

# รายงานฉบับสมบูรณ์

การวิเคราะห์ความสัมพันธ์ทางวิวัฒนาการของดีเอ็นเอจากไมโทคอนเดรียและ

โครโมโซมวายในประชากรภาคเหนือตอนล่างของประเทศไทย

(R2560B134)

ดร.สายศิริ มีระเสน

คณะวิทยาศาสตร์การแพทย์ มหาวิทยาลัยนเรศวร

สำนักหอสมุด มหาวิทยาลัยนเรศวร

วันลงทะเบียน.....

เลขทะเบียน..... 1047485

เลขเรียกหนังสือ..... จ. ๗๕

๖๐5  
๗5  
๕๕55  
๕๕๐๐

ทุนอุดหนุนการวิจัยจากงบประมาณแผ่นดิน ประจำปีงบประมาณ พ.ศ. ๒๕๖๐

**PHYLOGENETIC RELATIONSHIP ANALYSIS OF  
MITOCHONDRIAL DNA AND Y-CHROMOSOME  
IN LOWER NORTHERN PART OF  
THAILAND POPULATIONS  
(R2560B134)**



**Funding of the National Research Council of Thailand, 2017**

## ABSTRACT

The lower northern part of Thailand contains various genetically diverse ethnic populations. The sequences of the mitochondrial DNA hypervariable region were studied in three ethnic populations inhabiting Phitsanulok Province. One hundred and nine nucleotide sequences—53, 29, and 27 from Hmongs (Hill tribe), Lao Songs, and Thai-Siams, respectively—were collected. The haplotypes were generated from 1130 nucleotides of the entire control region. Eighty-six haplotypes were found in the three ethnic populations, and no shared haplotypes were found between populations. Point heteroplasmy was noted at position 311 (C-->Y). Haplotypes with ACAC-insertion at position 512 were observed in immigrant individuals from the Lao Song population. The Thai-Siam population showed higher genetic diversity than the other populations. The Hmong and Lao Song populations showed less genetic diversity than those living in their original area. The neutrality testing suggested that the result might be influenced by genetic drift. Five loci of Y-chromosome short tandem repeat, DYS19, DYS390, DYS391, DYS392 and DYS393, were analyzed in 281 unrelated males inhabited in lower northern part of Thailand. The allele frequencies were calculate by counting method. Each loci shown high level of gene diversity except DYS391. The combination of loci presented a total of 165 haplotypes which were 112 unique haplotypes. The haplotype diversity of overall were 0.9928. The overall haplotypes shown the genetic mixture of Thai population as a variety of the Southeast Asian. This result can provide the preliminary data for study the genetic relationship of population in Southeast Asia.

## LIST OF CONTENTS

| Chapter   | Page     |
|---|----------|
| <b>I INTRODUCTION</b>                                       | <b>1</b> |
| Statement of purpose  | 1        |
| Objectives of the study                                     | 2        |
| Expected outputs of the study                               | 2        |
| Expected outcomes   | 2        |
| <b>II LITERATURE REVIEW</b>                                 | <b>3</b> |
| Population and Ethnicity in Lower Northern part of Thailand | 3        |
| Genetic markers of Human Population study                   | 4        |
| Mitochondrial DNA polymorphism                              | 5        |
| Mitochondrial DNA properties                                | 7        |
| Human mtDNA Applications                                    | 10       |
| Y-chromosome polymorphisms                                  | 11       |
| Y-STR haplotype database                                    | 14       |
| Interpretation of Y-STR results                             | 15       |
| Statistical test  | 15       |
| 1. Number of polymorphic sites                              | 16       |
| 2. Mean number of pairwise differences ( $\pi$ )            | 16       |
| 3. Estimation of genetic distances between DNA sequences    | 16       |
| 3.1 Pairwise difference                                     | 16       |
| 3.2 Percentage difference                                   | 16       |
| 4. Estimation of distances between Y-STR haplotypes         | 17       |
| 4.1 No. of different alleles                                | 17       |
| 4.2 Sum of squared size difference                          | 17       |
| 5. Neutrality tests   | 18       |
| 5.1 Tajima's test of selective neutrality                   | 18       |
| 5.2 Fu's FS test of selective neutrality                    | 18       |

## **LIST OF CONTENTS (cont.)**

| <b>Chapter</b>                              | <b>Page</b> |
|---|-------------|
| 6. The population pairwise genetic distance | 18          |
| Population genetic study in global scale    | 19          |
| Population genetics study in Thailand       | 25          |
| <b>III RESEARCH METHODOLOGY</b>             | <b>30</b>   |
| Hypothesis                                  | 30          |
| Ethical consideration                       | 30          |
| Population and sampling                     | 31          |
| Sample size                                 | 32          |
| Blood collection                            | 32          |
| DNA extraction                              | 33          |
| DNA analysis                                | 33          |
| 1. The mitochondrial DNA study              | 33          |
| Statistical analysis                        | 34          |
| 2. The Y-STR study                          | 36          |
| <b>IV RESULTS AND DISCUSSION</b>            | <b>38</b>   |
| Samples collection                          | 38          |
| DNA analysis                                | 39          |
| MtDNA control regions analysis              | 39          |
| Y-STR analysis                              | 52          |
| <b>V CONCLUSION</b>                         | <b>55</b>   |
| <b>REFERENCE</b>                            | <b>55</b>   |
| <b>APPENDIXES</b>                           | <b>61</b>   |
| <b>BIOGRAPHY</b>                            | <b>80</b>   |

## LIST OF TABLE

| Table  | Page |
|--|------|
| 1 Summary of statistical parameter for the mtDNA   | 9    |
| 2 Characteristics of commonly used Y-chromosome STR loci   | 14   |
| 3 Summary of available online Y-STR databases  | 15   |
| 4 Sequences of primers used for amplification the entire sequence<br>of D-loop hypervariable region of mtDNA                                 | 35   |
| 5 Sequence of Y-STR primers  | 36   |
| 6 Characteristics of 5 Y-STR loci  | 37   |
| 7 Characteristic of three ethnic populations in Phitsanulok province   | 39   |
| 8 Distribution of segregating sites and mutations in HVR of mtDNA<br>control regions in the three ethnic populations in Phitsanulok Province | 41   |
| 9 Observed mutations in 1,130 nucleotides of mtDNA control region<br>in studied populations  | 42   |
| 10 Numbers of haplotypes, as well as unique and shared haplotypes<br>observed in the three ethnic populations                                | 45   |
| 11 Frequencies of haplotypes in the three ethnic populations.  | 48   |
| 12 Diversity and neutrality testing parameters of mtDNA control region<br>haplotypes in the three ethnic populations.                        | 49   |
| 13 The genetic differentiations by nucleotide diversity between<br>groups of 3 ethnic populations  | 50   |
| 14 Allele frequencies and gene diversity ( $h$ ) of 5 Y-STR loci   | 52   |
| 15 The frequencies of 5 loci Y-STR haplotypes  | 53   |

## LIST OF FIGURE

| Figure   | Page |
|--|------|
| 1 Human mitochondrial genome   | 7    |
| 2 Position of human mtDNA control region   | 7    |
| 3 Pattern of single nucleotide polymorphisms (SNPs) and haplotype                                    | 8    |
| 4 Single nucleotide polymorphisms and short tandem repeats   | 12   |
| 5 Relative position of Y-STR   | 13   |
| 6 UPGMA tree for 26 representation human populations   | 21   |
| 7 NJ tree for 26 representation human populations  | 22   |
| 8 Y-haplogroups of the world   | 23   |
| 9 The simplified tree of Y-haplogroups   | 23   |
| 10 The mitochondrial haplogroups of the world  | 24   |
| 11 The simplified tree of mitochondrial haplogroups  | 24   |
| 12 NJ tree for ten population's groups studies by $\beta^E$ -globin haplotype analysis base on $d_A$ | 25   |
| 13 The relationship of population based on $d_A$ in eight Thai populations                           | 26   |
| 14 Unrooted NJ tree of the 14 populations based on the net genetic distances ( $d_A$ )               | 27   |
| 15 The genetic relationship among Mon-Khmer speaking and Tai-Kadai speaking                          | 29   |
| 16 Geographic distribution of three ethnic populations   | 32   |
| 17 Position of primers which were used to amplify the entire control region of mtDNA                 | 34   |
| 18 Chromatograms of mtDNA control region from 2 sequences  | 43   |
| 19 Frequency distribution of mutation at variation sites in mtDNA control regions                    | 44   |
| 20 Haplotype distribution in three ethnic populations  | 46   |
| 21 NJ-tree of HVR-1 in 3 ethnic population   | 51   |

# CHAPTER I

## INTRODUCTION

### **Statement of purpose**

The region known as lower northern part of Thailand is the center of Indo-China region because its prominent location which is connecting the northern Thailand to the neighboring countries. There are several major and important highways passing through the province, including one leading to Myanmar and Laos and others from the central part to the northern Thailand. Since the region is a frequent-travelled route by many travelers of all ethnicities, it would influence high diversity in populations. The Siam tribe, Lao tribe and some hill tribes such as the Hmong can be found here. Each group still maintains their unique cultures and traditions. They had been preserved is their genetic structure which is crucial for conducting a genetic study of a population. On the other hand, population migration has changed the genetic structure. The genetic structure of the populations is going to be rapidly diversifying due to the Asian Economic Community Initiative (AEC Initiative) that bring more foreign influx. As a result, now more than ever, it is necessary to study the genetic structure of the population in this region before too much change occurs.

In this research, the genetics structure of population in the Lower Northern part of Thailand will be studied by two molecular markers: the mitochondrial DNA at the D-loop control region and the Y-chromosome short tandem repeats.

The data in this research aims to support the genetic reference database of Thai populations and there are not only the basis of population genetics in Thailand but might be applied in ethnology, forensic science and other fields such as the study on genetics disease and personal pharmaceutical.



**Objectives of the study**

1. To evaluate the genetic diversity and estimate the genetic relationship within/between populations of Thai Siam, Lao Song and Hmong in lower northern part of Thailand using mtDNA sequences and Y-chromosome Short Tandem Repeat (Y-STR) as molecular markers
2. To investigate the possible haplotypes of mtDNA or Y-chromosome STR polymorphism which are associated to ethnic populations.

**Expected output of the study**

To establish the genetic database of populations in Lower Northern part of Thailand.

**Expected outcome**

The database can be applied to ethnological science, forensic science, medical science and others.

## CHAPTER II

### LITERATURE REVIEW

#### **Ethnicity and population in lower northern part of Thailand**

Ethnology is the branch of anthropology that characterized and compared the differentiation and relation of people. It indicates the social properties and explains the status of populations. In the past, the physical traits and ancestor were used to classified human as “race” in term of human and social evolutions. It established the racism and controversy. However, there are many factors such as social mechanism, communication and interpopulation marriage, resulting in the complication of human classification. Nowadays, the anthropologist proposed to use social contexts such as language, costume, culture and traditions to classified people into “ethnic”, which have been more diverse. Their lifestyles had been changed rapidly according to the changing of social, economic and culture context. The ethnic groups in Thailand were classified into 5 language families which were Tai-Kadai, Austro-Asiatic, Sino-Tibetan, Hmong-Mien and Austronesian. Tai-Kadai is the largest ethnic group including those who lived in Thailand and used Thai language since Thailand was Siam, in Sukhothai era and others whose ancestor originated in China, Laos, Vietnam or Burma (Myanmar). The native Thai or **Thai Siam**, who speaks central Thai language, is the most common group in Thailand. There are more than 20 ethnic groups identified such as Tai Yai, Lue, Song, Yuan, Meung (northern Thai), Paktai (Southern Thai), Lao Lom, Lao Ngaew, Lao Klang, Lao Wieng, etc. (Schliesinger, 2001).

The lower northern part of Thailand consists of nine provinces –Tak, Phitsanulok, Sukothai, Petchabun, Pichit, Kamphaeng Phet, Nakhon Sawan and Uthai Thani. Phitsanulok, as a center of this region, has several highways passing through, from Central to North plain, and connecting western to north Eastern plain. It is the important area for economic affair. Several population migrated in to this region, since Sukothai era resulting in high varieties of populations (Satapanawatana, Sirasunthorn,

& Suwanpratest, 2011). Culture, belief and ethnicity presented the identity of each population. The Thai Siam who used Thai language with Ayutthaya or Sukothai dialect can be observed in Meung district and surrounding area of district. The Laos tribes, Lao Luang Prabang, Lao Vientiane and Lao Song for examples living in some villages in Bang Rakam and Wang Thong districts. Hill tribes such as Hmong settled at Chat Trakan and Nakhon Thai districts. Chinese and Indian inhabited in commercial area of Meung district. Americans, Europeans and Vietnamese are found in less frequency. The ethnic populations and their languages with accents in Phitsanulok province were identified and mapped on their geographical areas, (Ngourungsi et al., 1976).

**Lao Song**, the Tai-Kadai speaking group, is one of minority groups living in the area near the border of Southern China, Laos and Vietnam for more than 200 years. They migrated to Thailand in the early Rattanakosin period (1778-1779 A.D.). The first wave of Lao Song could be found in the central part of Thailand; Phetchaburi, Suphan Buri and Nakhon Prathom provinces. Then they travelled upwards to lower northern part of Thailand, Phitsanulok province (Srising, 1976).

**Hmong** is the most dispersed hill tribe in Southeast-Asia, from China to Myanmar, Vietnam, Lao and Thailand. They were belonged to Miao ethnicity who used Hmong-Mein (Miao-Yao) language family. In the 16<sup>th</sup> century, Hmong were forced to leave their original habitat and moved to Yunan because of the Han expansion. Moreover, they also moved to the northern part of Laos and Vietnam. During the Indo-China conflict, the areas were too affected by the war at the time, Hmong then continued their migration to several areas in northern part of Thailand. Some settled down in Phitsanulok province, around Nakhon Thai district (Schliesinger, 2000).

### **Genetic markers of human population study**

After the finishing of the Human Genome Project, the knowledge of human molecular genetics were applied in many areas, not only medical science or pharmacogenomics, but social science such as ethnology and anthropology were influenced the plenty genetic data. The genetic variation was used to study the genetic

structure of populations and several ethnic groups. The researcher can trace the demographic or the migration of interested population by tracking the pattern of genetic structure in each population and compare to the others.

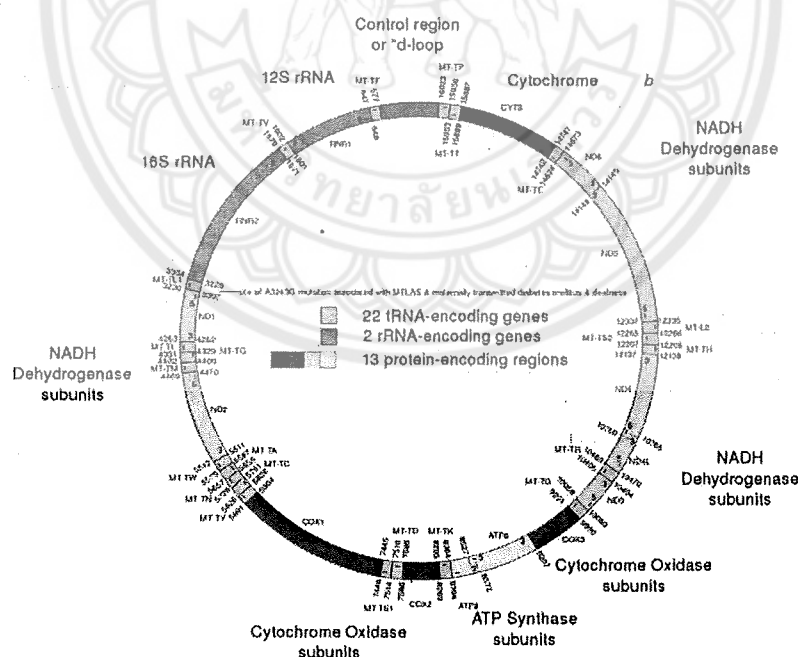
Many types of genetic markers were used to study the structure of populations. The autosomal markers such as the HLA-typing (Cao, et al., 2001), blood groups (Ward, et al., 1975), and hemoglobin E distribution (Fucharoen, et al., 2001a) were used to study the gene frequencies in populations. But there are some complexity of autosomal markers. The first is the recombination of inheritance. In each generation, half of autosomal DNA of individual are shuffled from their mother and father. That makes its difficulty for data analysis. Second, the selective effect of the chosen gene brings more sophisticated effect of population evolution. Then, the new choices of markers were developed. The mitochondrial DNA and Y-chromosome were chosen. Aiming to trace the history of population and compare the genetic relationship between populations.

The mitochondrial DNA and the Y-chromosome can be used as the lineage markers because, both are passed from generation to generation without gene recombination although the mutation event can occur in rare cases. Maternal lineage can be traced with mitochondrial DNA sequence information, whereas paternal lineage can be detected with Y-chromosome markers. The genetic information of both mitochondrial DNA and Y-chromosome are called as haplotype rather than genotype because there is only one set of haplotype in each individual.

### **Mitochondrial DNA**

The mitochondrial genome is a closed circle of double-stranded, antiparallel, double-helix DNA. The complete sequence was published in 1981 by Anderson and his colleagues. It consists of 16,569 base pairs which contains the information for encoding 13 proteins, 22 tRNAs and 2 rRNAs (Figure1). All are responsible for mitochondrial respiration (Asin-Cayuela & Gustafsson, 2007; Mercer et al., 2011). The genome of mtDNA is widely studied in the medicinal field because more than 250 mutations of mtDNA were identified as important causes of diseases due to oxidative

phosphorylation defects (Tuppen, Blakely, Turnbull, & Taylor, 2010). The two strands of mitochondrial DNA are called “H” for heavy strand and “L” for light strand because the H-strand contains higher purines (adenine and guanine) composition than L-strand as well as the purines are heavier molecular weight than pyrimidines. The replication process of mitochondria initiated on H-strand at a specific origin but L-strand initiated later. It is result in the formation of a displacement loop (D-loop) (Chang & Clayton, 1985). The mitochondrial replication occurs autonomously and is separated from nuclear genome (Clayton, 2000). The most sense sequence of RNA is lined on L-strand which was defined as the main coding strand. The base composition of the L-strand are 24.7% T, 31.2% C, 30.9% A and 13.1% G. The first complete sequence of human mitochondria also called the Cambridge reference sequence (CRS) (Anderson et al., 1981). However, the CRS was revised and found 11 nucleotide errors and 7 rare nucleotide polymorphisms. The CRS was revised and newly named as **revised Cambridge reference sequence (rCRS)** (Andrew, et al., 1999) which was favorable tool when compared with others.

