I (MTND6) and eight tRNAs are transcribed from the L-strand. The

mitochondrial genome can be divided into two sections: a large coding region, which is responsible for the production of various biological molecules involved in the process of energy production in the cell, and a smaller 1.1 kb non-coding fragment, called the control region.

The non-coding region which also known as the displacement (D)-loop is situated between the mitochondrial $tRNA^{pro}$ and $tRNA^{phe}$ genes and contains three hypervariable segment, HVS I, HVS II and HVS III. HVS I situated between position 16,024 to 16,365, whereas HVS II extends from position 73 to 340 and HVS III is situated between position 438 to 574 (Greenberg *et al.*, 1983, Wilson *et al.*, 1993). The transcription of both strands is directed by the promoters in the D-loop, and almost the entire length of both strands is used to produce polycistronic RNA transcripts, which are subsequently cleaved to produce tRNAs, rRNAs and mRNAs. The transcription and transition of mtDNA are controlled by nuclear genes (Falkenberg *et al.*, 2002; Shoubridge, 2002; Casas *et al.*, 2003; Rodeheffer, 2003).

Due to the lower efficiency of DNA repair as well as higher frequency of DNA replication errors (Wilson *et al.*, 1993), mitochondrial DNA, particularly the non-coding region is highly polymorphic. The generation of mtDNA variability can only occur through new mutations. In vertebrates, the mutation rate of mitochondrial genes is approximately 10 folds higher compared to nuclear genes (Cann and Wilson, 1983). Most of the sequence variations between individuals were found within two specific segments of the control region (Greenberg *et al.*, 1983), the HVS I and HVS II.