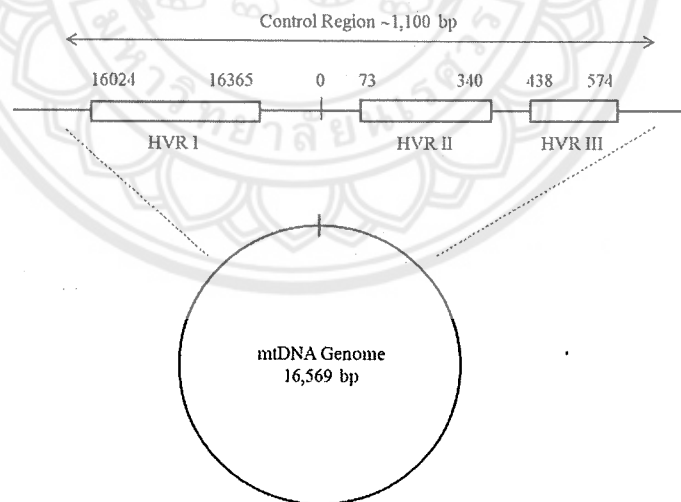


**Figure 1 Human mitochondrial genome, a 16,569-bp (base pair) circle of double-stranded DNA containing 37 genes, 2 ribosomal RNAs, 22 transfer RNAs and 13 polypeptides (Ganfyd, 2010)**

### Mitochondrial DNA properties

MtDNA has some important features such as high polymorphism, maternal inheritance and high copy number which were attractive to many fields of studies.

First, the mitochondrial DNA had higher mutation rate than the autosomal DNA, especially in the D-loop control region which was approximately 1,100 base pairs in size and divided into three distinct hypervariable regions (HVR) called HVR-1, HVR-2 and HVR-3 (Figure 2). HVR-1 was composed of 342 base pairs which was located on base 16,024 to 16,365, while HVR-2 was composed of 268 base pairs which was located on base 73 to 340 and HVR-3 was composed of 136 base pairs which was located on base 438 to 574. HVR-1 and HVR-2 consisted of many mutation hot spot positions (Parsons, Muniec, & Sullivan, 1997). The most polymorphic characters found in both HVR-I and HVR-II were single nucleotide polymorphisms (SNPs), although some position were length polymorphism such as poly-c at position 16184 to 16193 and position 303 to 309 (Meyer, Weiss, & von Haeseler, 1999).



**Figure 2 Position of human mtDNA control region**

The mutation rate ( $\mu$ ) of mitochondrial genome was approximately  $1.70 \times 10^{-8}$  substitutions per site per year (Ferris, et al., 1981) while mutation rate of whole genome was between  $1.6-3.4 \times 10^{-8}$  per site per generation approximately, assuming the generation time of 25 years (Nachman & Crowell, 2000). High mutation rate brings to high polymorphism, this leads to both HVR-1 and HVR-2 were the useful marker for evolution and forensic science studies. However, the polymorphic character of HVR-3 consists of both SNPs and length variation which were dinucleotide repeats. There were very few studies which targeted to HVR-3. The diversity of mtDNA were characterized due to pattern of SNP or polymorphism occurred in each position. Set of nucleotide polymorphisms was called haplotype (Figure 3).

| a. SNPs      | SNP1                         | SNP2 | SNP3 |
|--------------|------------------------------|------|------|
| DNA1         | ACTAAGC....GGTCCT....CTAAATG |      |      |
| DNA2         | ACTAAGC....GGACCT....CTAAATG |      |      |
| DNA3         | ACTGAGC....GGTCCT....CTAAATG |      |      |
| DNA4         | ACTGAGC....GGACCT....CTATATG |      |      |
| b. Haplotype |                              |      |      |
| Haplotype1   | ACTAAGC....GGTCCT....CTAAATG |      |      |
| Haplotype2   | .....A.....                  |      |      |
| Haplotype3   | .....G.....                  |      |      |
| Haplotype4   | .....G.....A.....T...        |      |      |

**Figure 3 Pattern of single nucleotide polymorphisms (SNPs) and haplotype**

The high polymorphic property of mtDNA brought to high genetic diversity ( $\pi$ ) of D-loop which was approximately  $1.1-1.8 \times 10^{-2}$ , but the total mitochondrial genome (excluded the D-loop) was approximately  $2.3-4.6 \times 10^{-3}$  (Table 1). The  $\pi$  values reported in table 1 were belonged to entire sequence, D-loop and rest (all except the D-loop), which reported in groups of Africans and non-Africans. It showed higher genetic diversity in Africans than others (Ingman, et al., 2000).

**Table 1 Summary of statistical parameter for the mtDNA.** Length, aligned sequence length excluding gaps; n, number of sequences; S, number of segregating sites; MPSD, mean pairwise sequence difference; and  $\pi$ , genetic diversity (Ingman, et al., 2000).

|        | Data set     | Length | n  | S   | MPSD | $\pi$                |
|--------|--------------|--------|----|-----|------|----------------------|
| Total  | All human    | 16,553 | 53 | 657 | 61.1 | $3.7 \times 10^{-3}$ |
|        | Non-Africans | 16,555 | 32 | 358 | 38.5 | $2.3 \times 10^{-3}$ |
|        | Africans     | 16,556 | 21 | 367 | 76.7 | $4.6 \times 10^{-3}$ |
| D-loop | All human    | 1,118  | 53 | 141 | 17.2 | $1.5 \times 10^{-2}$ |
|        | Non-Africans | 1,118  | 32 | 103 | 12.8 | $1.1 \times 10^{-2}$ |
|        | Africans     | 1,121  | 21 | 77  | 19.7 | $1.8 \times 10^{-2}$ |
| Rest   | All human    | 15,435 | 53 | 516 | 43.9 | $2.8 \times 10^{-3}$ |
|        | Non-Africans | 15,437 | 32 | 255 | 25.7 | $1.7 \times 10^{-3}$ |
|        | Africans     | 15,448 | 21 | 290 | 57.0 | $3.7 \times 10^{-3}$ |

The second, the maternal inheritance property was used for person identification in maternal lineage in forensic science. Theoretically, mtDNA inherited without gene segregation and recombination (Giles, Blanc, Cann, & Wallace, 1980) which resulted in constantly sharing their mother's mtDNA haplotypes of the offsprings (Cortellini, Verzeletti, Cerri, Cisana, & De Ferrari, 2015). Though the paternal mitochondria could penetrate into oocytes, they were eliminated by the lysosomes in oocyte cytoplasm (Sato & Sato, 2013). The remaining paternal mtDNA transmission was very rare, but it could occur (Schwartz & Vissing, 2002).

The last advantage of mitochondrial DNA was its high copy number in a single cell, 100 to 1,800 copies approximately (Bogenhagen and Clayton, 1974). When there was high copy number of mtDNA, there was more chance to successfully amplify mitochondrial DNA in case of old or damaged samples. However, mitochondria copy number varied from cell types, mostly found in cardiac muscles and liver (Malka, Lombès, & Rojo, 2006).

Another property of mtDNA is independent replication. The DNA replication of mitochondria is independent of nuclear genome. The properties which high mutation rate and high copy number resulting the mix type of mtDNA called heteroplasmy.



In some case, the mixture of mitochondria, heteroplasmy, can be found in cells or tissue. The degree of heteroplasmy was related to the mitochondrial disease such as the Leber's Hereditary Optic Neuropathy (LHON) (Holt, Miller, & Harding, 1989), Alzheimer disease (AD) and Parkinson disease (PD) (Shoffner et al., 1993) and also aging (Linnane, Ozawa, Marzuki, & Tanaka, 1989). However, heteroplasmy transmission is also important tools in forensic sciences like nuclear genome. The case of Tsar Nicolas II is the classic example on heteroplasmy which transmitted in maternal linages (Gill et al., 1994; Ivanov et al., 1996).

### **Human mtDNA applications**

In molecular genealogy, the accumulation of nucleotide mutations since the time of the most recent common ancestor (MRCA) has to be analyzed to reconstruct the tree of human genetic relationship. In order to locate the mutation sites, the position must be compared with reference. The rCRS (Andrew et al., 1999) is the favorable reference sequence. The original sample of rCRS, the HELA cells, belonged to the European person but many studies of the human genealogy suggested that the human origin was African. The Copernican, an ancient *Homo sapiens*, mitochondrial DNA was suggested to use as reference sequence and called the Reconstructed Sapiens Reference Sequence (RSRS) (Behar, et al., 2012). Nevertheless, RSRS is still not favorable reference in forensic science and population study due to plenty of analyzed sequence which already compared with rCRS.

Now the data of mitochondrial DNA are available online in many database such as NCBI (<http://www.ncbi.nlm.nih.gov>), MITOMAP (<http://www.mitomap.org/bin/view.pl/MITOMAP/WebHome>), EMPOP (<http://empop.org/>), mtDB (<http://www.mtodb.igp.uu.se/>) and *mitosearch* (<http://www.mitosearch.org/>). Many of them are included the varieties of phylogenetic analysis tools.

## Y-chromosome polymorphisms

The human Y-chromosome is the second smallest human chromosome with length of approximately 60 million nucleotides. The tips of the Y-chromosome which are called the pseudoautosomal regions, recombine with their sister X-chromosome homologous regions. The remainder of the Y-chromosome (approximately 95%) is known as the non-recombining portion of the Y-chromosome, or NRY. The NRY remains the same from father to son unless a mutation occurs.

Y-chromosome DNA testing is important for a number of different applications of human genetics including forensic evidence examination, paternity testing, historical investigations, studying human migration patterns throughout history, and genealogical research. The value of the Y-chromosome in forensic DNA testing is that it is found only in males. Since a vast majority of crimes where DNA evidence is helpful, particularly sexual assault, involve males as the perpetrators, DNA tests are designed to only examine the male portion and it can be valuable.

The Y-chromosome has become a popular tool for tracing historical human migration patterns through male lineages. Anthropological, historical, and genealogical questions can be answered through Y-chromosome results. Random mutations are the only mechanisms for variation over time between paternally related males.

While, the mitochondrial polymorphisms are the nucleotide substitution, the polymorphisms on Y-chromosome are the repeated sequence or short tandem repeats (Y-STR) (Figure 4). The Y-STRs are distributed along the chromosomes. Otherwise, they are some nucleotide substitution sites on Y-chromosomes. Since Y-STRs change more rapidly (mutation rate approximately 1 in  $10^3$ ) compared to Y-SNPs (mutation rate approximately 1 in  $10^9$ ), Y-STR results exhibit more variability and thus have greater use in forensic applications. Typically, Y-STRs are described as defining haplotypes while Y-SNPs define haplogroups. Y-SNPs can be useful in DNA ancestry studies.

|      | SNP  | STR |           |
|------|--|-----|-----------|
| Man1 | ATCAG...CGCG/ <u>GTA</u> / <u>GTA</u> /GTGACTACGT  |     | 2 repeats |
| Man2 | AT <u>G</u> AG...CGCG/ <u>GTA</u> / <u>GTA</u> / <u>GTA</u> /GTGACTACGT                              |     | 3 repeats |
| Man3 | AT <u>G</u> AG...CGCG/ <u>GTA</u> / <u>GTA</u> / <u>GTA</u> / <u>GTA</u> / <u>GTA</u> /GTAGTGACTACGT |     | 5 repeats |

**Figure 4 Single nucleotide polymorphisms and short tandem repeats**

Although more than 400 Y-STR loci have been mapped along the human Y-chromosome, only a small set of core loci is routinely analyzed. In 1997, the European forensic community settled on a core set of Y-STR markers or “minimal haplotype” that include DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393 and DYS385 a/b (Figure 5). Most Y-chromosome data to date has been generated with these loci. In 2003, the U.S. Scientific Working Group on DNA analysis Methods (SWGDM) recommended a core set of Y-STRs that includes the nine markers in the minimal haplotype plus DYS438 and DYS439 (Willuweit and Roewer, 2007). These loci are available in commercial Y-STR kits. Although new markers will be added to databases as their value is demonstrated and they become part of commercially available kits, these 11 established markers are likely to continue to be important in future Y-STR research. The characteristics of commonly used Y-chromosome STR loci show in table 2.

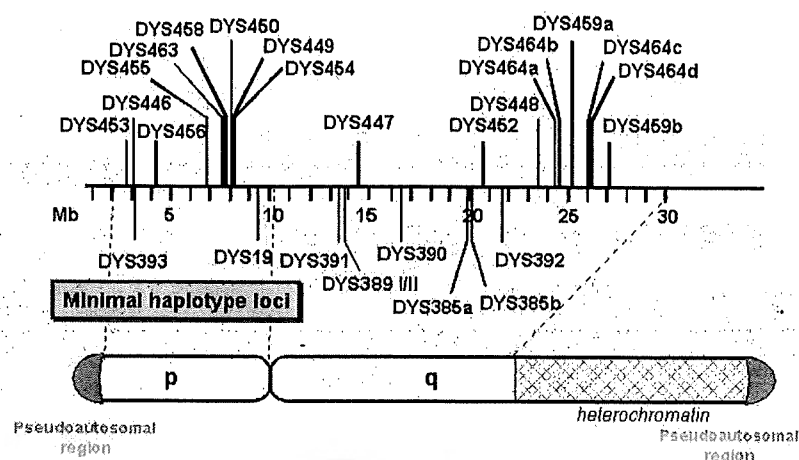


Figure 5 Relative position of Y-STR (Butler & Reeder, 2014)

### Y-STR haplotype databases

A number of online Y-STR databases exist (Table 3). The forensic databases contain collections of anonymous individuals and can be used to estimate the frequency of specified Y-STR haplotypes. The genetic genealogy databases, such as Ysearch and Ybase, contain Y-STR haplotype information gathered by genetic genealogy companies with different sets of loci from males trying to make genealogical connections. Thus, the haplotypes in these genealogy databases are associated with specific individuals and family names.

The largest and most widely used forensic and general population genetics Y-STR database, known as the Y-STR Haplotype Reference Database (YHRD) contains results from more than 65,000 samples representing greater than 500 different groups of sample submissions from various populations around the world. Several genetic genealogy Y-STR haplotype database are also available online. These databases are typically not only used for Y-STR forensic haplotype frequency estimates, but could be helpful in trying to associate a family surname with a particular haplotype if this information was desired in an investigation. These genetic genealogy database contain information from the minimal haplotype loci, a subset of the minimal haplotype loci, or additional Y-STR, and therefore it can always be searched across all loci of interest.

**Table 2 Characteristics of commonly used Y-chromosome STR loci (Butler, 2009, p.370)**

| STR marker  | Position (MB) | Repeat Motif | Allele Range | Mutation rate |
|-------------|---------------|--------------|--------------|---------------|
| DYS393      | 3.19          | AGAT         | 8-17         | 0.08%         |
| DYS456      | 4.33          | AGAT         | 13-18        | 0.72%         |
| DYS458      | 7.93          | GAAA         | 14-20        | 1.1%          |
| DYS19       | 10.13         | TAGA         | 10-19        | 0.24%         |
| DYS391      | 12.61         | TCTA         | 6-14         | 0.29%         |
| DYS635      | 12.89         | TSTA         | 17-27        | 0.42%         |
| DYS437      | 12.98         | TATR         | 13-17        | 0.15%         |
| DYS439      | 13.03         | AGAT         | 8-15         | 0.54%         |
| DYS389 I/II | 13.12         | TCTR         | 9-17/24-34   | 0.19%/0.30%   |
| DYS438      | 13.38         | TTTTC        | 6-14         | 0.05%         |
| DYS390      | 15.78         | TCTR         | 17-28        | 0.24%         |
| GATA-H4     | 17.25         | TAGA         | 8-13         | 0.25%         |
| DYS385 a/b  | 19.26         | GAAA         | 7-28         | 0.22%         |
| DYS392      | 21.04         | TAT          | 6-20         | 0.06%         |
| DYS448      | 22.78         | AGAGAT       | 17-24        | 0.18%         |

For applicable data, a large amount of data have to be collected from each separated population. For example, in Koreans (Park, et al., 2005), Singapore (Yong, Lee and Yap, 2006), Japanese (Hashiyada, et al., 2008), Chinese Han (Hu, 2006) and Bangladeshi (Alam, et al., 2010). Unfortunately, in 879 population data of YHDB database, it is only one Thai population with 757 samples collected from central part Thailand (Siriboonpiputtana, et al., 2010).

**Table 3 Summary of available online Y-STR databases**

| Database                                  | Number of samples (population groups)          | Web site  |
|---|--|---|
| Y-STR Haplotype Reference Database (YHRD) | 65,165 (517 groups)                            | <a href="http://www.YHRD.org">http://www.YHRD.org</a>                     |
| US Y-STR Database (US Y-STR)              | 13,906 (5 groups)                              | <a href="http://www.usystrdatabase.org">http://www.usystrdatabase.org</a> |
| Ysearch                                   | 65,869 (genealogists)                          | <a href="http://www.ysearch.org">http://www.ysearch.org</a>               |
| Ybase                                     | 13,830 (genealogists)                          | <a href="http://www.ybase.org">http://www.ybase.org</a>                   |
| Sorenson Genealogy (SMGF)                 | Molecular Foundation<br>>30,000 (genealogists) | <a href="http://www.smgf.org">http://www.smgf.org</a>                     |

### Interpretation of Y-STR results

Since the Y-chromosome is passed from father to son with unchanged, the observation of match with Y-STR does not carry some power of discrimination as an autosomal STR. The lack of recombination between Y-chromosome markers means that allele of each Y-STR locus have to be combine into a haplotype for searching available databases as well as estimating the rarity of a particular haplotype. The statistical data of Y-STR are derived from population data and done by counting method, where the number of haplotype in the database is used. The size of the database used for the counting method makes a difference when trying to estimate the rarity of Y-STR profile. The larger number of unrelated individuals in the database, the better the statistics will be for a frequency estimation.

### Statistical test

To estimate the genetic diversity of population, several statistical value were used to test. The data of mitochondrial DNA can be analyze under the form of haplotype and Y-STR can be analyze under the form of haplogroup. The genetic structure can be estimate both within and between populations.

Within population, the parameter measurement are below:

### 1. Number of polymorphic sites

The number of usable loci show more than one allele per locus.

### 2. Mean number of pairwise differences ( $\pi$ )

Mean number of differences between all pairs of haplotypes in the sample will be determined It is given by

$$\pi = \frac{n}{n-1} \sum_{i=1}^k \sum_{j=1}^k p_i p_j d_{ij}$$

where  $d_{ij}$  is an estimate of the number of mutations having occurred since the divergence of haplotypes  $i$  and  $j$ ,  $k$  is the number of haplotypes,  $p_i$  is the frequency of haplotype  $i$ , and  $n$  is the sample size. The total variance (over the stochastic and the sampling process), assuming no recombination between sites and selective neutrality, is obtained as

$$V(\hat{\pi}) = \frac{3n(n+1)\pi + 2(n^2 + n + 3)\hat{\pi}^2}{11(n^2 - 7n + 6)}$$

(Tajima, 1993)

### 3. Estimation of genetic distances between DNA sequences

Definitions

L : Number of loci

Gamma correction : This correction is proposed when the mutation rates cannot be assumed as uniform for all sites. It seems to be the case of the control region of human mtDNA. In such a case, a Gamma distribution of mutation rates is often assumed. The shape of this distribution (the unevenness of the mutation rates) is mainly controlled by a parameter  $a$ , which is the inverse of the coefficient of variation of the mutation rate. The smaller the  $a$  coefficient, the more uneven the mutation rates. A uniform mutation rate corresponds to the case where  $a$  is equal to infinity.

- $n_d$  : Number of observed substitutions between two DNA sequences  
 $n_s$  : Number of observed transitions between two DNA sequences  
 $n_v$  : Number of observed transversions between two DNA sequences  
 $\omega$  : G+C ratio, computed on all the DNA sequences of a given sample

### 3.1 Pairwise difference

Outputs the number of loci for which two haplotypes are different

$$\hat{d} = n_d$$

$$V(\hat{d}) = \hat{d}(L - \hat{d})/L$$

### 3.2 Percentage difference

Outputs the percentage of loci for which two haplotypes are different

$$\hat{d} = n_d/L$$

$$V(\hat{d}) = \hat{d}(1 - \hat{d})/L$$

## 4. Estimation of distances between Y-STR haplotypes

### 4.1 No. of different alleles

The simply count the number of different alleles between two haplotypes.

$$\hat{d}_{xy} = \sum_{i=1}^L \delta_{xy}(i)$$

Where  $\delta_{xy}(i)$  is the Kronecker function, equal to 1 if the alleles of the  $i$ -th locus are identical for both haplotypes, and equal to 0 otherwise. When estimating genetic structure indices, this choice amounts at estimating weighted  $F_{ST}$  statistics over all loci.

### 4.2 Sum of squared size difference

Counts the sum of the squared number of repeat difference between two haplotypes.

$$\hat{d}_{xy} = \sum_{i=1}^L (a_{xi} - a_{yi})^2$$

Where  $a_{xi}$  is the number of repeats of the microsatellite for the  $i$ -th locus.

When estimating genetic structure indices, this choice amounts at estimating an analog of Slatkin's  $R_{ST}$  (1995).



## 5. Neutrality tests

### 5.1 Tajima's test of selective neutrality

Tajima's (1989a) test is based on the infinite-site model without recombination, appropriate for short DNA sequences or RFLP haplotypes. It compares two estimators of the mutation parameter theta ( $\theta = 2Mu$ , with  $M=2N$  in diploid populations or  $M=N$  in haploid populations of effective size  $N$ ). The test statistic  $D$  is then defined as

$$D = \frac{\hat{\theta}_{\pi} - \hat{\theta}_s}{\sqrt{\text{Var}(\hat{\theta}_{\pi} - \hat{\theta}_s)}}$$

where  $\hat{\theta}_{\pi} = \hat{\pi}$  and  $\hat{\theta}_s = S / \sum_{i=1}^{n-1} (\frac{1}{i})$ , and  $S$  is the number of segregating sites in the sample. The limits of confidence intervals around  $D$  may be found in Table 2 of Tajima's paper for different sample sizes.

The significance of the  $D$  statistic is tested by generating random samples under the hypothesis of selective neutrality and population equilibrium. The  $P$  value of the  $D$  statistic is then obtained as the proportion of random  $FS$  statistics less or equal to the observation. We also provide a parametric approximation of the  $P$ -value assuming a beta-distribution limited by minimum and maximum possible  $D$  values. Note that significant  $D$  values can be due to factors other than selective effects, like population expansion, bottleneck, or heterogeneity of mutation rates.

### 5.2 Fu's $FS$ test of selective neutrality

Like Tajima's test, Fu's test (Fu, 1997) is based on the infinite-site model without recombination, and thus appropriate for short DNA sequences or RFLP haplotypes. The principle of the test is very similar to that of Chakraborty described above. Here, we evaluate the probability of observing a random neutral sample with a number of alleles similar or smaller than the observed value given the observed number of pairwise differences, taken as an estimator of  $\theta$ . In more details, Fu first calls this probability  $S' = \Pr(K \geq k_{obs} : \theta = \hat{\theta}_{\pi})$  and defines the  $F_S$  statistic as the logit of

$$F_S = \ln\left(\frac{S'}{1 - S'}\right)$$

Fu (1997) has noticed that the  $FS$  statistic was very sensitive to population demographic expansion, which generally lead to large negative  $FS$  values.

The significance of the  $FS$  statistic is tested by generating random samples under the hypothesis of selective neutrality and population equilibrium, using a coalescent simulation algorithm adapted from Hudson (1990, cited by Excoffier, Laval and Schneider, 2005). The P-value of the  $FS$  statistic is then obtained as the proportion of random  $FS$  statistics less or equal to the observation. Using simulations, Fu noticed that the 2% percentile of the distribution corresponded to the 5% cutoff value (i.e. the critical value of the test at the 5% significance level). We indeed confirmed this behavior by our own simulations. Even though this property is not fully understood, it means that a  $FS$  statistic should be considered as significant at the 5% level, if its P-value is below 0.02, and not below 0.05.

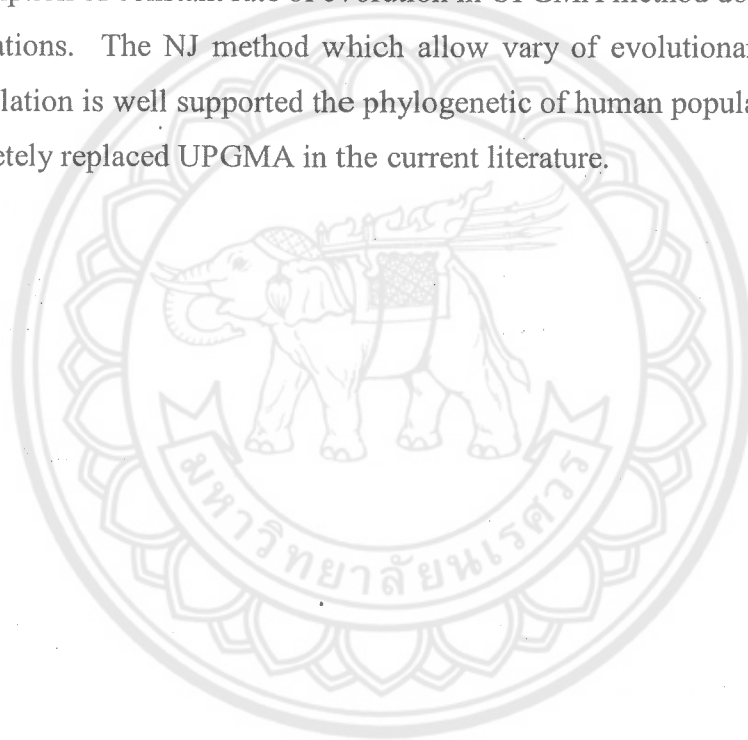
#### **6. The population pairwise genetic distance**

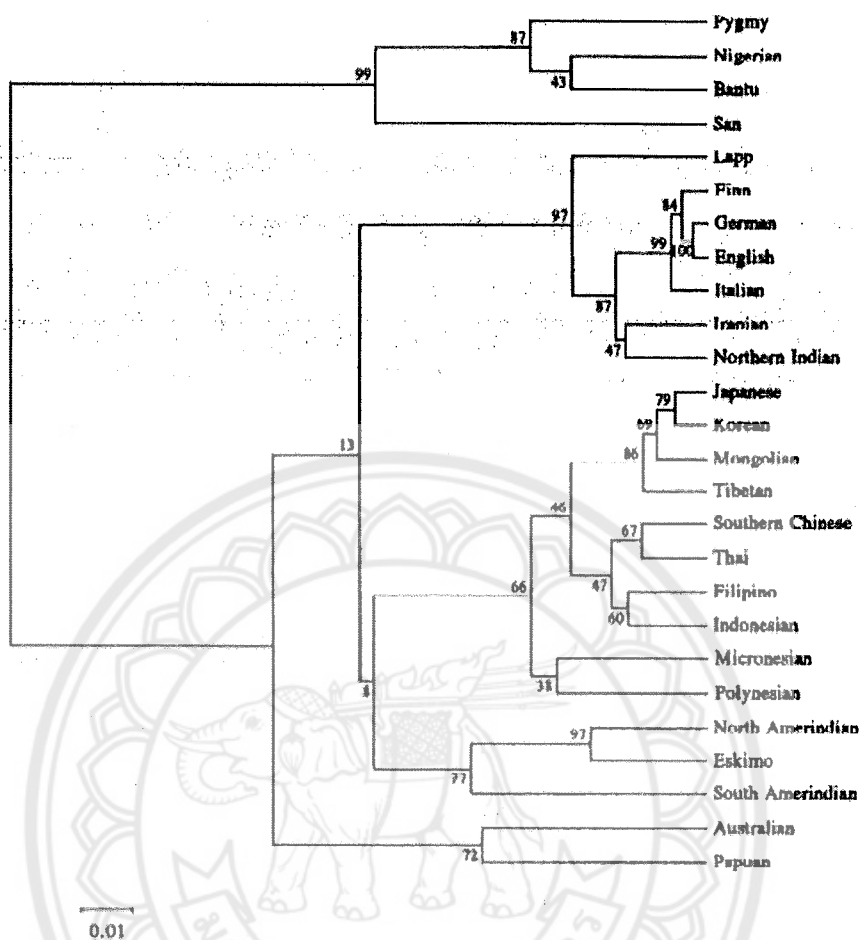
The pairwise  $F_{ST}$  can be used as short-term genetic distances between populations, with the application of a slight transformation to linearize the distance with population divergence time. The pairwise  $F_{ST}$  values are given in the form of a matrix. The null distribution of pairwise  $F_{ST}$  values under the hypothesis of no difference between the populations is obtained by permuting haplotypes between populations. The P-value of the test is the proportion of permutations leading to a  $F_{ST}$  value larger or equal to the observed one. The P-values are also given in matrix form.

#### **Population genetic study in global scale**

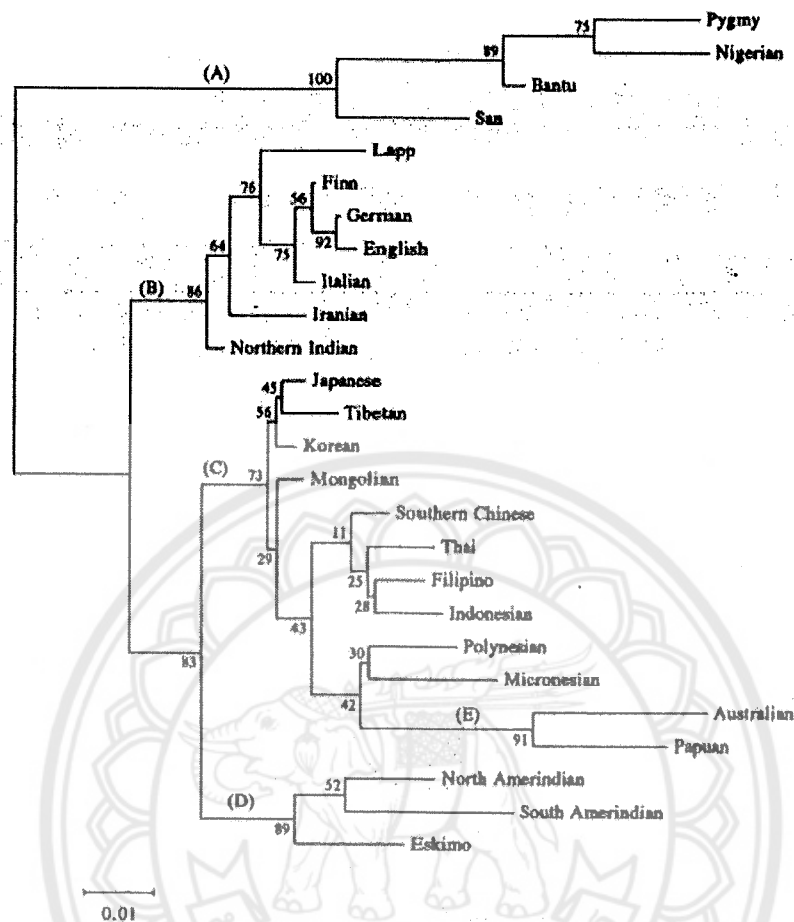
The phylogenetic or the evolutionary relationships of human populations can be established when a large number of population data are examined. By using autosomal polymorphic loci, the major clade of human population separates African from non-African. After that the North Eurasian supercluster (Caucasians, Northeast Asians, and Amerindians) it split from the Southeast Asian supercluster (Southeast Asians, Australians, Papua New Guineans, and Pacific Islanders) (Cavalli-Sforza, et al., 1988). However, two phylogenetic study methods, UPGMA and NJ, presented two different figure of genetic relationship between populations (figure 6 and 7) (Nei and

Roychoudhury, 1993). The UPGMA tree, the second split occurs between the Northeurasian supercluster and the Southeast Asian supercluster, whereas in this tree it occurs between Caucasians and non-Caucasians. The third level of splitting is between the Caucasian cluster and the remainder, whereas the fourth level is between the American cluster (Amerindians and Eskimos) and the cluster of Asians and Oceanians. The splitting pattern of the former tree was poorly supported by bootstrap tests (arrows in figure 6). Moreover, in the human history, the genetic drift in mode of founder effect gives a dramatic change in population size and genetic diversity between populations then the assumption of constant rate of evolution in UPGMA method does not hold for human populations. The NJ method which allow vary of evolutionary rate among different population is well supported the phylogenetic of human populations. NJ has almost completely replaced UPGMA in the current literature.





**Figure 6** UPGMA tree for 26 representation human populations. Data calculated from genetic distance ( $D_A$ ) of 29 autosomal polymorphic loci. The scale of branch is in units of  $D_A/2$ . The bootstrap probability are shown on interior brunch.(Nei & Roychoudhury, 1993)



**Figure 7** NJ tree for 26 representation human populations. Data calculated from genetic distance ( $D_A$ ) of 29 autosomal polymorphic loci. The scale of branch is in units of  $D_A/2$ . The bootstrap probability are shown on interior branch. Major groups of human populations are Africans (A), Caucasians (B), Greater Asians (C), Amerindians (D) and Australopapuan (E). (Nei & Roychoudhury, 1993)

In present day, a large amount of mitochondrial DNA haplotypes and Y-STR haplogroups data are multiply in global scale. Numerous population were reported and the bioinformatics software, both on-line and stand-alone, were used to data analyses. They were all shown not only the genetic relationship between populations but the populations in each continent sharing some genetic characters. The data were simplified by McDonald (2005) (figure 8-11).

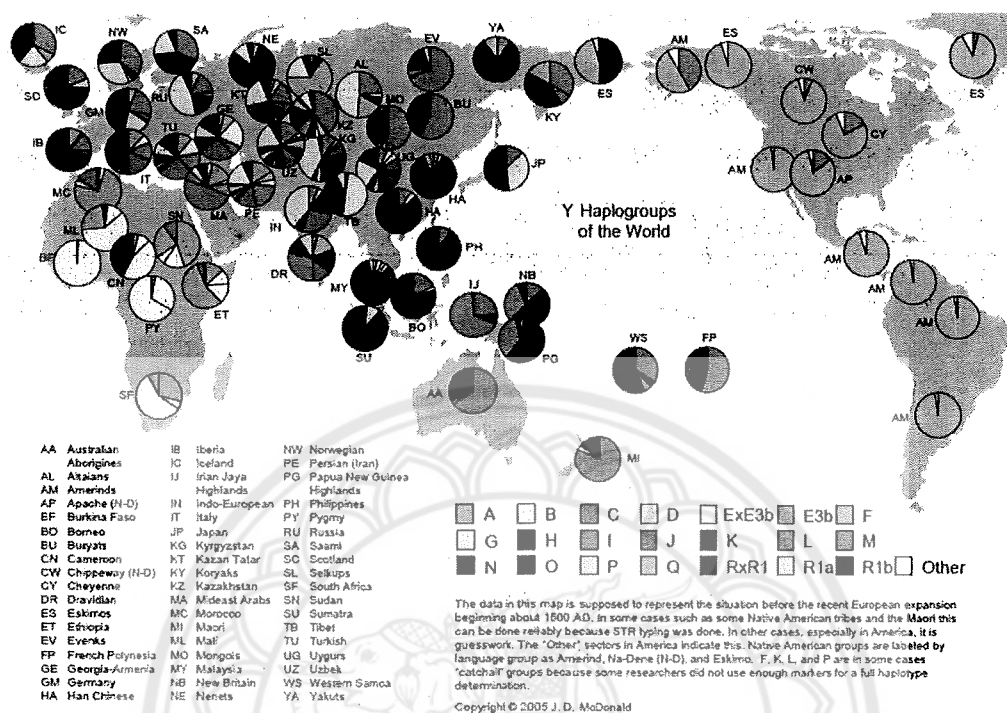


Figure 8 Y-haplogroups of the world (McDonald, 2005).