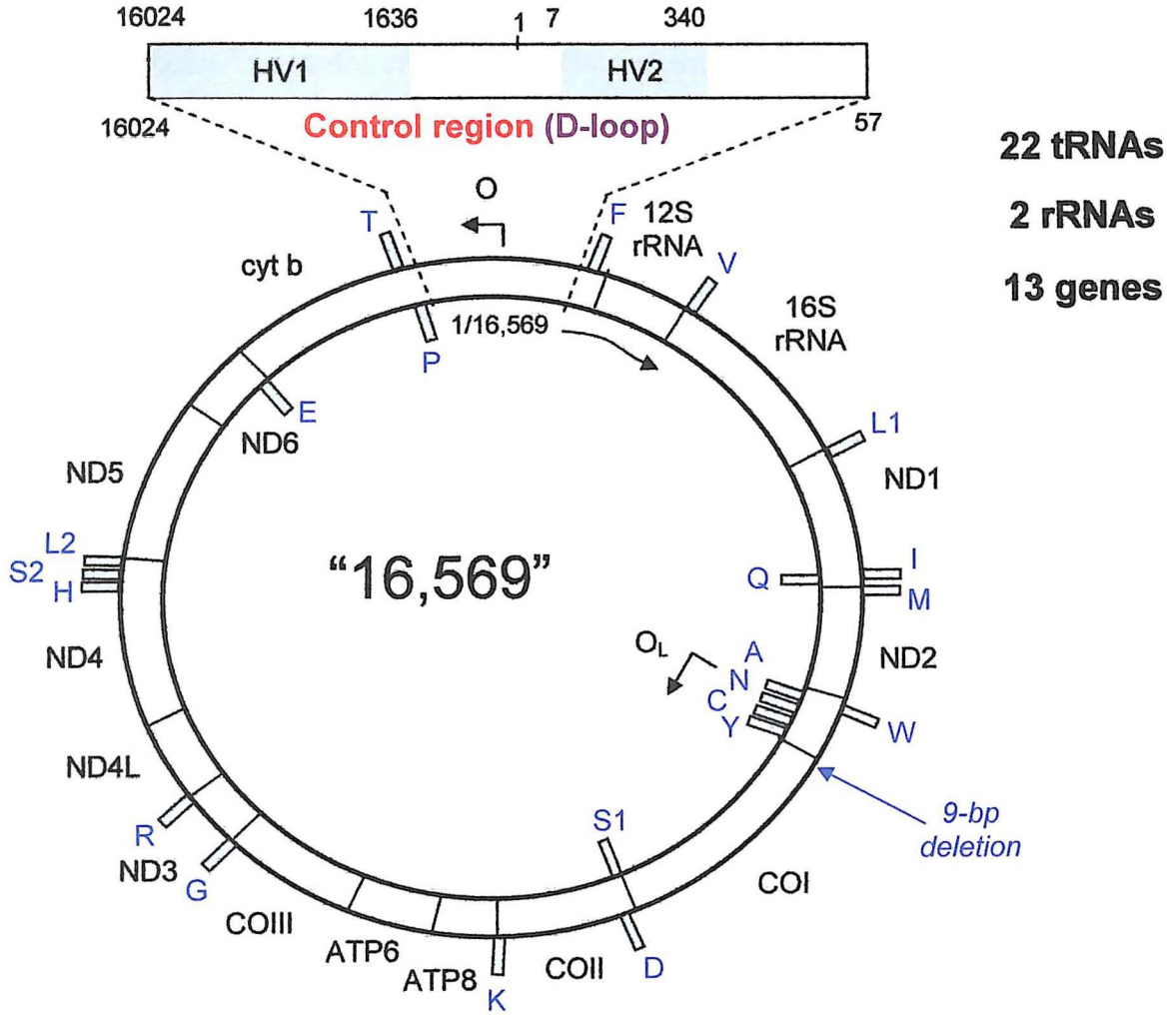


Figure 1.1: The Organisation of human mitochondrial DNA

(J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press)

1.1.2 Maternal inheritance of mitochondrial DNA

The mtDNA in mammals is maternally inherited (Giles *et al.*, 1980). The number of mtDNA molecules in oocytes exceeds 100,000, whereas sperm contains only 100 to 1500 mtDNAs (Chen *et al.*, 1995b, Manfredi *et al.*, 1997; Diez-Sanchez *et al.*, 2003). Paternal mitochondria enter the oocyte, but they are promptly removed, possibly by ubiquitin dependent proteolysis (Manfredi *et al.*, 1997; Sutovsky *et al.*, 2000). The presence of paternal mtDNA has been observed in abnormal embryos in association with certain in vitro fertilization techniques (John *et al.*, 2000), but no paternal mtDNA has been found in children born after intracytoplasmic sperm injection (ICSI), suggesting that paternal mtDNA is normally removed successfully even when introduced into the oocyte by this method (Marchington *et al.*, 2002).

It is possible, however, that the mechanism responsible for the elimination of paternal mtDNA may fail occasionally, potentially leading to maternal or paternal mtDNA mosaicism in an individual (Schwartz and Vissing, 2002). The number of mtDNA molecules within a cell experiences radical changes during the development of the oocytes, resulting in the sampling of only a small proportion of mtDNAs and subsequent replication of the sampled molecules. This phenomenon is called the mitochondrial "bottleneck" and its most likely purpose is to lower the probability of new mutations in mtDNA being allowed to pass on along the germ line (Jenuth *et al.*, 1996; Jansen and de Boer, 1998; Marchington *et al.*, 1998).

1.1.3 Advantages and application of mtDNA in Forensic DNA analysis

The mtDNA has drawn attention in forensic casework, population study and archeology for many reasons. Unlike nuclear DNA, mtDNA is maternally inherited (Case and Wallace, 1981; Giles *et al.*, 1980; Hutchinson *et al.*, 1974). Excluding mutation, the mtDNA sequence of siblings and all maternal relatives is identical. This characteristic can be helpful in forensic cases, such as analyzing the remains of a missing person, where known maternal relatives can provide reference samples for direct comparison to the questioned mtDNA type (Ginther *et al.*, 1992; Holland *et al.*, 1993).

The other striking feature of mtDNA is the existence of a high number of copies per cell, as compared with only two copies of autosomal chromosomes (Pfeiffer *et al.*, 1998). Because of the high copy number of mtDNA molecules in a cell, mtDNA typing is particularly advantageous compared to nuclear DNA typing in certain forensic analyses. In cases where the nuclear DNA typing is not likely to be successful when dealing with samples such as hair shafts, bones, teeth, and other samples that are severely decomposed or have been exposed to substantial environmental insults, mtDNA is the only hope. Besides, mtDNA is less prone to degradation than nuclear DNA because of the circular form of the mtDNA genome and its subcellular sequestration makes it less susceptible to exonuclease degradation (Parson *et al.*, 1998).