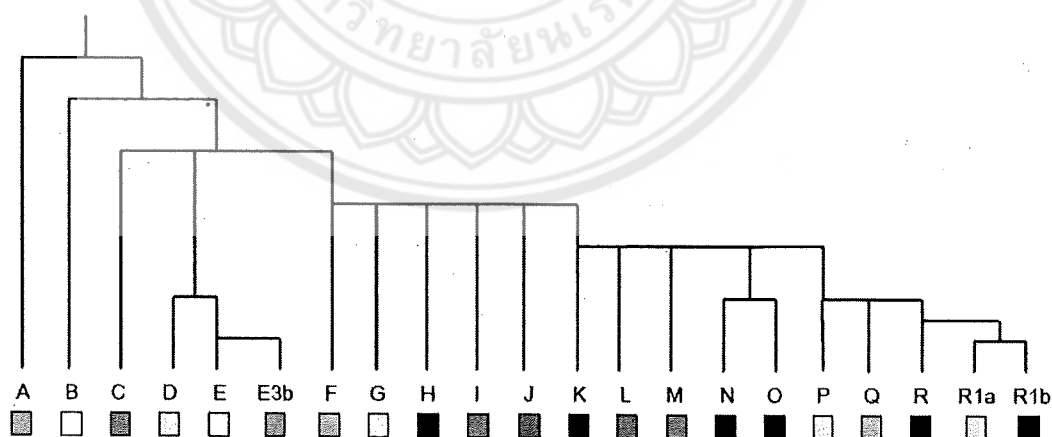
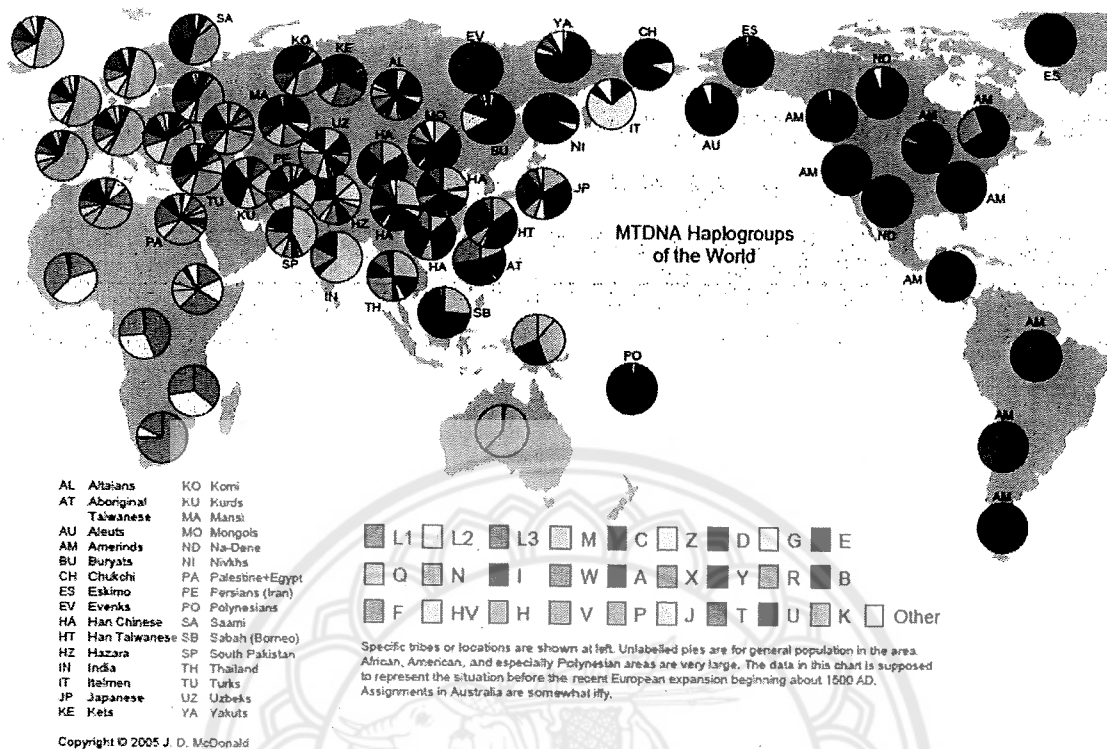
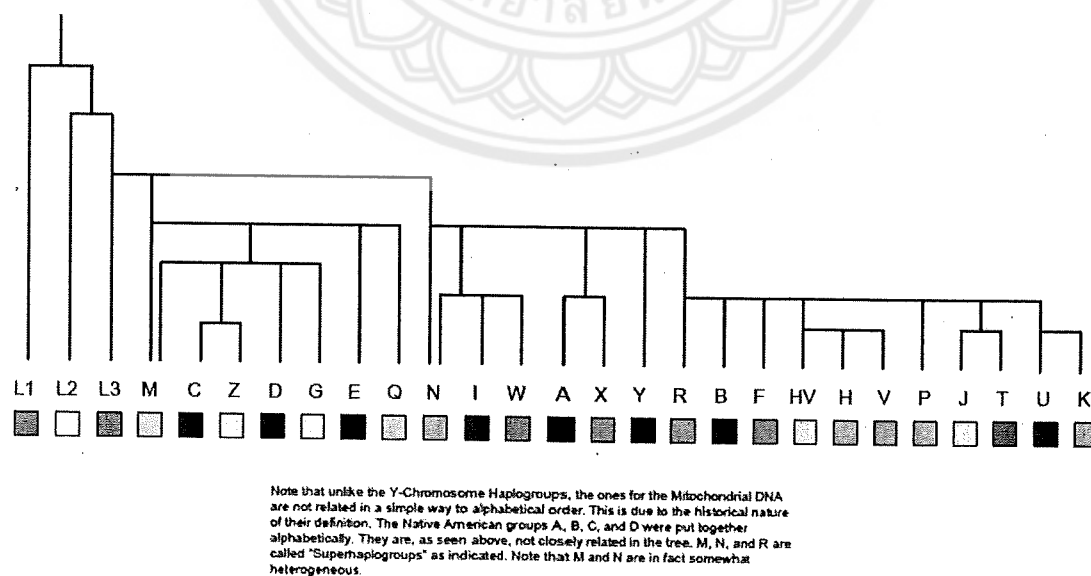


Figure 9 The simplified tree of Y-haplogroups (McDonald, 2005).



**Figure 10** The mitochondrial haplogroups of the world (McDonald, 2005)

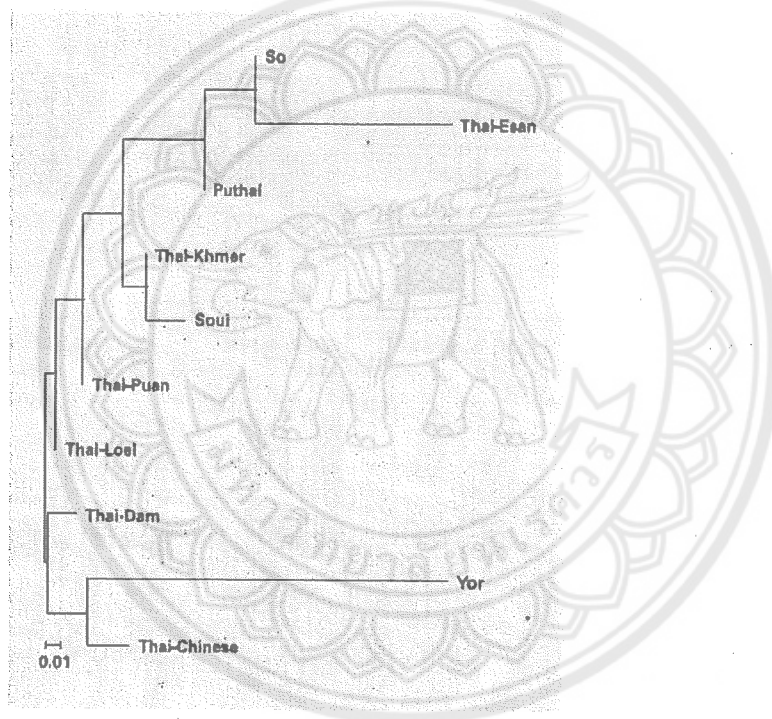


**Figure 11** The simplified tree of mitochondrial haplogroups (McDonald, 2005).

2 Q14  
603  
.M5  
6655  
2560  
1047485

## Population genetics studies in Thailand

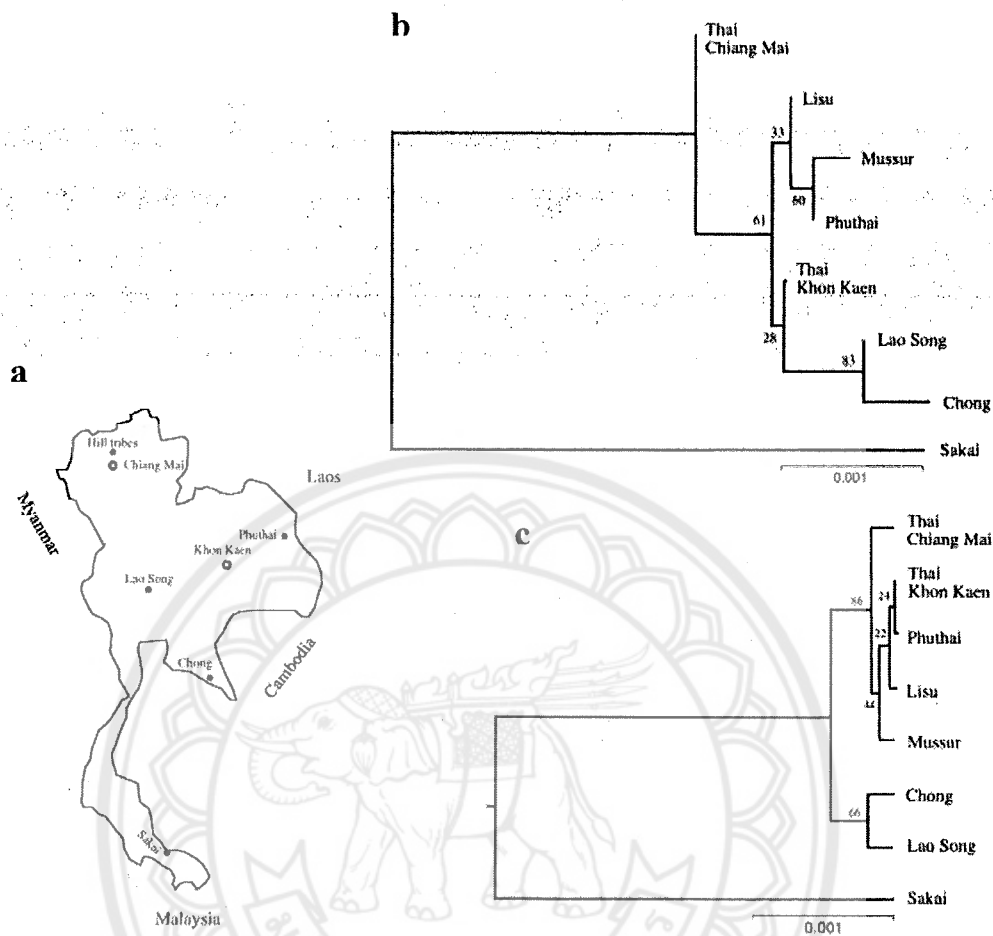
Phylogenetic studied in Thailand began by using the autosomal marker such as  $\beta^E$ -globin gene (Fucharoen, et al., 2001a). Ten minority groups in the Northeastern part of Thailand were examined. The haplotype of  $\beta^E$ -globin show the closely relationship between Soui-Thai Khmer and between So-Thai Esan and Phuthai present in separated clusters. This report was clarified the origin of hemoglobin-E hypothesis. In addition, the NJ tree of this report show the separated cluster between Thai-Chinese populations and Tai-kadia populations which are Thai Dam, Thai Loei and Thai Puan (figure 12).



**Figure 12 NJ tree for ten population's groups studies by  $\beta^E$ -globin haplotype analysis base on  $d_A$  (Fucharoen, et al., 2001a).**

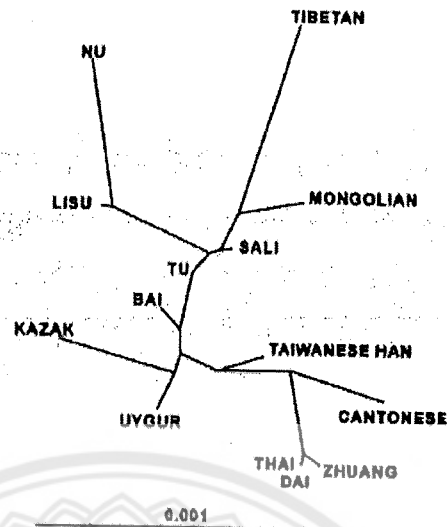
At the same time, the HVR-I of mitochondrial DNA was studied in some native Thai. The phylogenetic study show the closely relationship between Thai, Phuthai and the hill tribes which inhabit in Chiang Mai. It supported the hypothesis which Thai population migrated from the Southern China. However, Thai Khonkaen was grouped with Lao Song and Chong for NJ tree though this clade was support poorly (Figure 13). Again, the NJ tree and UPGMA tree draw a different relationship pattern. According to these results, the Sakai was characterize as basal lineage (Fucharoen, et al., 2001b).





**Figure 13** The relationship of population based on  $d_A$  in eight Thai populations. (a) Geographic distribution of eight Thai populations. (b) NJ tree and (c). UPGMA tree show the relationship of population based on  $d_A$  (Fucharoen, G. et al., 2001b)

Several studies shown the relationship between populations in Southeast Asian. In 2002, the HVR-I of mitochondrial DNA was studied in 13 Chinese population and 1 Northern Thai population. The unrooted NJ tree showed the close affinity between Thai, Dai and Zhuang. They were separated from other Chinese populations (Figure 14) (Yao et al., 2002). This result support the hypothesis that Northern Thai population migrated from Southern part of China.



**Figure 14** Unrooted NJ tree of the 14 populations based on the net genetic distances (dA)(Yao et al., 2002).

In 2009, the mitochondrial DNA haplotypes from Northern part (Chiang Mai) were characterized and published in EMPOP database. The composition and genetic structure were significantly different from other Southeast Asian populations (Zimmermann et al., 2009). At first, his data was aimed to use in forensic science applications, after that these data were chosen to compare with 16 East Asian groups by the linguistic respect: Tai-Kadai, the Austro-Asiatic (including the Mon Khmer), the Sino-Tibetan (including Tibeto-Burmans and Han), the Hmong-Mien (Miao-Yao), the Austronesian and the Altaic (Boodner, et al., 2011). This is the first report which revealed the great mitochondrial variety of population around Thailand. The relationship among neighboring countries of Thailand was influenced on trading and war. This was consequence of migration all over the long times. The small populace can be observed in whole country. The ethnicity was characterized by the linguistic data.

On the other hand, the DNA data together with the linguistic and culture was widely used for the genetic structure study in Southeast Asia. However, the literature of Thai populations are limit in only some areas. For instance, ten populations of Khon Mueang in the Northern part of Thailand, were studied to compare with the ethnic groups; namely Lawa, Mon (Mon-Khmer-speaking groups), Shan, Yuan, Lue, Khuen and Yong (Tai-speaking groups). The Khon Meung showed the mixture of genetic structure in the closely geographic ethnic groups (Kutanan et al., 2011a). Different from the study in Mon-Khmer speaking populations compare to Tai-Kadai speaking, they showed the divergence of genetic relationship among them (Figure 15) (Kutanan, et al., 2011b). Meanwhile the populations in Southeast Asian, especially in Thailand, has a great variety of genetic structure but they frequently show the influence of linguistics which affect to population isolation, especially in Mon-Khmer speaking (Xu et al., 2010; Kampuansai et al., 2012). Moreover, the genetic data can be proved the ethnicity of population. For example, there are Lau living in Nan province which their ethnic history are unclear. The mitochondrial DNA, showed the different genetic structure between the Lua from Nan province and other areas whereas the genetic of them was similar to H'tin who migrated from Laos (Kutanan & Kangwanpong, 2010)



## CHAPTER III

### RESEARCH METHODOLOGY

This chapter presents the methodology of this research including materials, apparatus and methods. The detail of each topic is described below.

#### Hypothesis

This study aimed to determine the genetics structure of populations and their genetic relationship, the hypothesis were set for this study as follow;

Null hypothesis: the genetic structure between individual within population were not difference ( $H_0: A_1 = A_2$ )

Alternative hypothesis: the genetic structure between individual within population were difference ( $H_A: A_1 \neq A_2$ )

Null hypothesis: the genetic structure between population were not difference ( $H_{0(n)}: A_n = B_n$ )

Alternative hypothesis: the genetic structure between population were difference ( $H_{A(n)}: A_n \neq B_n$ )

#### Ethical consideration

The thesis proposal was proposed and approved by the Human Ethic Committee of Naresuan University (COA No. 211/2015). The volunteer will be inform the objectives and benefits of study and effects to them before sign the consent form. Names and personal information will transform to codes and data analyzes by groups to ensure blinding the individual information.

### Populations and sampling

The preliminary questionnaires were sent to 85 sub-districts in 7 provinces of lower northern part of Thailand to locate the ethnic populations according to reports of Ngourungsi and colleagues, 1976; Burusphat, 2000; Jirananthanaporn and colleagues, 2008. The target populations were selected under the criteria which were:

1. The residential habitat of populations were small rural village and separated from others.
2. The ethnicity of populations were identified due to languages and cultures.
3. There were community leaders and medical specialist who cooperated the sample collection process.

Three populations were selected which were:

1. Thai Siam who lived in Phrom Phiram district, Phitsanulok province. (Lat.17.1669996, Long. 100.0945124)
2. Lao Song who lived in Bang Rakam district, Phitsanulok province (Lat.16.6367235, Long.100.1505811)
3. Hmong who lived in Nakhon Thai district, Phitsanulok province (Lat. 16.8124112, Long. 100.9695965) (Figure 16).

The volunteers were explained about the details of research and voluntaries signed the consent form at the collection sites. The data of ethnicity and migration history of each individual were record.

The inclusion criteria were as follow:

1. The individuals had their ethnicity history at least 3 generations.
2. The volunteers were 20 years old or more, healthy, good verbal communication and willing to participate in this study.
3. There were unrelated individuals.

### Sample size

The required sample size in this study based on the numbers of ethnic groups. The volunteers were employed by the consecutive number of sample size per each group within one visit or at least 30 individuals per population.



**Figure 16 Geographic distribution of three ethnic populations; Thai Siam (A) from Phrom Phiram District, Lao Song (B) from Bang Rakam District and Hmong (C) from Nakhon Thai district in Phitsanulok province Thailand.**

### Blood collection

Blood samples were draw from vain by nurse or licenses medication official. Three to five ml. of blood will be stored in EDTA vacutainer tubes.

## DNA extraction

The genomic DNA were extracted using the PureLink™ Genomic DNA Kits (Invitrogen) and quantified by 1% agarose gel electrophoresis. They were stored at 8°C.

## DNA analysis

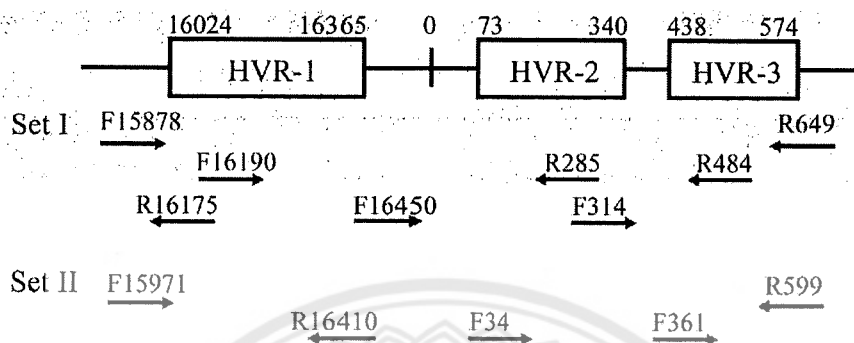
The DNA samples were used to amplify the D-loop control region of mtDNA and Y-STR.