mtDNA also has a higher mutation rate compared to the nuclear genome (Parsons *et al.*, 1997). The mutation rate in mtDNA is at least 10-fold relative to that in nuclear DNA, and as a result, mtDNA evolves rapidly (Brown *et al.*, 1979; Wallace *et al.*, 1987). Studies have also shown that the substitution rate in the control region where HVS I and HVS II are situated, is about ten times higher than that in the remainder genome (Parsons *et al.*, 1997) and gives rise to more polymorphic site at this region. The high rate of mtDNA evolution, thought to be a consequence of the lower fidelity of mitochondrial DNA polymerase (Kunkel and Loeb 1981) and the increased susceptibility of the mitochondrial genome to DNA damage (Yakes and Houten, 1997), is reflected by the variability of this region, which allows comparisons between closely related species (Tamura and Nei, 1993) or intraspecies populations (Budowle *et al.*, 1999). The non-coding region especially the HVS I and HVS II can be sequenced to provide a high degree of information for discriminating between unrelated individuals. These regions (HVS I and HVS II) are of interest for human identity testing since polymorphisms are very much concentrated in these regions

1.2 Mitochondrial DNA variation in human populations

The evolutionary relationships between species can be deduced from comparisons and phylogenetic reconstructions of DNA sequences. Because of its high mutation rate, maternal inheritance and lack of recombination, mtDNA has been extensively used for understanding modern human evolution and migrations. The evolutionary orders of mammals (Arnason *et al.*, 2002) and the divergence times of humans and other primates (Glazko & Nei, 2003) have been studied using mtDNA, and similar studies have also been carried out to assess the evolution of different human populations (Richards *et al.*, 2000; Cavalli-Sforza & Feldman, 2003).

Evolutionary mtDNA analysis, in addition to Y-chromosomal, and autosomal analyses have generally confirmed the "out of Africa" hypothesis, and discover that human mtDNA lineages form a single, monophyletic phylogenetic tree which is rooted in Africa (Johnson *et al.*, 1983; Ingman *et al.*, 2000; Jorde *et al.*, 2000; Caramelli *et al.*, 2003). The high mutation rate of mtDNA has resulted in the sequential accumulation of a large number of nucleotide substitutions in mtDNA lineages that have diverged on the same time scale as human populations have colonized the different geographical regions of the world. These mutations are often assumed to be neutral and to have evolved mostly by genetic drift, but evidence is now accumulating that selection may have had a significant role during human mtDNA evolution (Mishmar *et al.*, 2003).

1.2.1 Mitochondrial DNA haplogroups

The analyses of human mtDNA lineages were based on sequence variation in the D-loop or control region, specifically in the two hypervariable segments HVS-I and HVS-II (Richards *et al.*, 1996; Wilkinson-Herbots *et al.*, 1996), and also on restriction fragment length polymorphism (RFLP) analyses of the entire genome. The analysis of non-coding region by sequencing and the coding region by RFLP, allowing the scientist screening of 15-20 % of the mtDNA sequence for variations (Chen *et al.*, 1995b).

The major human mtDNA lineages have been classified into haplogroups according to the information provided by RFLPs and the sequence of the hypervariable regions in the D-loop. The principal clusters have been denoted with capital letters, with additional letters or numbers used to denote lineages within the principal clusters (Torroni *et al.*, 1996). More recently, phylogenetic reconstruction of the human mtDNA using complete mtDNA sequences has allowed haplogroups to be defined by individual nucleotide substitutions, which provides the highest possible level of accuracy (Ingman *et al.*, 2000; Finnilä *et al.*, 2001; Maca-Meyer *et al.*, 2001; Herrnstadt *et al.*, 2002a; Mishmar *et al.*, 2003).

1.2.2 Continent-specific mtDNA lineages

Analyses of the distribution of the major mtDNA haplogroups in different continents and phylogenetic analyses have facilitated reconstruction of the human mtDNA dispersal out of Africa during evolution (Figure 2). Haplogroup L is prominent in Africa and comprising about 3/4 of all African mtDNAs. Besides, it is the oldest human mtDNA lineages (Wallace *et al.*, 1999, Ingman *et al.*, 2000) and closest to the most recent common ancestor of all human populations ("the mitochondrial Eve"), who presumably lived in central Africa approximately 130,000–200,000 years before present (YBP). Haplogroup L is further divided into several sublineages (L0, L1, L2, and L3).

Superhaplogroup M and N are the two macrolineages diverged from L, presumably in northeastern Africa or in the Middle East, approximately 65,000 years ago. The European haplogroups H, I, J, K, T, U, V, W, and X (Torroni *et al.*, 1996) were subsequently derived primarily from macrolineage N, whereas M and N contributed equally to the radiation of mtDNA into the Asian specific haplogroups A, C, D, G, Z, and Y. The American continent was populated from northeastern Asia through the Bering land bridge by humans with haplogroups A, C, and D (Derbeneva *et al.*, 2002, Silva *et al.*, 2002), whereas haplogroup B may have arrived later and via a coastal route (Starikovskaya *et al.*, 1998, Mishmar *et al.*, 2003). Thus, all the major continents were inhabited and the haplogroups had achieved their present geographical distribution some 35,000–15,000 YBP, although demographic shifts have also occurred since then (Bandelt *et al.*, 2001, Salas *et al.*, 2002).

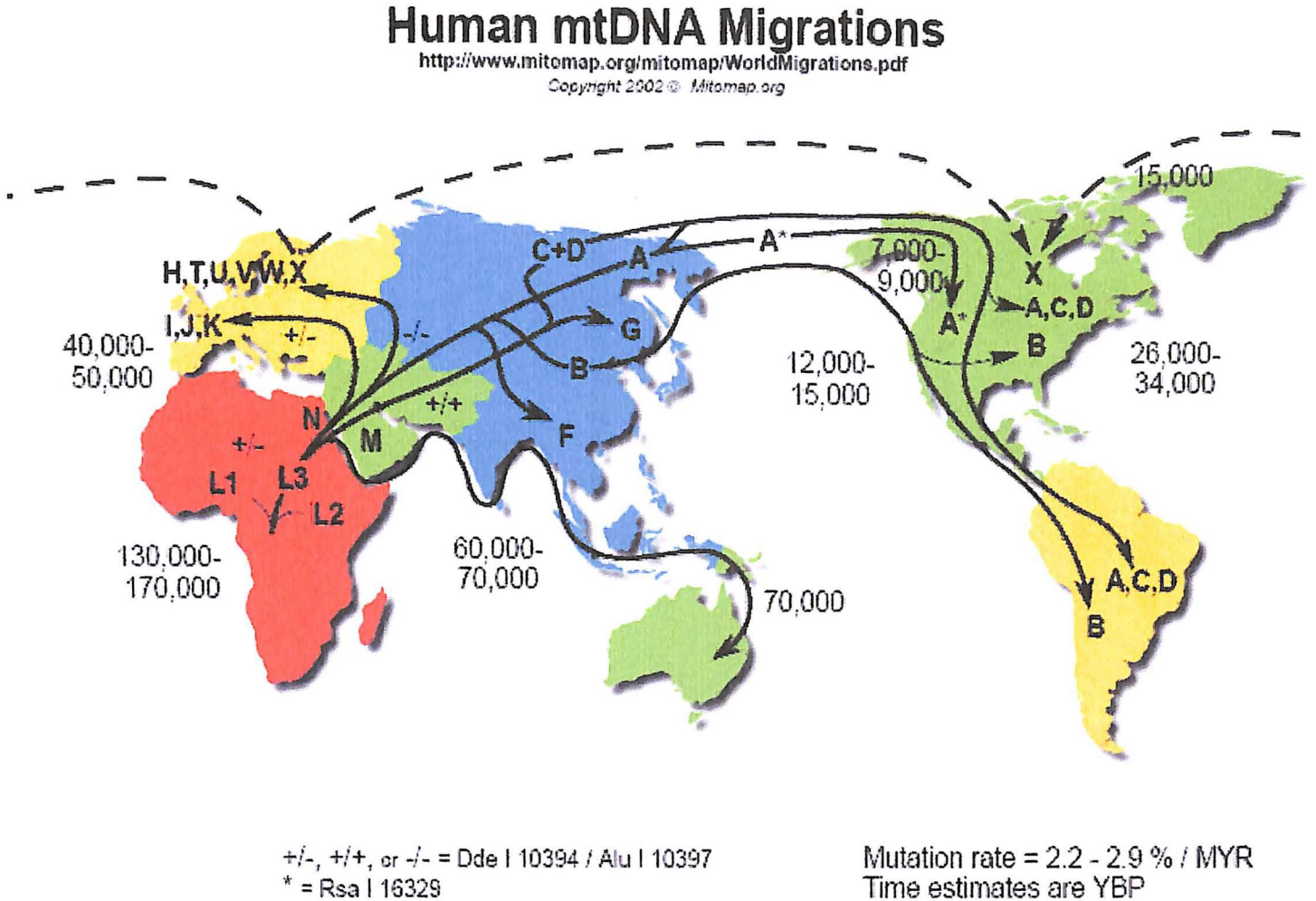


Figure 1.2: The world mtDNA human migration patterns and each continent specific haplogroup (www.mitomap.org)