### 1. The mitochondrial DNA study

Genomic DNA was extracted using PureLink™ Genomic DNA Kits (Invitrogen, USA). The entire sequences of D-loop hypervariable region of mtDNA were amplified using primers F15878 and R649 (Brandstätter et al., 2004). The PCR mixture consisting of 2 µl DNA template, 1X OnePCR™ (GeneDirex) and 200 µM each primers were carried out in 50 µl. The PCR were performed by 95°C for 10 minutes and then 36 cycles of 94°C for 1 minute, 56°C for 1 minute, and 72°C for 2 minutes and followed by 72°C for 5 minutes. The PCR product were purified by PureLink® PCR Purification Kit (Invitrogen) and sequencing by BigDye® Terminator v3.1 cycle sequencing kit (1<sup>st</sup> BASE, Singapore). Series of 8 overlapping PCR-amplified primers set-I; F15878, R649, F16190, R16175, F16450, R484, R285 and F314, were used for the D-loop hypervariable sequences determination (Figure 17, Table 4). In case of mutation or polymorphisms at the primers binding site or incomplete sequence, The PCR of primer F15971 and R599 were used instead. The PCR mixture consisting of 2 µl DNA template, 1X OnePCR™ (GeneDirex) and 200 µM each primers were carried out in 50 µl. The PCR was performed by 95°C for 10 minutes and then 36 cycles of 94°C for 30 seconds, 56°C for 30 seconds, and 72°C for 1 minute and followed by 72°C for 5 minutes. The overlapping primers set-II; F15971, R599, R16410, F34, R599 and F361 were used (Irwin et al., 2007). The generated sequences of each sample were assembled using the MEGA7 software (Kumar et al., 2016). To ensure the quality of sequences, the consensus sequences were generated twice and both of consensus sequences were compared to generate the final sequences. The nucleotide positions of



the mtDNA control region were numbered according to the revised Cambridge Reference Sequence (rCRS) (Andrew et al., 1999).



**Figure 17 Position of primers which were used to amplify the entire control region of mtDNA. (modified from Brandstätter et al., 2004)**

### Statistical analysis

The mtDNA control region sequences from position 16024 to 16569 and 1 to 576 were analyzed. MEGA7 (Tamura et al., 2013) was used to identify for DNA variation as follow:

Variable sites (*v*) or polymorphic sites are nucleotide positions which contains 2 or more nucleotides. Some variable sites can be singleton or parsimony-informative.

Singleton (*s*) is a variable site which only one occurrence.

Parsimony-informative sites (*pi*) is a variable site which at least 2 of them occur with a minimum frequency of two.

The percentage of *v*, *pi* and *s* were calculated by number of sites divided by range of nucleotide in each regions.

Base substitution were determine. The percentage of mutations; transitions (ts), transversions (tv), insertions-deletions (in/del) were calculated by number of mutations divided by number of variable sites in each regions.

DnaSP ver.510.1 software (Librado and Rozas, 2009) was used to calculated the numbers of haplotype, haplotype diversity ( $H_d$ ) and its variance (Nei 1987, equations 8.4 and 8.12 but replacing  $2n$  by  $n$ ), nucleotide diversity ( $\pi$ ) (Nei 1987, equations 10.6) and its variance (Nei 1987, equation 10.7). Two neutrality estimators, Fu's  $F_s$  (Fu, 1997) and Tajima's  $D$  (Tajima, 1989) were also calculated by DnaSP.

NJ-tree was reconstructed by MEGA7. The evolutionary distances were computed using the number of differences method (Nei, 2000).

**Table 4 Sequences of primers used for amplification the entire sequence of D-loop hypervariable region of mtDNA**

| Primer names | Sequence 5' to 3'              |
|--------------|--------------------------------|
| F15878       | AAA TGG GCC TGT CCT TGT AG     |
| R649         | TTT GTT TAT GGG GTG ATG TGA    |
| F15971       | TTA ACT CCA CCA TTA GCA CC     |
| R599         | TTG AGG AGG TAA GCT ACA TA     |
| F16190       | CCC CAT GCT TAC AAG CAA GT     |
| F16715       | TGG ATT GGG TTT TTA TGT A      |
| F16450       | GCT CCG GGC CCA TAA CAC TTG    |
| R484         | TGA GAT TAG TAG TAT GGG AG     |
| R285         | GTT ATG ATG TCT GTG TGG AA     |
| F314         | CCG CTT CTG GCC ACA GCA CT     |
| R16410       | GAG GAT GGT GGT CAA GGG A      |
| F34          | GGG AGC TCT CCA TGC ATT TGG TA |
| R921         | ACT TGG GTT AAT CGT GTG ACC    |
| F361         | ACA AAG AAC CCT AAC ACC AGC    |

## 2. The Y-STR study

### Sample collection

281 Blood samples were collected from blood donor volunteers at The Fort Somdet Phranaresuan Maharat Hospital. The blood collection process and personal information of volunteers were preserved followed by the Standard Operation Protocol of the Human Ethic Committee of Naresuan University.

### DNA extraction

Genomic DNA was extracted using MasterPure™ DNA Purification Kit (Epicentre) and quantified with Qubit® dsDNA HS Assay Kit (Invitrogen) using Qubit Fluorometer.

### Y-STR amplification

Five loci of Y-STR were amplified by PCR (Table 6). The condition for PCR reaction are 0.2 µg DNA, 1x PCR buffer, 2.2 mM MgCl<sub>2</sub>, 2 mM dNTP, 0.3 mM primer (each) and 0.13 Unit/µl *Taq* polymerase (Invitrogen®, USA). The thermal conditions for each pair of primers are 94°C for 5 minutes and then 38 cycles of 94°C for 1 minute, 57°C for 1 minute, and 72°C for 1 minute and followed by 72°C for 5 minutes. The genotype of each loci were determine by 8.5% acrylamide gel and followed by silver staining for detection. Characteristics of Y-STR loci were determined (Table 6) (Butler, 2005).

The allelic ladders were constructed by combining all observed allele in each loci. PCR product of each loci were purified using HiYield™ Gel/PCR DNA Fragment Extraction Kit (RBC Bioscience) and sequenced using BigDye® Terminator v3.1 cycle sequencing kit (1<sup>st</sup> Base, BASE Life Sciences Holdings, Singapore).

The allele frequencies and gene diversity were calculated as follow:

$$\text{allele frequency} = \frac{\text{no. of alleles}}{\text{total alleles}}$$

$$\text{gene diversity} = \frac{n(1 - \sum x^2)}{(n - 1)}$$

Where  $n$  = sample size,  $x$  = gene frequencies in population

**Table 5 Sequence of Y-STR primers (Thomas, Bradman, & Flinn, 1999)**

| Loci       | Primer Sequences (5'-to-3')        |
|------------|------------------------------------|
| 1 DYS19-L  | CTA CTG AGT TTC TGT TAT AGT        |
| DYS19-R    | ATG GCA TGT AGT GAG GAC A          |
| 2 DYS390-L | TAT ATT TTA CAC ATT TTT GGG CC     |
| DYS390-R   | TGA CAG TAA AAT GAA CAC ATT GC     |
| 3 DYS391-L | CTA TTC ATT CAA TCA TAC ACC CAT AT |
| DYS391-R   | ACA TAG CCA AAT ATC TCC TGG G      |
| 4 DYS392-L | AAA AGC CAA GAA GGA AAA CAA A      |
| DYS392-R   | CAG TCA AAG TGG AAA GTA GTC TGG    |
| 5 DYS393-L | GTG GTC TTC TAC TTG TGT CAA TAC    |
| DYS393-R   | AAC TCA AGT CCA AAA AAT GAG G      |

**Table 6 Characteristics of 5 Y-STR loci**

|   | Loci   | Repeat motif | Allele Range |
|---|--------|--------------|--------------|
| 1 | DYS19  | TAGA         | 10-19        |
| 2 | DYS390 | TCTR         | 17-28        |
| 3 | DYS391 | TCTA         | 6-14         |
| 4 | DYS392 | TAT          | 6-20         |
| 5 | DYS393 | AGAT         | 8-17         |

## CHAPTER IV

### RESULTS AND DISCUSSION

#### Samples collection

The preliminary questionnaires were sent to 85 sub-districts in 23 districts, 7 provinces. Forty were returned. There were 25 sub-districts which information of ethnic populations were available. Twenty sub-districts are located in Phitsanulok province. Three locations —Phrom Phiram, Bang Rakam and Nakhon Thai— were selected because they are small rural villages and community histories which can present their ethnicity are available. Total 167 samples from unrelated individuals were collected (Table 7). Age of individuals ranges from 17 to 90 years. Sex ratio was approximately 1.6 (F:M) in all populations.

Data from 75 families' data and blood specimens were collected from Phrom Phiram district. The dominant ethnicity is Thai (88%). Fifty-three families of Thai were in lower northern part of Thailand (Phitsanulok, Phichit, and Uttaradit for at least three generations, while eight families (12.12%) moved from other region of Thailand. There were also Lao, Thai-Laos, and Thai-Chinese families inhabited in this location. Some families moved from other region of Thailand such as, Chaiyaphum, Khon Kaen, Lopburi and Ratchaburi. There were three (4.55%) matrilocal and two (3.03%) patrilocal families. This population is Thai or so called Thai-Siam who spoke central-Thai dialect, according to linguistic characterization (Ngourungsi et al., 1976). The specimens which were identified as Lao or others were not used for DNA analysis.

Family history and blood specimen from 80 individuals were collected from Bang Rakam district. Almost of their individuals were Lao Song (75%). Twenty-nine individuals from Lao Song (48.3%) lived in Phitsanulok and Phichit at least last 3 generations, while 17 family (28.33%) moved from central part of Thailand —Nakhon Pathom, Ratchaburi, Phetchaburi, Samut Sakhon, and Supanburi. Nine patrilocal (15%) and five matrilocal (8.33%) families were observed. Other ethnics found in this location

—Thai, Lao Wiang, Lao Puan, Chinese, and others were excluded for DNA analysis. In this population, Lao Song still keep their heritages such as language and weaving. Lao Song have their own language used within families or relatives but also used central-Thai language in normal situations. The basis traditions of Lao Song were believed in nature and spirit. “Thaen” —the nature and ancestors spirit— were paying homage in every household, even the society and living style were changed (Pitiphat & Phunsuwan, 1997). The wedding tradition —Kin Dong was carried hierarchy and relationship within populations. These factor were strong tradition which keep the identity of Lao Song in lower northern part of Thailand (Srising, 1976). The same as Lao Song who lived in others region of Thailand such as Phetchaburi (Wadkeaw, 1978) and Supanburi (Khiawkham, 2003).

Seventy individuals were interviewed and blood specimen were collected from Nakhon Thai district. All of them are Hmong. Sixty families (85.71%) are living in Nakhon Thai district and surrounding area —Chart Trakarn, Khao Kho, Lumkao, and Dan Sai districts, while 9 families (12.86) migrated from others region —Nan, Chiang Rai and Yunan. Only one Hmong-Thai specimen was observed and excluded for DNA analysis. This data indicated that Hmong migrated along with their whole families, neither patrilocal nor matrilocal society. In present day, the development of economics, social, government policies and medias were the factor which changed the life styles of Hmong (Sanyakul, 2004). They communicated and traded with Thai people. Younger person used Thai language. However, the wedding tradition of Hmong were still closed in their own ethnic (Mongkokthawat, 2005).

**Table 7 Characteristic of three ethnic populations in Phitsanulok province**

| Locations                | No.<br>(F:M)  | Avg. Age<br>(SD) | Dominant<br>ethnicity (n, %) | Inhabitant<br>(%) | Immigrant<br>(%) | Matrilocal<br>family (%) | Patrilocal<br>family (%) |
|--------------------------|---------------|------------------|------------------------------|-------------------|------------------|--------------------------|--------------------------|
| Phrom Phiram<br>district | 75<br>(35:40) | 55.5<br>(12.5)   | Thai<br>(66, 88)             | 53<br>(80.30)     | 8 (12.12)        | 3 (4.55)                 | 2 (3.07)                 |
| Bang Rakam<br>district   | 80<br>(53:27) | 60<br>(11.9)     | Lao Song<br>(60, 75)         | 29<br>(48)        | 17 (28.33)       | 5 (8.33)                 | 9 (15)                   |
| Nakhon Thai<br>district  | 70<br>(41:29) | 42.6<br>(17.4)   | Hmong<br>(70, 100)           | 60 (85.71)        | 9 (12.86)        | 1(1.43)                  | -                        |

## DNA analysis

### MtDNA control regions analysis

The DNA of 66 Thai, 60 Lao Song, and 70 Hmong were extracted and quantified. D-loop control region from position 15878 to 649 were amplified and sequenced by set of 8-13 primers. MEGA7 software was used to generate consensus sequences from multiple primer sequences.

One hundred and nine sequences from unrelated individuals were aligned and compared with rCRS reference sequence. The D-loop control region from position 16,024–16,569 and from 1–576 (1130 nucleotides) was analyzed. One hundred and sixty-one segregated sites (14.25%) were found (Table 8). Ninety-six parsimony-informative sites (8.5%) and 61 singletons (5.40%) were noted. The HVR-1 contained the highest number of variable and parsimony-informative sites. One hundred and sixty-two mutational sites were counted, consisting of 137 transitions and 26 transversions. The combination of transitions and transversions were found at 7 sites. This result indicated that transitions were common in the entire control region, but transversions were more frequent in HVR-3 than in other regions. The ratio of the transitions to transversions was approximately 5:1. However, the ratio of the transitions to transversions in HVR-1 was 9:1, which was lower than that reported in a previous report on Thai population (12:1) (Sangthong et al., 2015). This ratio revealed the mutation that is accumulated in isolated populations.

A total of 1227 mutations were observed in this analysis, which contained 1,108 nucleotide substitutions, of which most were transitions (Table 9). The T→C mutation was the most frequent transition (36.01%) and A→G was the second frequencies. Only 10.65% transversions were observed, and most of them were A→C (6.77%) and C→A (2.08%), respectively. Four insertion/deletion (in/del) mutations were found at position 249 (A-deletion) (10.38%), 512 (ACAC-insertion) (0.36%), and 524–525 (CC-deletion) (0.36%).

**Table 8 Distribution of segregating sites and mutations in HVR of mtDNA control regions in the three ethnic populations in Phitsanulok Province (N = 109)**

|                   | HVR-1       | HVR-2      | HVR-3      | between HVR          | Total       |
|-------------------|-------------|------------|------------|----------------------|-------------|
| Position          | 16024-16365 | 73-340     | 438-576    | 16366-71,<br>341-437 |             |
| Range (bp)        | 342         | 268        | 137        | 373                  | 1130        |
| Seg. sites (%)    | 76 (22.22)  | 42 (15.67) | 22 (16.06) | 21 (5.63)            | 161 (14.25) |
| Par-Info. (%)     | 59 (17.25)  | 20 (7.46)  | 10 (7.30)  | 7 (1.88)             | 96 (8.50)   |
| Singletons (%)    | 18 (5.26)   | 18 (6.72)  | 11 (8.03)  | 14 (3.75)            | 61 (5.40)   |
| Transitions (%)   | 72 (94.74)  | 35 (83.33) | 10 (45.45) | 19 (90.48)           | 137 (85.09) |
| Transversions (%) | 8 (10.53)   | 6 (14.29)  | 10 (45.45) | 2 (9.52)             | 26 (16.15)  |
| in/del (%)        | 0           | 1 (2.38)   | 3 (13.64)  | 0                    | 4 (2.48)    |

The poly-C (or C-stretch) was found around positions 16,184–16,193 and 303–309. However, they were reported as common length heteroplasmy sites in several populations (Brandstätter et al., 2004; Ramos et al., 2013). The length variation heteroplasmy interrupted the signal of chromatogram sequences (Bendall & Sykes, 1995) results an incomplete sequencing. In this report, several sequences were exclude because of poly-C at position 16181 to 16198. In the other hand, sequences which contained T16189 interruption in poly-C fragment resulted in exactly numbers of C-stretch (Figure 18).

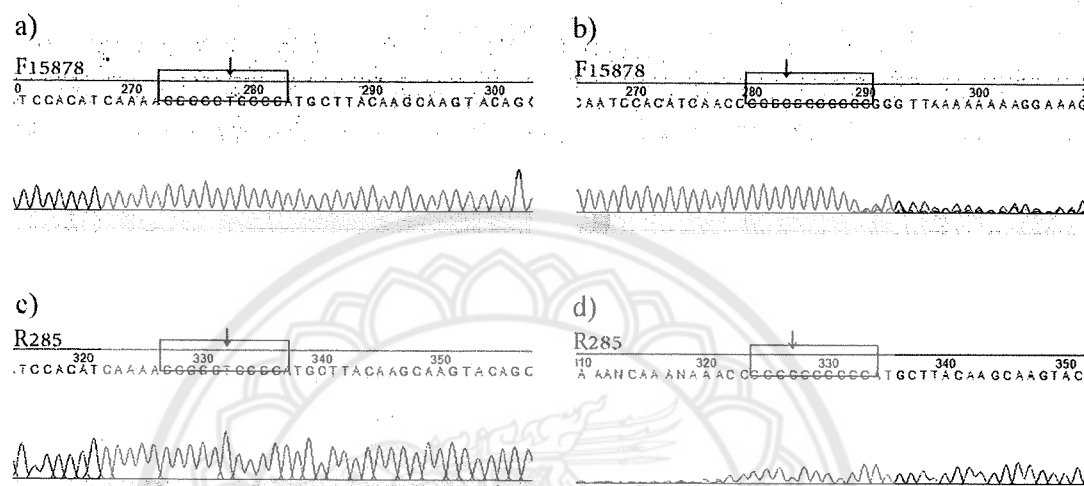


**Table 9 Observed mutations in 1,130 nucleotides of mtDNA control region in studied populations (n=109)**

| Type of mutations  |       |  | %     |
|--------------------|-------|--|-------|
| Base substitutions | 1,108 |  |       |
| Transitions        | 990   |  | 89.35 |
| A→G                | 275   |  | 24.82 |
| G→A                | 98    |  | 8.84  |
| T→C                | 399   |  | 36.01 |
| C→T                | 218   |  | 19.68 |
| Transversions      | 118   |  | 10.65 |
| A→C                | 75    |  | 6.77  |
| A→T                | 1     |  | 0.09  |
| G→C                | 2     |  | 0.18  |
| G→T                | 0     |  | 0.00  |
| T→A                | 1     |  | 0.09  |
| T→G                | 0     |  | 0.00  |
| C→A                | 23    |  | 2.08  |
| C→G                | 16    |  | 1.44  |
| Deletions          | 115   |  | 10.38 |
| Insertions         | 4     |  | 0.36  |
| Total              | 1,227 |  |       |

In this report, the poly-C positions were then excluded during haplotyping and statistical analysis. Point heteroplasmy was found in only one sequence at position 311(C→Y) in the Lao Song population. However, 2 individuals in Thai-Siam population were characterized with position 311 heteroplasmy, but they were not included in this result because of low quality of DNA sequences due to length heteroplasmy at position 309. There were the heteroplasmy hotspots, mostly in HVR-1 (Irwin et al., 2009), but not at position 309. The ACAC-insertion at position 512 was

also reported as length heteroplasmy (Ramos et al., 2013). In this study, there were 4 individuals with ACAC-insertion, but no heteroplasmy was noted.