1.3 Chinese Population of Peninsular Malaysia

Malaysia which is located in the heart of Southeast Asia and is divided into two main regions: Peninsular Malaysia, which lies just south of Thailand, and East Malaysia, which is situated north of Indonesia on the island of Borneo. The population of Malaysia in July 2007 is estimated to be 24,821,286 and the Chinese who are the second largest ethnic in Malaysia form about 23.7% of the total population (https://www.cia.gov/library/publications/the-world-factbook/geos/my.html). Most of them are the descendants of the Chinese from Mainland China who arrived between the fifteenth and the mid-twentieth centuries. The Chinese population in Malaysia is mostly Buddhist or Taoist. They speak a variety of Chinese dialects including Mandarin, Hokkien, Cantonese, Hakka and Teochew. A large majority of Chinese population in Malaysia, especially those from larger cities such as Kuala Lumpur, Petaling Jaya and Penang speak English as well. There has also been an increasing number of the present generation Chinese who consider English as their first language. The Chinese have historically been dominant in the Malaysian business community.

2. REVIEW OF LITERATURE

Over the past 15 years, both restriction fragment length polymorphism (RFLP) analysis and direct sequencing have been used to study the patterns of mitochondrial DNA (mtDNA) variation in Asian populations. RFLP analysis has been the primary method of detecting sequence variation in the coding regions of the mtDNA genome (Horai *et al.*, 1984; Horai and Matsunaga 1986; Cann *et al.*, 1987; Harihara *et al.*, 1988; Stoneking *et al.*, 1990; Ballinger *et al.*, 1992; Torroni *et al.*, 1993a, 1994a; Passarino *et al.*, 1993; Kivisild *et al.*, 1999; Schurr *et al.*, 1999), whereas sequencing studies have primarily targeted the noncoding control region (CR) to detect nucleotide diversity among individuals (Lum *et al.*, 1994; Sykes *et al.*,1995; Horai *et al.*, 1996; Kolman *et al.*, 1996; Lum and Cann 1998; Richards *et al.*, 1998; Redd and Stoneking 1999). These studies have clearly shown that a number of distinctive mtDNA lineages, or haplogroups, are present in these populations, with some of them having undergone considerable differentiation. Furthermore, many sequence polymorphisms are known to have certain ethnic groups or geographic regions (e.g., Papua New Guinea [PNG] populations; Stoneking *et al.*, 1990; Ballinger *et al.*, 1992; Redd and Stoneking 1999), thereby permitting the reconstruction of the genetic history of these mtDNAs, and potentially that of the populations in which they occur.

There are several different mtDNA lineages that are present in Southeast Asia, namely, haplogroups B, F, and M. These haplogroups are present in Vietnamese and aboriginal groups from Malaysia (Orang Asli) and Borneo to varying degrees, as well as in other Southeast Asian groups. Their distribution in this geographic region allows us to make certain inferences about the genetic relationships of the populations that possess them.

Furthermore, these mtDNA lineages appear to demarcate major expansions of Asian peoples, whether the initial colonizers of this region or the most recent expansion of Austronesian speakers.

Superhaplogroup M is a dominant mtDNA cluster among the populations of Mainland Asia as well as among Native Americans (Ballinger *et al.*, 1992). It is defined by the presence of a *Dde* I site at np 10394 and an *Alu* I site at np 10397. Superaplogroup M occurs in all Southeast Asian populations at varying frequencies (25%-45%), with the highest frequencies occurring in the Malays and Sabah Aborigines (~60%). As shown in several analyses of mtDNA haplotypes (Kivisild *et al.*, 1999; Macaulay *et al.*, 1999; Quintana-Murci *et al.*, 1999), Superhaplogroup M is actually a macrohaplogroup, meaning that M represents the founding or stem haplogroup from which all subsequent haplogroups bearing the +*Dde*I/+*Alu* I sites evolved. In the same way, haplogroup L encompasses all LI and L2 mtDNAs found in African populations (Chen *et al.*, 2000; Watson *et al.*, 1997).

Therefore, any mtDNA with the +*Dde* I/+*Alu* I sites can be said to belong to this macrohaplogroup. Thus, haplogroups C, D, E, G, and Z (Torroni *et al.*, 1992, 1993a, 1994a; Schurr *et al.*, 1999) can be considered smaller branches of superhaplogroup M. Most of these smaller haplogroups have retained their independent status as mtDNA lineages because the designations for haplogroups A-L preceded the naming of macrohaplogroup M. However, some researchers have been renaming these smaller haplogroups as variants of M itself (MI-M7) (e.g., Quintana-Murci *et al.*, 1999; Kivisild *et al.*, 1999; Forster *et al.*, 2001), a trend that may ultimately supplant the other designations.

Another mtDNA lineage that is fairly widespread throughout Southeast Asia is haplogroup F. However these two haplogroups are not as genetically diverse as superhaplogroup M. Haplogroup is defined by the associated *–Hinc* II (12406) and *–Hpa* I (12406) polymorphisms, with many of these also having the associated *–Hae* II (9052) and *–Hha I* (9053) mutations (Ballinger *et al.*, 1992; Torroni *et al.*, 1993a, 1994a). Haplogroup F appears in a number of Asian populations, including Filipinos (Cann *et al.*, 1987) and Aboriginal Taiwanese (Melton *et al.*, 1995, 1998), and occurs at its highest frequencies in Southeast Asia, specifically 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 linguistic ties (Austro-Asiatic family; Bellwood 1979). The frequency distribution of this mtDNA lineage shows that it decreases in both northerly and southeasterly directions from Southeast Asia, but is widespread in East Asia itself, being seen as far north as central Siberia (Evenks) (Torroni *et al.*, 1993a) and as far south as Borneo (Kadazan).

Based on its distribution, haplogroup F may have been disseminated throughout East Asia through some kind of a population expansion. It does not appear to be as diverse as haplogroup M, based on the RFLP composition of its haplotypes (Horai *et al.*, 1984; Harihara *et al.*, 1988; Ballinger *et al.*, 1992; Torroni *et al.*, 1993a, 1994a), and, in fact, is found within the geographic range of superhaplogroup M. In addition, while the distribution of haplogroup F overlaps that of haplogroup B, its greatest haplotypic diversity occurs in Vietnam, not in Taiwan, the Philippines, and Indonesia, as is the case for haplogroup B (Melton *et al.*, 1995, 1998; Richards *et al.*, 1998).

As shown by various studies, deletion mtDNAs from haplogroup B are broadly distributed in Asian populations. Haplogroup B mtDNAs are also found in Vietnamese, Malaysian, and Bornean populations (Ballinger *et al.*, 1992) as well as Philippines, Indonesia, and Melanesia (Melton *et al.*, 1995, 1998). A population tree of Southeast Asian (Ballinger *et al.*, 1992) populations reveals that Malays and Sabah (Borneo) Aborigines are closer to each other than to Malaysian Chinese and Vietnamese.