**Figure 18** Chromatograms of mtDNA control region from 2 sequences (a and c) and (b and d) amplified by primers F15878 and R285. Boxes covered the nucleotide position 16181 to 16189. The arrow indicated nucleotide position 16189.

The frequency distribution of mutation at each variation site showed a similar pattern in all populations (Figure 19). Some positions, particularly HVR-1, showed higher frequency of mutation ( $>0.100$ ) in all the three ethnic populations (Supplementary Table S1). The positions at which mutational frequency was over 0.3 were 16,223; 489; 523; 524; and 525. The latter four positions were in HVR-3. The distinctively high mutational frequency positions were 73 (A→G) and 263 (A→G) with frequency 1.00 and 0.991, respectively. A previous study reported that position 263A of rCRS was a rare polymorphism (Andrew et al., 1999), but the data for position 73 is still lacking.

**Figure 19** Frequency distribution of mutation at variation sites in mtDNA control regions in (A) Hmong, (B) Lao Song, (C) Thai-Siam, and (D) overall populations. The Y-axis represented mutation frequencies, while the X-axis represented the nucleotide position as compared to rCRS.

### Haplotype analysis

Within 3 ethnic populations, 86 haplotypes were found, composed of 73 unique haplotypes (Table 10, Figure 20). Thai-Siam showed the highest percentage of unique haplotype (85.19%). No shared haplotype was found between populations. It might indicate that each populations were different in maternal genetic structure. In Hmong population, 9 shared haplotypes were found with high frequencies (Table 11). The obviously high-frequency haplotypes in Hmong is Haplotye17 (0.094) and 34 (0.075). This result indicated the close relationship of maternal lineage in Hmong population. There were 2 shared haplotypes found in Lao Song and Thai-Siam. The highest frequency haplotype in overall populations was Haplotype 37, found in Lao Song population. This haplotype contained ACAC-insertion at position 512 in HVR-3. All 4 individuals who had haplotype 37 had maternal migration history from Nakhon Prathom province, the central part of Thailand. This insertion was reported as heteroplasmy, as we mentioned before, there were several reports suggested that the number of heteroplasmy might be eliminated from populations by genetic drift. However, this result suggested that the ACAC-insertion might be unique pattern found in Lao Song population and might be capable to identify Lao Song population.

**Table 10 Numbers of haplotypes, as well as unique and shared haplotypes, observed in the three ethnic populations**

|           | Haplotypes | Unique haplotypes (%) | Shared haplotypes (%) | n   |
|-----------|------------|-----------------------|-----------------------|-----|
| Hmong     | 36         | 27 (50.94)            | 9 (16.98)             | 53  |
| Lao Song  | 25         | 23 (79.31)            | 2 (6.90)              | 29  |
| Thai-Siam | 25         | 23 (85.19)            | 2 (7.41)              | 27  |
| Total     | 86         | 73 (66.97)            | 13 (11.93)            | 109 |

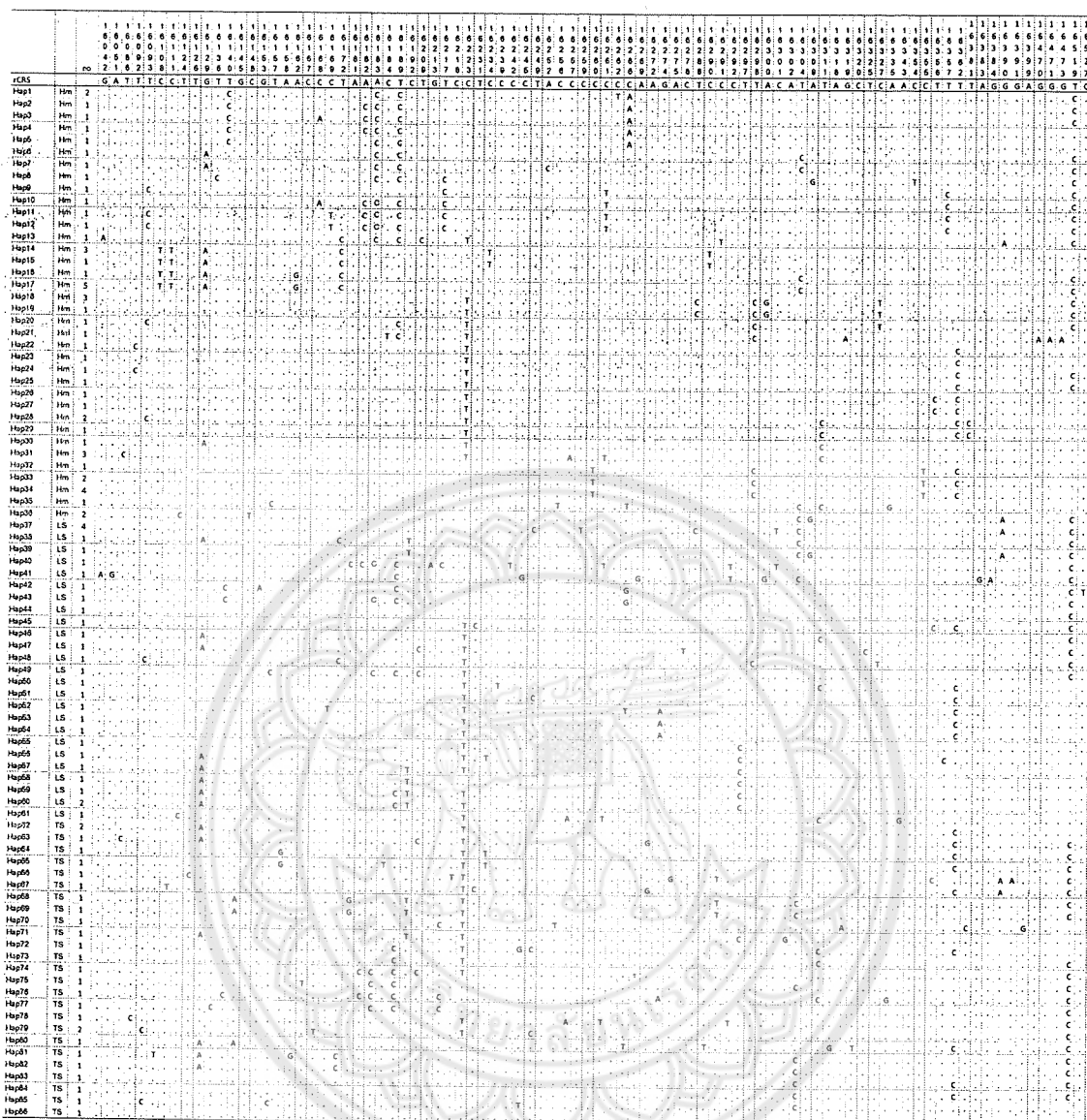


Figure 20.1 Haplotype distribution in three ethnic populations

**Figure 20.2 Haplotype distribution in three ethnic populations**

**Table 11** Frequencies of haplotypes of mtDNA control region in the three ethnic populations. Number of samples of Hmong, Lao Song, and Thai-Siam populations were 53, 29, and 27, respectively

| Shared haplotypes | Ethnicity | Numbers | Frequencies |
|-------------------|-----------|---------|-------------|
| Haplotype 1       | Hmong     | 2       | 0.038       |
| Haplotype 14      | Hmong     | 3       | 0.057       |
| Haplotype 17      | Hmong     | 5       | 0.094       |
| Haplotype 18      | Hmong     | 3       | 0.057       |
| Haplotype 28      | Hmong     | 2       | 0.038       |
| Haplotype 31      | Hmong     | 3       | 0.057       |
| Haplotype 33      | Hmong     | 2       | 0.038       |
| Haplotype 34      | Hmong     | 4       | 0.075       |
| Haplotype 36      | Hmong     | 2       | 0.038       |
| Haplotype 37      | Lao Song  | 4       | 0.138       |
| Haplotype 60      | Lao Song  | 2       | 0.069       |
| Haplotype 62      | Thai-Siam | 2       | 0.074       |
| Haplotype 79      | Thai-Siam | 2       | 0.074       |

The diversity parameters of mtDNA control region haplotype in the three ethnic populations are presented in Table 12. The insertion and deletion sites were ignored (Tillmar et al., 2010; Jankova-Ajanovska et al., 2014). There were 1119 positions in the final dataset. The average number of nucleotide differences ( $k$ ) was approximately 13. The Thai-Siam population showed the highest haplotype diversity ( $H_d$ ). The nucleotide diversity ( $\pi$ ) varied from 0.01130 to 0.01174.

Several lines of evidence indicated decrease in genetic diversity in migrated populations. A previous report on the Hmong population in China, the original habitat of Hmong, revealed that the  $H_d$  varied from 0.937 to 1.000, which was consistent with the Hmong population  $H_d$  observed in this study. Meanwhile,  $\pi$  of Hmong population in this report was relatively low when compared to that in a similar report, which varied from 0.00174 to 0.00218 (Wen et al., 2005). Lao Song population in this study revealed

lower percentage of unique haplotype as compared to that in Suphan Buri (92.0%,  $N = 25$ ), where their first wave colonized (Fucharoen et al., 2001). The evidence of migration of these 2 populations resulted in a decrease in genetic diversity by founder effect. Additionally, there were higher values of genetic diversity in the Thai-Siam population who had no history of migration.

The Tajima's  $D$  and Fu's  $F_s$  were estimated to prove the null hypothesis of neutrality of genetic effect. The negative values of both parameters in all populations suggested that they were evolved without selective effect and growth in constant population size. The evolution of population was influenced by the founder effect of population expansion.

**Table 12 Diversity and neutrality testing parameters of mtDNA control region haplotypes in the three ethnic populations**

|  | Hmong      | Lao Song   | Thai-Siam  |
|--|------------|------------|------------|
| No. of sequences                         | 53         | 29         | 27         |
| Number of segregated sites               | 87         | 78         | 93         |
| Ave. no. of nucleotide differences, $k$  | 13.137     | 12.714     | 12.669     |
| No. of haplotypes                        | 36         | 25         | 25         |
| No. of unique haplotypes (% individuals) | 27 (50.94) | 23 (79.31) | 23 (85.19) |
| No. of share haplotypes                  | 9          | 2          | 2          |
| Haplotype diversity, $H_d$               | 0.979      | 0.983      | 0.994      |
| Variance of $H_d$                        | 0.00008    | 0.00029    | 0.00014    |
| Standard deviation of $H_d$              | 0.009      | 0.017      | 0.012      |
| Nucleotide diversity; $\pi$              | 0.01174    | 0.01130    | 0.01132    |
| Variance of $\pi$                        | 0.0000001  | 0.0000007  | 0.0000006  |
| Standard deviation of $\pi$              | 0.00035    | 0.00082    | 0.00079    |
| Tajima's $D$ ( $p < 0.05$ )              | -1.16069   | -1.40425   | -1.86488   |
| Fu's $F_s$                               | -11.092    | -9.431     | -11.729    |



**Table 13 The genetic differentiations by nucleotide diversity of mtDNA control region between groups of 3 ethnic populations**

|           | Hmong   | Lao Song | Thai Siam |
|-----------|---------|----------|-----------|
| Hmong     |         |          |           |
| Lao Song  | 0.04072 |          |           |
| Thai Siam | 0.03564 | 0.02312  |           |

### HVR-1 analysis

Most of the previous reports revealed only HVR-1 data because of the forensic tradition of mtDNA control region study. The HVR-1  $H_d$  and  $\pi_i$  of the three ethnic populations were calculated by the sequences from position 16,024–16,365. The HVR-1  $H_d$  in Hmong, Lao Song, and Thai-Siam populations was 0.957, 0.975, and 0.989, respectively. The HVR-1  $\pi_i$  in Hmong, Lao Song, and Thai-Siam populations was 0.02365, 0.02112, and 0.02110, respectively. Several studies on several populations in Thailand reported the range of the HVR-1  $H_d$ , which varied from 0.931 to 1.000, and that of  $\pi_i$  varied from 0.0169 to 0.0235 (Fucharoen et al., 2001; Lertrit et al., 2008; Bodner et al., 2011; Boonsoda et al., 2013; Kutanan et al., 2014). These values indicated that the genetic diversity of the three ethnic populations in this study were relatively high. Moreover, the genetic distance, referred to as the differentiation of genetic structure between populations, between the three populations (Table 13) was relatively high as compared to that revealed in previous reports (Fucharoen et al., 2001; Kutanan et al., 2014). The previous study on Tai-Kadai populations—Yuan, Leu, and Yong in the northern part of Thailand—reported that the range of percentage of unique haplotypes was 26.3–67.7 and  $H_d$  range was 0.878–0.992 ( $N = 39$ –62). The negative value of  $F_s$  also suggested the demographic expansion of populations (Kampuansai et al., 2007).

The NJ-tree of HVR-1 mtDNA showed the heterogeneity among 3 ethnic populations (Figure 21). Radial topology indicated the closed pattern of pairwise nucleotide different between populations. All 3 populations could not be separated from

each other in NJ-tree, it may cause by the immigration history. Both Lao Song and Hmong migrated from near geographical area, the south of China. Their genetic diversities were closely related. Meanwhile, there were admixture of population in Thailand due to multiple wars in Southeast Asia which brings to immigrations. When the age of population is close, NJ-tree of HVR-1 could not be used for clustering of population. We assumed that all 3 populations might share the common ancestors. The network analysis is the useful method to study demographic of population. We suggested to use network analysis to compare the relationship of population under the 2 assumptions. The first, their common ancestors might not be in present population and another one is that the genetic diversity might be the result of genetic drift.

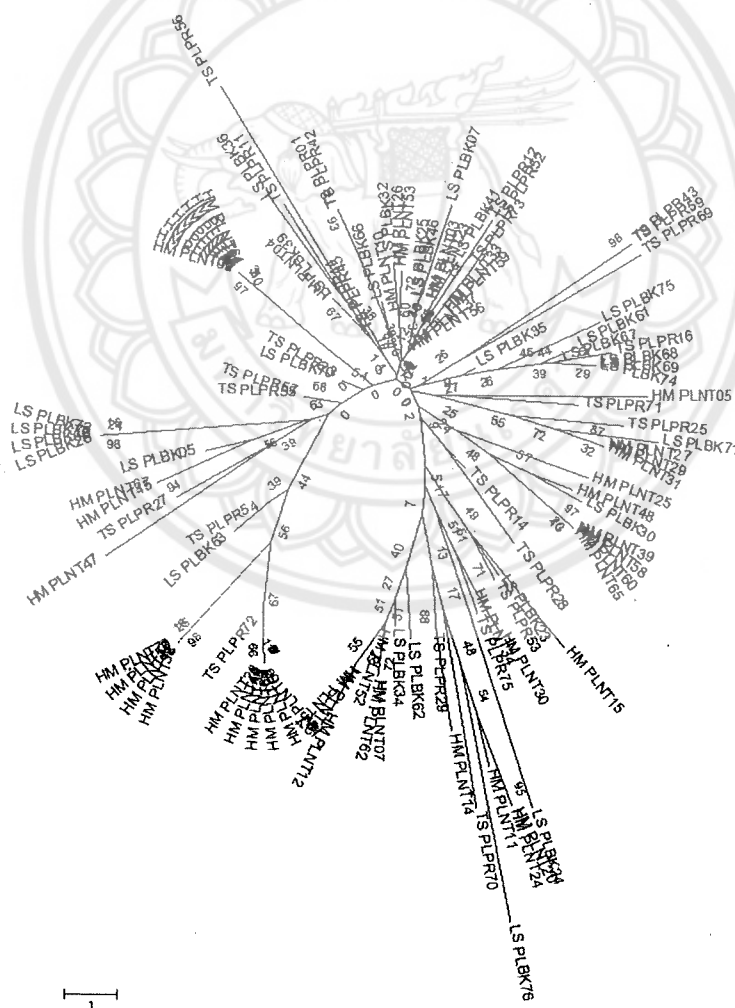


Figure 21 NJ tree of HVR-1 in 3 ethnic populations

### Y-STR analysis

Allele frequencies of 5 Y-STR loci in 281 Y-chromosome of the lower northern part of Thailand populations were summarized in Table 14. The gene diversity of all loci were high ( $>0.5$ ) except DYS391. The most common alleles in each loci show similar trend as reported in Southeast Asian populations such as in Central Thai population (Siriboonpiputtana et al., 2010), South China (Hou et al., 2001), Korean (Park et al., 2005) and Singapore (Yong, Lee, & Yap, 2006) except in Japan which the highest frequencies was allele 11 of DYS392. The range of allele distribution in each loci are smaller than global report (de Knijff et al., 1997). The haplotype analysis found 165 haplotypes (58.71%) with 112 unique haplotypes (39.86%). The frequencies of highest common haplotype show in Table 15. The haplotype diversity in this population was 0.9928 which was high informative. However, this results are only a preliminary data. More loci have to be observed for steady statistic. At least 9 common loci have to be considered (Lessig et al., 2003). The assessment of Y-STR haplotype need a large number of haplotypes in appropriate populations. The haplotype data of the Lower Northern part of Thailand are continuously examine.