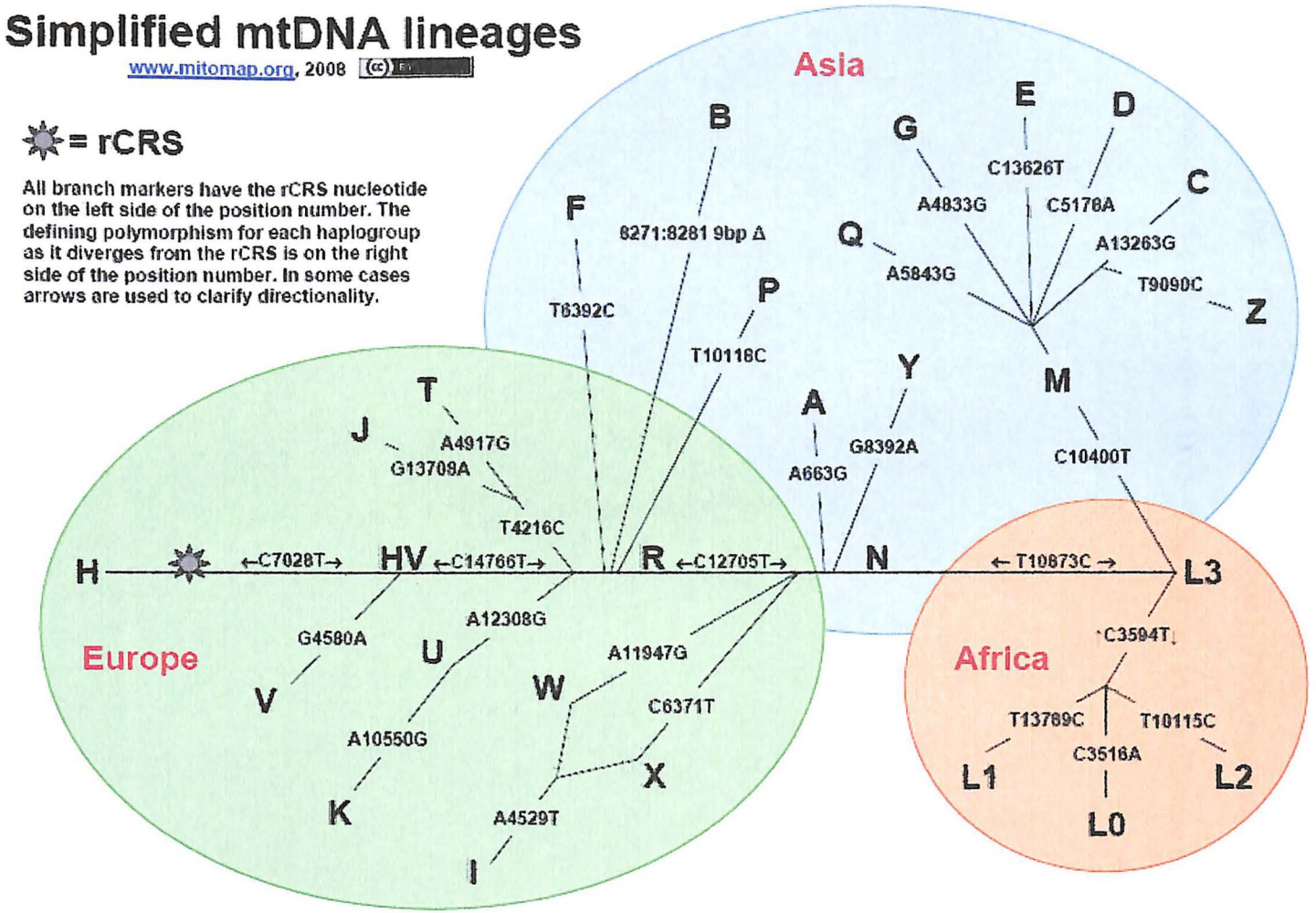


Figure 2.1: Simplified tree of mitochondrial haplogroups (www.mitomap.org)

3. OBJECTIVES OF THE STUDY

The objectives of the study are:

- To determine the HVS II sequence polymorphisms of the Malaysian Chinese population
- To classify the mtDNA haplogroup based on HVS II sequence variation
- To calculate the frequency of HVS-II haplotypes in the Chinese population.

4. MATERIALS AND METHODS

4.1 MATERIALS

4.1.1 PCR primers

The PCR primers (HVC1 and HVD1) were synthesized by Invitrogen. The primers were manually designed by using the Cambridge References Sequence. Table 4.1 list the primer sequences applied in this study which is used for the amplification of HVS-II region as well as for sequencing analysis

Table 4.1: Sequences of each primer used in the study

Primers	Sequence (5' to 3')
HVC1 –F (LO48)	CTC ACG GGA GCT CTC CAT GC
HVD1– R (H408)	CTG TTA AAA GTG CAT ACC GCC A

4.1.2 ReddyMix™ PCR Master Mix

ReddyMix™ PCR Master Mix was purchased from ABgene, UK. It is a ready-to-use master mix for PCR reactions, which is a convenient way of amplifying DNA fragments without the need of thawing individual components, reducing the risk of contamination and pipetting errors. ReddyMix™ PCR Master Mix also contains a dye and precipitant to facilitate the gel loading. Each vial contains 1.8 mL of a 1.1 X working concentration PCR Master Mix which is sufficient for 40 X 50 µL reactions. The final reaction volume is 50 µL after the addition of DNA template, primers and ddH_2O. The composition of ReddyMix™ PCR Master Mix is shown in Table 4.2

Table 4.2: Composition of ReddyMix™ PCR Master Mix

Reagents	
1.25 units	ThermoPrime plus DNA polymerase
75 mM	TrisHCL
20mM	$(NH_4)_2SO_4$
1.5mM	$MgCl_2$
0.1% (v/v)	Tween® 20
0.2mM	each of dATP, dCTP, dGTP and dTTP
Precipitant and red dye for electrophoresis	

4.1.3 QIAquick PCR Purification Kit

QIAquick PCR Purification Kit was purchased from QIAGEN Inc., USA. QIAquick Kits contain a silica membrane assembly for binding of DNA in high-salt buffer and elution with low-salt buffer or water. The purification procedure removes primers, nucleotides, enzymes, mineral oil, salts, agarose, ethidium bromide, and other impurities from DNA samples. This kit consisted of: QIAquick Spin Columns, Buffers (150 mL Buffer PBI, 55 mL Buffer PE, 15 mL Buffer EB) and collection tubes (2 ml).

4.1.4 DNA Ladder

1 kb DNA ladder which has been used in this study was purchased from Invitrogen, USA. The ladder contains a total of twenty bands: twelve bands ranging in size from 1000 bp to 12,000 bp in 1000-bp increments and eight bands ranging in size from 100 to 1650 bp. The concentration of DNA ladder was 1 μg/μl and was stored at -20 °C. DNA ladder was used to determine the size of unknown sample by comparing the distance of sample traveled relative to the marker.

4.1.5 Loading Dye

Loading Dye was purchased from QIAGEN. It is used for analysis of purified DNA samples using electrophoresis. It contains 3 marker dyes (bromophenol blue, xylene cyanol and orange G) that facilitate estimation of DNA migration distance and optimization of agarose gel run time.

4.1.6 10X TBE buffer

To prepare this buffer, 53.9 g of Tris-base and 3.72 g of EDTA was dissolved in dH_2O. Then, 27.5 g of boric acid was added to the mixture solution. pH of the solution was adjusted to 8.3 by addition of boric acid and finally was autoclaved before being used in electrophoresis.

4.2 METHODS

4.2.1 Sterilisation

All plastic ware, glassware, medium, deionized distilled water, PCR tubes, Eppendorf tubes, pipette tips and instruments that are needed to be in sterile condition were autoclaved at 20 psi for 15 minutes.

4.2.2 Separate Working Areas

The separate work area was used for adding DNA sample to the PCR Master Mix solution. To prevent transfer of exogenous human DNA into the PCR setup area, dedicated equipment and supplies such as pipettes, consumable tips and gloves used for PCR setup should not be taken out of the PCR setup work area. Besides, all PCR reactions and primer dilutions were carried out in a laminar hood. All work surfaces were thoroughly cleaned with isopropyl alcohol or ethanol prior and after all works. Post-PCR products were handled in a separate amplified DNA working area in order to prevent contamination of the amplified DNA. This working area is a physically separated area used only for those activities that involve the handling of amplified DNA. This includes DNA typing, gel electrophoresis of amplified DNA, waste disposal of amplified DNA solutions, and storage of amplified DNA.

4.2.3 DNA Extraction from Buccal Swabs

The genomic DNA was extracted from buccal swabs collected from 37 Chinese individuals by using standard phenol / chloroform methods by previous final year students.

4.2.4 Polymerase Chain Reaction (PCR)

The HVS-II region was amplified using a pair of primers, HVC1 and HVD1. PCR reaction mixture was prepared in a 0.2 mL tube which consists of 50 µL of the final reaction volume. The final PCR reaction mixture composed of 2 µL of each forward and reverse primer, 1 µL of DNA template, 1 µL of deionized distilled water and 44 µL of ReddyMix™ PCR Master Mix. The PCR program involved an initial denaturation at 95 °C for 3 minutes, followed by 30 cycles of denaturation at 95 °C for 30 sec, annealing at 60 °C for 30 sec, extension at 72 °C for 30 sec and a final extension step at 72 °C for 5 minutes.

Table 4.3: PCR reaction volume for amplification of HVS-II

Reagent		**Volume (µL)**
ReddyMix Master Mix		44.0
DNA Template		1.0
Primers	F	2.0
	R	2.0
ddH_2O		1.0
Total		50.0

Table 4.4: PCR Program

Condition	Temperature (°C)	Time	
Initial denaturation	95	3 min	
Denaturation	95	30 sec	30 cycles
Annealing	60	30 sec	
Extension	72	30 sec	
Final extension	72	5 min	
Hold	4		

4.2.5 Agarose Gel Electrophoresis

To ensure that PCR process was successful, the amplified products were loaded into 1 % agarose gel. The agarose gel was prepared by dissolving 1 g of agarose powder in 100 mL of 0.5 X TBE buffer (diluted from the 10 X TBE buffer prepared earlier). The mixture was first heated for 3 minutes in the microwave oven until completely melt and allowed to cool to about 60 °C before adding 1 µL of ethidium bromide. The gel was then poured slowly into the casting tray containing a sample comb and left to solidify for about 30 minutes. After the gel has solidified, the comb was removed, and the gel was inserted horizontally into the electrophoresis chamber. TBE buffer (0.5 X) was poured into the chamber to submerge the gel. A total of 2 µL of PCR products were pipetted into the wells and 2 µL of 1 kb DNA marker was loaded into the first well. The PCR products were then electrophoresed at 90 V for 45 minutes. After the electrophoresis has completed, the gel was visualized under ultraviolet transilluminator and a pictures was taken for record.