**Table 14 Allele frequencies and gene diversity (*h*) of 5 Y-STR loci**

| allele | DYS19  | DYS390 | DYS391 | DYS392 | DYS393 |
|--------|--------|--------|--------|--------|--------|
| 7      |        |        | 0.0071 |        |        |
| 8      |        |        | 0      |        |        |
| 9      |        |        | 0.0463 |        |        |
| 10     |        |        | 0.7936 | 0.0071 | 0.0036 |
| 11     |        |        | 0.1530 | 0.1388 | 0      |
| 12     | 0.0249 |        |        | 0      | 0.3416 |
| 13     | 0      |        |        | 0.5018 | 0.4733 |
| 14     | 0.2135 |        |        | 0.2527 | 0.1103 |
| 15     | 0.4270 |        |        | 0.0819 | 0.0605 |
| 16     | 0.2242 |        |        | 0.0178 | 0.0107 |
| 17     | 0.1103 |        |        |        |        |
| 18     |        |        |        |        |        |
| 19     |        |        |        |        |        |
| 20     |        | 0.0356 |        |        |        |
| 21     |        | 0.0249 |        |        |        |
| 22     |        | 0.1281 |        |        |        |
| 23     |        | 0.1851 |        |        |        |
| 24     |        | 0.3701 |        |        |        |
| 25     |        | 0.2562 |        |        |        |
| n.     | 281    | 281    | 281    | 281    | 281    |
| h      | 0.7146 | 0.7507 | 0.3473 | 0.6632 | 0.6484 |

**Table 15 Frequencies of 5 loci Y-STR haplotypes**

| Haplotype | DYS19 | DYS390 | DYS391 | DYS392 | DYS393 | No. | frequencies |
|-----------|-------|--------|--------|--------|--------|-----|-------------|
| 1         | 14    | 24     | 10     | 13     | 13     | 12  | 0.042705    |
| 2         | 14    | 25     | 10     | 13     | 13     | 9   | 0.032028    |
| 3         | 16    | 24     | 10     | 13     | 12     | 8   | 0.02847     |
| 4         | 16    | 25     | 10     | 13     | 13     | 6   | 0.021352    |
| 5         | 15    | 23     | 10     | 13     | 12     | 5   | 0.017794    |
| 6         | 15    | 23     | 10     | 14     | 13     | 5   | 0.017794    |
| 7         | 15    | 25     | 10     | 14     | 12     | 5   | 0.017794    |

Present study show some polymorphism or haplotype are associated with several disease (Smart, Bolnick, & Tutton, 2017). For examples, in Asian populations, mutation at position 16189 (T->C) and several haplogroups were associated with an increasing risk of type 2 diabetes (Kwak & Park, 2016). In Thailand population, the haplogroup B5a1 was associated with the Leber hereditary optic neuropathy (LHON) , significantly, and may increase risk of visual loss. While, the haplogroup F was trended to positive protection from the disease (Kaewsutthi et al., 2011).

Hmong is an example of culturally diverse population not only in Southeast-Asian countries but they also lived in United States, France, Australia and others. Genetic studies in Hmong population will connected to genetic in medicine and health care management. For example, the DNA polymorphism analysis in carcinogen-metabolism, single-carbon metabolism and DNA repair in Hmong population who lived in Minnesota, USA showed the higher risk of nasopharyngeal cancer, gastric cancer, hepatic cancer and cervical cancer than white population (Kniffmeyer et al., 2004; Ross, Xie, Kniffmeyer, Bushhouse, & Robison, 2003).

## CHAPTER V

### CONCLUSION

The present study demonstrated the genetic structure and diversity of Thai-Siam, Hmong, and Lao Song populations living in Phitsanulok Province, Thailand. The haplotype study of mtDNA control region revealed no admixture between populations. The unique patterns of mutation in mtDNA control region were found at position 73 (A→G) and 512 (ACAC-insertion), which can be used as population markers. Thai-Siam population represented the highest diversity in all the three populations, while the genetic diversities of Hmong and Lao Song populations, who had migration history, decreased as compared to their original populations. These results confirmed the history of long-term isolation of Hmong and Lao Song populations. The Y-STR diversity of lower northern part population overall also indicated the genetic mixture of Thai population as a variety of the Southeast Asian. These results can provide the preliminary data for study the genetic relationship of population in Southeast Asia.



## REFERENCE

- Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., . . . Staden, R. Y., I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature*, 290, 457-465.
- Andrew, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M., & Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics*, 23, 147.
- Asin-Cayuela, J., & Gustafsson, C. M. (2007). Mitochondrial transcription and its regulation in mammalian cells. *Trends in Biochemical Sciences*, 32(3), 111-117. doi:<http://doi.org/10.1016/j.tibs.2007.01.003>
- Bendall, K. E., & Sykes, B. C. (1995). Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. *American Journal of Human Genetics*, 57(2), 248-256.
- Butler, J. M. (2005). *Forensic DNA typing* (2 ed.). USA: Elsevier.
- Chang, D. D., & Clayton, D. A. (1985). Priming of human mitochondrial DNA replication occurs at the light-strand promoter. *Proceedings of the National Academy of Sciences*, 82(2), 351-355.
- Clayton, D. A. (2000). Transcription and replication of mitochondrial DNA. *Human Reproduction*, 15(suppl\_2), 11-17. doi:10.1093/humrep/15.suppl\_2.11
- Cortellini, V., Verzeletti, A., Cerri, N., Cisana, S., & De Ferrari, F. (2015). Evaluation of mtDNA stability across the maternal line: A study on three generations in two families. *Forensic Science International: Genetics Supplement Series*, 5, e19-e20. doi:<https://doi.org/10.1016/j.fsigss.2015.09.008>
- de Knijff, P., Kayser, M., Caglia, A., Corach, D., Fretwell, N., Gehrig, C., . . . Roewer, L. (1997). Chromosome Y microsatellites: population genetic and evolutionary aspects. *International Journal of Legal Medicine*, 110(3), 134-140. doi:10.1007/s004140050052
- Ganfyd. (2010, 30 January 2010). File:Human Mitochondrial DNA en.svg. Retrieved from [http://www.ganfyd.org/index.php?title=File:Human Mitochondrial DNA en.svg#file](http://www.ganfyd.org/index.php?title=File:Human_Mitochondrial_DNA_en.svg#file)
- Giles, R. E., Blanc, H., Cann, H. M., & Wallace, D. C. (1980). Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences*, 77(11), 6715-6719.
- Gill, P., Ivanov, P. L., Kimpton, C., Piercy, R., Benson, N., Tully, G., . . . Sullivan, K. (1994). Identification of the remains of the Romanov family by DNA analysis. *Nat Genet*, 6(2), 130-135.
- Holt, I. J., Miller, D. H., & Harding, A. E. (1989). Genetic heterogeneity and mitochondrial DNA heteroplasmy in Leber's hereditary optic neuropathy. *Journal of Medical Genetics*, 26(12), 739.
- Hou, Y. P., Zhang, J., Li, Y. B., Wu, J., Zhang, S. Z., & Prinz, M. (2001). Allele sequences of six new Y-STR loci and haplotypes in the Chinese Han population. *Forensic Science International*, 118(2-3), 147-152. doi:[http://doi.org/10.1016/S0379-0738\(01\)00383-8](http://doi.org/10.1016/S0379-0738(01)00383-8)
- Ivanov, P. L., Wadhams, M. J., Roby, R. K., Holland, M. M., Weedn, V. W., & Parsons, T. J. (1996). Mitochondrial DNA sequence heteroplasmy in the Grand



- Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nat Genet*, 12(4), 417-420.
- Kaewsutthi, S., Phasukkijwatana, N., Joyjinda, Y., Chuenkongkaew, W., Kunhapan, B., Tun, A. W., . . . Lertrit, P. (2011). Mitochondrial Haplogroup Background May Influence Southeast Asian G11778A Leber Hereditary Optic Neuropathy. *Investigative Ophthalmology & Visual Science*, 52(7), 4742-4748. doi:10.1167/iovs.10-5816
- Khiawkham, P. (2003). *Cultural changes in community rituals and life-relate traditions : The case of Lao Song communities in Suphan Buri Province*. (Master of Art (Anthropology)), Chulalongkorn University.
- Kniffmeyer, W. R., Langer, E., Davies, S. M., Envall, J., Robison, L. L., & Ross, J. A. (2004). Genetic polymorphisms in the Hmong population. *Cancer*, 100(2), 411-417. doi:10.1002/cncr.11913
- Kutanen, W., Kampuansai, J., Fuselli, S., Nakbunlung, S., Seielstad, M., Bertorelle, G., & Kangwanpong, D. (2011b). Genetic structure of the Mon-Khmer speaking groups and their affinity to the neighbouring Tai populations in Northern Thailand. *BMC Genetics*, 12-56.
- Kwak, S. H., & Park, K. S. (2016). Role of mitochondrial DNA variation in the pathogenesis of diabetes mellitus. *Frontiers in bioscience (Landmark edition)*, 21, 1151-1167. Retrieved from <http://europepmc.org/abstract/MED/27100497>
- Lessig, R., Willuweit, S., Krawczak, M., Wu, F.-C., Pu, C.-E., Kim, W., . . . Roewer, L. (2003). Asian online Y-STR Haplotype Reference Database. *Legal Medicine*, 5, Supplement, S160-S163. doi:[https://doi.org/10.1016/S1344-6223\(02\)00100-1](https://doi.org/10.1016/S1344-6223(02)00100-1)
- Linnane, A., Ozawa, T., Marzuki, S., & Tanaka, M. (1989). MITOCHONDRIAL DNA MUTATIONS AS AN IMPORTANT CONTRIBUTOR TO AGEING AND DEGENERATIVE DISEASES. *The Lancet*, 333(8639), 642-645. doi:[http://dx.doi.org/10.1016/S0140-6736\(89\)92145-4](http://dx.doi.org/10.1016/S0140-6736(89)92145-4)
- Malka, F., Lombès, A., & Rojo, M. (2006). Organization, dynamics and transmission of mitochondrial DNA: Focus on vertebrate nucleoids. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1763(5-6), 463-472. doi:<https://doi.org/10.1016/j.bbamcr.2006.04.001>
- McDonald, J. D. (Producer). (2005, 15 April, 2017). Y-Haplogroups of the world. Retrieved from <http://www.scs.illinois.edu/~mcdonald/WorldHaplogroupsMaps.pdf>
- Mercer, Tim R., Neph, S., Dinger, Marcel E., Crawford, J., Smith, Martin A., Shearwood, A.-Marie J., . . . Mattick, John S. (2011). The Human Mitochondrial Transcriptome. *Cell*, 146(4), 645-658. doi:<http://doi.org/10.1016/j.cell.2011.06.051>
- Meyer, S., Weiss, G., & von Haeseler, A. (1999). Pattern of Nucleotide Substitution and Rate Heterogeneity in the Hypervariable Regions I and II of Human mtDNA. *Genetics*, 152(3), 1103.
- Mongkokthawat, P. (2005). *Cultural Identity and Change of Hmong at Ban Huay Nam Sai, Nakhonchai district, Phitsanulok province*. (Master degree in Anthropology), Chulalongkorn University.
- Nachman, M. W., & Crowell, S. L. (2000). Estimate of the Mutation Rate per Nucleotide in Humans. *Genetics*, 156, 297-304.

- Nei, M., & Roychoudhury, A. K. (1993). Evolutionary relationships of human population on a global scale. *Mol. Biol. Evol.*, 10(5), 927-943.
- Park, M. J., Lee, H. Y., Yoo, J.-E., Chung, U., Lee, S. Y., & Shin, K.-J. (2005). Forensic evaluation and haplotypes of 19 Y-chromosomal STR loci in Koreans. *Forensic Science International*, 152(2-3), 133-147. doi:<http://doi.org/10.1016/j.forsciint.2004.07.016>
- Parsons, T. J., Muniec, D. S., & Sullivan, K. (1997). A high observed substitution rate in the human mitochondrial DNA control region. *Nature Genetics*, 15, 363-368. doi:10.1038/ng0497-363
- Pitiphat, S., & Phunsuwan, S. (1997). *Lao Song: Dynamics of cultural system in two decades*. Retrieved from
- Ross, J. A., Xie, Y., Kiffmeyer, W. R., Bushhouse, S., & Robison, L. L. (2003). Cancer in the Minnesota Hmong population. *Cancer*, 97(12), 3076-3079. doi:10.1002/cncr.11443
- Sanyakul, W. (2004). *Factors affecting the development of economic culture of the Hmong in Kheknoi communities, Amphoe Khaokho, Changwat Phetchaboon*. (Master of Art (Anthropology)), Kasetsart University.
- Satapanawatana, J., Sirasunthorn, P., & Suwanpratest, O. (2011). *Phitsanulok: Province of Opportunities and Challenges*. Bangkok, Thailand: East Printing House Co. Ltd.
- Sato, M., & Sato, K. (2013). Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1833(8), 1979-1984. doi:<https://doi.org/10.1016/j.bbamcr.2013.03.010>
- Schliesinger, J. (2000). *Ethnic groups of Thailand: non-Tai speaking peoples*. Bangkok: White Lotus Press.
- Schliesinger, J. (2001). *Tai group of Thailand, Volume 1: Introduction and overview*. Bangkok: White Lotus Press.
- Schwartz, M., & Vissing, J. (2002). Paternal Inheritance of Mitochondrial DNA. *New England Journal of Medicine*, 347(8), 576-580. doi:10.1056/NEJMoa020350
- Shoffner, J. M., Brown, M. D., Torroni, A., Lott, M. T., Cabell, M. F., Mirra, S. S., ... Wallace, D. C. (1993). Mitochondrial DNA Variants Observed in Alzheimer Disease and Parkinson Disease Patients. *Genomics*, 17(1), 171-184. doi:<https://doi.org/10.1006/geno.1993.1299>
- Siriboonpiputtana, T., Jomsawat, U., Rinthachai, T., Thanakitgosate, J., Shotivaranon, J., Limsuwanachot, N., ... Rerkamnuaychoke, B. (2010). Y-chromosomal STR haplotypes in Central Thai population. *Forensic Science International: Genetics*, 4(3), e71-e72. doi:<http://doi.org/10.1016/j.fsigen.2009.06.007>
- Smart, A., Bolnick, D. A., & Tutton, R. (2017). Health and genetic ancestry testing: time to bridge the gap. *BMC Medical Genomics*, 10(1), 3. doi:10.1186/s12920-016-0240-3
- Srising, S. (1976). *Thai Song Dam (Lao Song) culture in Phitsanulok and Phichit Provinces*. Retrieved from Phitsanulok, Thailand:
- Thomas, M. G., Bradman, N., & Flinn, H. M. (1999). High throughput analysis of 10 microsatellite and 11 diallelic polymorphisms on the human Y-chromosome. *Hum Genet*, 105. doi:10.1007/s004399900181
- Tuppen, H. A. L., Blakely, E. L., Turnbull, D. M., & Taylor, R. W. (2010). Mitochondrial DNA mutations and human disease. *Biochimica et Biophysica*

- Acta (BBA) - Bioenergetics*, 1797(2), 113-128.  
doi:<http://doi.org/10.1016/j.bbabi.2009.09.005>
- Wadkeaw, M. (1978). *The Study of Social Structure of Lao Song*. (Master of Arts (Social Development)), Kasetsart University Library.
- Yao, Y.-G., Nie, L., Harpending, H., Fu, Y.-X., Yuan, Z.-G., & Zhang, Y.-P. (2002). Genetic Relationship of Chinese Ethnic Populations Revealed by mtDNA Sequence Diversity. *AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY*, 118, 63-76.
- Yong, R. Y. Y., Lee, L. K. H., & Yap, E. P. H. (2006). Y-chromosome STR haplotype diversity in three ethnic populations in Singapore. *Forensic Science International*, 159(2-3), 244-257.  
doi:<https://doi.org/10.1016/j.forsciint.2005.05.010>
- Zimmermann, B., Bodner, M., Amory, S., Fendt, L., Röck, A., Horst, D., . . . Sanguansermsri, T. (2009). Forensic and phylogeographic characterization of mtDNA lineages from northern Thailand (Chiang Mai). *Int J Legal Med*, 123, 495-501.



## APPENDIXES



COA No. 112/2015  
IRB No. 021/58



**คณะกรรมการจริยธรรมการวิจัยในมนุษย์**

**มหาวิทยาลัยนเรศวร**

99 หมู่ 9 ตำบลท่าโพธิ์ อำเภอเมือง จังหวัดพิษณุโลก 65000 เบอร์โทรศัพท์ 05596 8642

**เอกสารรับรองโครงการวิจัย**

คณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยนเรศวร ดำเนินการให้การรับรองโครงการวิจัยตามแนวทางหลักจริยธรรมการวิจัยในคนที่เป็นมาตรฐานสากล ได้แก่ Declaration of Helsinki, The Belmont Report, CIOMS Guideline และ International Conference on Harmonization in Good Clinical Practice หรือ ICH-GCP

**ชื่อโครงการ :** การวิเคราะห์ความสัมพันธ์ทางวิวัฒนาการของดีเอ็นเอจากไมโทคอนเดรียและโครโมโซมวายในประชากรภาคเหนือตอนล่างของประเทศไทย  
**Study Title :** Phylogenetic Relationship Analysis of mitochondrial DNA and Y-chromosome in Lower Northern Part of Thailand Populations.  
**ผู้วิจัยหลัก :** ดร. ร้อยโทหญิงสายศิริ มีระเสน  
**Principal Investigator :** Lt. Salsiri Mirasena, Ph.D.  
**สังกัดหน่วยงาน :** คณะวิทยาศาสตร์การแพทย์  
**ผู้ร่วมวิจัย :** นางสาวอรุณี สุธะสุนันท์  
ดร.มลิวรรณ นาคขุนทด  
**วิธีทบทวน :** คณะกรรมการเต็มชุด (Full Board Review)  
**รายงานความก้าวหน้า :** ส่งรายงานความก้าวหน้าอย่างน้อย 1 ครั้ง/ปี หรือ ส่งรายงานฉบับสมบูรณ์หากคำปณิธานโครงการเสร็จสิ้นก่อน 1 ปี

**เอกสารรับรอง**

1. AF 01-10 เวอร์ชัน 3.0 วันที่ 18 พฤษภาคม 2558
2. AF 02-10 เวอร์ชัน 3.0 วันที่ 18 พฤษภาคม 2558
3. AF 03-10 เวอร์ชัน 3.0 วันที่ 18 พฤษภาคม 2558
4. AF 04-10 เวอร์ชัน 3.0 วันที่ 18 พฤษภาคม 2558
5. AF 05-10 เวอร์ชัน 3.0 วันที่ 18 พฤษภาคม 2558
6. สรุปโครงการเพื่อการพิจารณาทางจริยธรรมการวิจัยในมนุษย์ เวอร์ชัน 3.0 วันที่ 18 พฤษภาคม 2558
7. โครงร่างงานวิจัย เวอร์ชัน 3.0 วันที่ 18 พฤษภาคม 2558
8. ประวัติผู้วิจัย เวอร์ชัน 2.0 วันที่ 17 เมษายน 2558
9. รายละเอียดงบประมาณโดยย่อ เวอร์ชัน 2.0 วันที่ 17 เมษายน 2558
10. แบบสัมภาษณ์โครงการวิจัย เวอร์ชัน 2.0 วันที่ 17 เมษายน 2558
11. Questionnaire เวอร์ชัน 2.0 วันที่ 17 เมษายน 2558

ลงนาม   
(นายแพทย์สมพงษ์ อึ้งอัมพรวิไล)  
ประธานคณะกรรมการจริยธรรมการวิจัยในมนุษย์

**วันที่รับรอง :** 27 พฤษภาคม 2558  
**Date of Approval :** May 27, 2015  
**วันหมดอายุ :** 27 พฤษภาคม 2559  
**Approval Expire Date :** May 27, 2016

ทั้งนี้ การรับรองนี้ไม่เ็นข้อบังคับที่ระบุด้านหลังทุกข้อ (ดูด้านหลังของเอกสารรับรองโครงการวิจัย)

แบบสอบถามข้อมูลเบื้องต้นโครงการ  
การวิเคราะห์ความสัมพันธ์ทางวิวัฒนาการของดีเอ็นเอจากไมโทคอนเดรียและโครโมโซมวาย  
ในประชากรภาคเหนือตอนล่างของประเทศไทย  
ภาควิชาชีววิทยา คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร

กรุณาขีด ✓ หรือเติมข้อความลงในช่อง

ตอนที่ 1 ข้อมูลพื้นฐานผู้ให้ข้อมูล

1. ผู้ให้ข้อมูล ชื่อ-นามสกุล .....
2. ตำแหน่ง .....
3. ที่อยู่ หมู่ ..... บ้าน ..... ตำบล.....อำเภอ.....จังหวัด..... 4. หมายเลขโทรศัพท์ที่ติดต่อได้ .....

ตอนที่ 2 ข้อมูลชาติพันธุ์เบื้องต้น

1. ท่านทราบว่าในพื้นที่ของท่านมีผู้อาศัยที่มีชาติพันธุ์ต่างๆ หรือไม่

☐ มี

☐ ไม่มี (ข้ามไปตอบข้อ 7)

2. ในพื้นที่ของท่านมีผู้อาศัยที่มีชาติพันธุ์ใดบ้าง (ตอบได้หลายข้อ)

☐ ไทย (ภาคกลาง)

☐ ไทยใหญ่

☐ ไทยดำ

☐ ผู้ไทย

☐ ไทย (โคราช)

☐ ลาวโซ่ง (ไทยดำ, ไทยทรงดำ)

☐ ลาวครั่ง

☐ ลาวล้านนา

☐ ลาวหลวงพระบาง

☐ ลาวพวน(ไทยพวน)

☐ ลาวเวียงจันทน์

☐ ลาวเงี้ยว

☐ ม้ง

☐ เย้า

☐ กระเหรี่ยง

☐ มูเซอ

☐ ลีซู

☐ จีน

☐ อินเดีย

☐ เวียดนาม

☐ พม่า

☐ อื่นๆ คือ .....

3. ในพื้นที่ของท่านมีประชากรตามข้อ 2 อยู่มากเท่าใด (โดยประมาณ)

.....ครัวเรือน

.....คน

4. ในพื้นที่ของท่านมีผู้รู้หรือเอกสารที่ทำให้ติดตาม ข้อมูลชาติพันธุ์ย้อนไปในอดีตได้หรือไม่

☐ มี

☐ ไม่มี

5. หากผู้วิจัยต้องการสอบถามหรือหาข้อมูลเกี่ยวกับชาติพันธุ์ในพื้นที่เพิ่มเติม สามารถติดต่อได้จากผู้ใด

ชื่อ-นามสกุล .....

โทรศัพท์ .....

6. หากผู้วิจัยต้องการเก็บตัวอย่างเลือดจากอาสาสมัครประมาณ 3-5 cc. (มีค่าตอบแทนตามสมควร)

จะเป็นไปได้หรือไม่

☐ ได้

☐ ไม่ได้

7. หากท่านมีข้อมูลของชาติพันธุ์ในพื้นที่ของท่าน (หรือพื้นที่อื่น) เช่น ถิ่นกำเนิด การอพยพโยกย้ายของประชากร ช่วงเวลาในการโยกย้ายของประชากร ขอความกรุณาให้ข้อมูลเพื่อเป็นแนวทางของการวิจัยต่อไป

ขอขอบพระคุณในการให้ข้อมูล

For researcher

Code .....

Date .....

Collector .....

Questionnaire

**Phylogenetic Relationship Analysis of Mitochondrial DNA and Y-chromosome  
in Lower Northern Part of Thailand Populations**

Please give the information for these questions

**Part 1 General information**

1. Sex ☐ male ☐ female

2. Age ..... years

3. Ethnicity .....

4. Resident place District ..... Province .....

5. Birth place District ..... Province .....

**Part 2 Family background**

1. Do you know your family history? ☐ Yes ☐ No

2. Mother's family background

2.1 Ethnicity .....

2.2 Birth place District ..... Province .....

2.3 How many generation of your mother's family do you know?  
..... generations

2.4 Family background (migration, briefly) .....  
.....

3. Father's family background

3.1 Ethnicity .....

3.2 Birth place District ..... Province .....

3.3 How many generation of your mother's family do you know?  
..... generations

3.4 Family background (migration, briefly) .....  
.....

สำหรับผู้เก็บข้อมูล  
รหัสข้อมูล .....  
วันที่.....  
ผู้บันทึก.....

แบบสัมภาษณ์โครงการวิจัย  
การวิเคราะห์ความสัมพันธ์ทางวิวัฒนาการของดีเอ็นเอจากไมโทคอนเดรีย  
และโครโมโซมวายในประชากรภาคเหนือตอนล่างของประเทศไทย

คำชี้แจง โปรดตอบคำถามต่อไปนี้

ตอนที่ 1 ข้อมูลทั่วไป

1. เพศ ☐ ชาย ☐ หญิง
2. อายุ ..... ปี
3. เชื้อชาติ .....
4. ที่อยู่ปัจจุบัน อำเภอ ..... จังหวัด .....
5. สถานที่เกิด อำเภอ ..... จังหวัด .....

ตอนที่ 2 ข้อมูลเฉพาะ

1. ท่านทราบประวัติครอบครัวในอดีตของท่านหรือไม่ ☐ ทราบ ☐ ไม่ทราบ
2. ประวัติครอบครัวของท่านมีเชื้อสายมาจากชาติพันธุ์ใดต่อไปนี้  

|  |   |   |                                    |
|--|---|---|------------------------------------|
| <input type="checkbox"/> ไทย (ภาคกลาง) | <input type="checkbox"/> ไทยใหญ่            | <input type="checkbox"/> ไทยดำ          | <input type="checkbox"/> ผู้ไทย    |
| <input type="checkbox"/> ไทย (โคราช)   | <input type="checkbox"/> ลาวโซ่ง (ไทยทรงดำ) | <input type="checkbox"/> ลาวครั่ง       | <input type="checkbox"/> ลาวล้านนา |
| <input type="checkbox"/> ลาวหลวงพระบาง | <input type="checkbox"/> ลาวพวน(ไทยพวน)     | <input type="checkbox"/> ลาวเวียงจันทร์ | <input type="checkbox"/> ลาวเงี้ยว |
| <input type="checkbox"/> ม้ง           | <input type="checkbox"/> เย้า               | <input type="checkbox"/> กระเหรี่ยง     | <input type="checkbox"/> มูเซอ     |
| <input type="checkbox"/> ลีซู          | <input type="checkbox"/> จีน                | <input type="checkbox"/> อินเดีย        | <input type="checkbox"/> เวียดนาม  |
| <input type="checkbox"/> พม่า          | <input type="checkbox"/> อื่นๆ คือ .....    |   |                                    |

3. ครอบครัวฝ่ายมารดา

- 3.1 เชื้อชาติ .....
- 3.2 ภูมิลำเนาของครอบครัวอยู่ที่ อำเภอ ..... จังหวัด .....
- 3.3 ท่านทราบอดีตของครอบครัวฝ่ายมารดาย้อนขึ้นไป ..... ชั่วรุ่น .....
- 3.4 ประวัติครอบครัว (การอพยพย้ายถิ่น โดยสังเขป) .....

3. ครอบครัวฝ่ายบิดา

- 3.1 เชื้อชาติ .....
- 3.2 ภูมิลำเนาของครอบครัวอยู่ที่ อำเภอ ..... จังหวัด .....
- 3.3 ท่านทราบอดีตของครอบครัวฝ่ายบิดาย้อนขึ้นไป ..... ชั่วรุ่น .....
- 3.4 ประวัติครอบครัว (การอพยพย้ายถิ่น โดยสังเขป) .....



**Table A1 Basis data of Thai Siam from Phrom Phiram district, Phitsanulok province**

| No.     | เพศ | อายุ | เชื้อชาติ | ที่เกิด   | มารดา          |                     | บิดา      |                     |
|---------|-----|------|-----------|-----------|----------------|---------------------|-----------|---------------------|
|         |     |      |           |           | เชื้อชาติ      | ภูมิลำเนา           | เชื้อชาติ | ภูมิลำเนา           |
| PLPR-01 | F   | 56   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-02 | F   | 49   | ไทย       | พรหมพิราม | ไทย            | บางกระทุ่ม พิษณุโลก | ไทย       | บางกระทุ่ม พิษณุโลก |
| PLPR-03 | F   | 57   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-04 | F   | 52   | ไทย       | พรหมพิราม | ไทย            | บางกระทุ่ม พิษณุโลก | ไทย       | บางกระทุ่ม พิษณุโลก |
| PLPR-05 | M   | 62   | ลาวเวียง  | พรหมพิราม | ลาวเวียง       | พรหมพิราม พิษณุโลก  | ลาวเวียง  | พรหมพิราม พิษณุโลก  |
| PLPR-06 | F   | 62   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-07 | F   | 44   | ไทย       | พรหมพิราม | ไทย            | อยุธยา              | ไทย       | อยุธยา              |
| PLPR-08 | F   | 60   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-09 | M   | 46   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-10 | F   | 28   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-11 | F   | 52   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-12 | F   | 54   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-13 | F   | 56   | ไทย       | พรหมพิราม | ไทย            | ลพบุรี              | ไทย       | ลพบุรี              |
| PLPR-14 | F   | 53   | ไทย       | พรหมพิราม | ไทย            | วังทอง พิษณุโลก     | ไทย       | บางมูลนาค พิจิตร    |
| PLPR-15 | F   | 61   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-16 | F   | 59   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-17 | F   | 60   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-18 | F   | 57   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-19 | F   | 22   | ไทย       | ทับคล้อ   | ไทย            | ทับคล้อ พิจิตร      | ไทย       | ทับคล้อ พิจิตร      |
| PLPR-20 | F   | 54   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-21 | F   | 59   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-22 | F   | 55   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-23 | F   | 63   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-24 | M   | 56   | ไทย       | พรหมพิราม | ไทย            | โคกสำโรง ลพบุรี     | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-25 | F   | 59   | ลาว       | พรหมพิราม | ลาว<br>(อีสาน) | ขอนแก่น             | ลาวพวน    | ลพบุรี              |
| PLPR-26 | F   | 57   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-27 | F   | 60   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |
| PLPR-28 | F   | 28   | ไทย       | พรหมพิราม | ไทย            | มุกดาหาร            | ไทย       | มุกดาหาร            |
| PLPR-29 | M   | 29   | ไทย       | พรหมพิราม | ไทย            | พรหมพิราม พิษณุโลก  | ไทย       | พรหมพิราม พิษณุโลก  |

Table A1 (cont.)

| No.     | เพศ | อายุ | เชื้อชาติ | ที่เกิด               | มารดา     |                        | บิดา      |                        |
|---------|-----|------|-----------|-----------------------|-----------|------------------------|-----------|------------------------|
|         |     |      |           |                       | เชื้อชาติ | ภูมิลำเนา              | เชื้อชาติ | ภูมิลำเนา              |
| PLPR-30 | F   | 40   | ไทย       | พรหมพิราม             | ไทย       | พิชัย อุตรดิตถ์        | ไทย       | พิชัย อุตรดิตถ์        |
| PLPR-31 | F   | 45   | ไทย       | บัว                   | ไทย       | บัว น่าน               | ไทย       | บัว น่าน               |
| PLPR-32 | F   | 56   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-33 | F   | 46   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-34 | F   | 60   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-35 | M   | 54   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-36 | F   | 42   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | จีน       | เมืองพิษณุโลก          |
| PLPR-37 | M   | 47   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | ลพบุรี                 |
| PLPR-38 | F   | 37   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-39 | F   | 37   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-40 | F   | 48   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-41 | F   | 52   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-42 | M   | 67   | ไทย       | พรหมพิราม             | ไทย       | อุตรดิตถ์              | ลาว       | ชัยภูมิ                |
| PLPR-43 | M   | 68   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | อุตรดิตถ์              |
| PLPR-44 | M   | 49   | ไทย       | พรหมพิราม             | ไทย       | อุตรดิตถ์              | ไทย       | อุตรดิตถ์              |
| PLPR-45 | M   | 75   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-46 | M   | 68   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-47 | M   | 54   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-48 | M   | 56   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-49 | M   | 53   | ไทย       | พรหมพิราม             | ไทย       | หนองโดน ออยุธยา        | ไทย       | ชัยภูมิ                |
| PLPR-50 | M   | 64   | ลาว       | พรหมพิราม             | ลาว       | นครไทย พิษณุโลก        | ลาว       | ปราจีน                 |
| PLPR-51 | M   | 46   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | บ้านหมี่ ลพบุรี        |
| PLPR-52 | M   | 50   | ไทย       | พรหมพิราม             | ไทย       | โพธิ์ประทับช้าง พิจิตร | ไทย       | โพธิ์ประทับช้าง พิจิตร |
| PLPR-53 | M   | 68   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-54 | M   | 53   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-55 | M   | 64   | ไทย       | ศรีนคร<br>สุโขทัย     | ไทย       | ศรีนคร สุโขทัย         | ไทย       | ศรีนคร สุโขทัย         |
| PLPR-56 | M   | 76   | ไทย       | พรหมพิราม             | ไทย       | พรหมพิราม พิษณุโลก     | ไทย       | พรหมพิราม พิษณุโลก     |
| PLPR-57 | M   | 51   | ไทย       | โพธิ์ธาราม<br>ราชบุรี | ไทย       | โพธิ์ธาราม ราชบุรี     | ไทย       | โพธิ์ธาราม ราชบุรี     |

Table A1 (cont.)

| No.     | เพศ | อายุ | เชื้อชาติ | ที่เกิด   | มารดา     |                    | บิดา           |                    |
|---------|-----|------|-----------|-----------|-----------|--------------------|----------------|--------------------|
|         |     |      |           |           | เชื้อชาติ | ภูมิลำเนา          | เชื้อชาติ      | ภูมิลำเนา          |
| PLPR-58 | M   | 24   | ไทย       | พรหมพิราม | ไทย       | พิจิตร             | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-59 | M   | 68   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | อุตรดิตถ์          |
| PLPR-60 | M   | 67   | ไทย       | พรหมพิราม | ไทย       | อุตรดิตถ์          | ลาว            | ชัยภูมิ            |
| PLPR-61 | M   | 80   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-62 | M   | 52   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-63 | M   | 54   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-64 | M   | 58   | ไทย       | พรหมพิราม | จีน       | บางปะหัน ออยุธยา   | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-65 | M   | 68   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-66 | M   | 53   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-67 | M   | 64   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ลาว<br>(อีสาน) |                    |
| PLPR-68 | M   | 65   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-69 | M   | 38   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-70 | M   | 75   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-71 | M   | 61   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-72 | M   | 75   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |
| PLPR-73 | M   | 74   | ไทย       | พรหมพิราม | ไทย       | ชุมแสง นครสวรรค์   | ไทย            | ชุมแสง นครสวรรค์   |
| PLPR-74 | M   | 78   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ลาว            | ลพบุรี             |
| PLPR-75 | M   | 71   | ไทย       | พรหมพิราม | ไทย       | พรหมพิราม พิษณุโลก | ไทย            | พรหมพิราม พิษณุโลก |

**Table A2 Basis Data of Lao Song from Bang Rakam district, Phitsanulok province**

| รหัสข้อมูล | เพศ | อายุ | เชื้อชาติ | สถานที่เกิด | มารดา     |                  | บิดา      |                  |
|------------|-----|------|-----------|-------------|-----------|------------------|-----------|------------------|
|            |     |      |           |             | เชื้อชาติ | ภูมิลำเนา        | เชื้อชาติ | ภูมิลำเนา        |
| PLBK-01    | M   | 43   | ลาวเวียง  | พิษณุโลก    | ลาวเวียง  | บางระกำ พิษณุโลก | ลาวเวียง  | บางระกำ พิษณุโลก |
| PLBK-02    | F   | 42   | ลาวเวียง  | พิษณุโลก    | ลาวเวียง  | บางระกำ พิษณุโลก | อีสาน     | บางระกำ พิษณุโลก |
| PLBK-03    | F   | 60   | ไทย       | พิษณุโลก    | ไทย       | บางระกำ พิษณุโลก | ลาวเวียง  | บางระกำ พิษณุโลก |
| PLBK-04    | F   | 60   | ลาวเวียง  | พิษณุโลก    | ลาวเวียง  | บางมูลนาค พิจิตร | ลาวอีสาน  | -                |
| PLBK-05    | M   | 63   | ลาวเวียง  | พิษณุโลก    | ลาวเวียง  | ภาชี ออยุธยา     | ลาวเวียง  | ภาชี ออยุธยา     |
| PLBK-06    | M   | 60   | ลาวเวียง  | พิษณุโลก    | ลาวเวียง  | ภาชี ออยุธยา     | ลาวเวียง  | บางระกำ พิษณุโลก |
| PLBK-07    | M   | 81   | ไทย       | พิษณุโลก    | ลาวเวียง  | ภาชี ออยุธยา     | ไทย       | ท่าวุ้ง ไร่ชัย   |
| PLBK-08    | F   | 72   | ไทย       | นครปฐม      | ลาวโซ่ง   | บางแพ ราชบุรี    | จีน       |                  |
| PLBK-09    | M   | 63   | ไทย       | พิษณุโลก    | ไทย       | บางระกำ พิษณุโลก | ลาวเวียง  | บางระกำ พิษณุโลก |
| PLBK-10    | F   | 65   | ไทย       | พิษณุโลก    | ไทย       | บางระกำ พิษณุโลก | ลาวเวียง  | บางระกำ พิษณุโลก |
| PLBK-11    | M   | 71   | ไทย       | พิษณุโลก    | ลาวพวน    | หนองแซง สระบุรี  | จีน       |                  |
| PLBK-12    | M   | 51   | ไทย       | พิษณุโลก    | ไทย       | อยุธยา           | จีน       | สามง่าม พิจิตร   |
| PLBK-13    | F   | 46   | ไทย       | พิษณุโลก    | โคราช     | นครราชสีมา       | ไทย       | บางระกำ พิษณุโลก |
| PLBK-14    | M   | 82   | ไทย       | พิษณุโลก    | ไทย       | อยุธยา           | ไทย       | อยุธยา           |
| PLBK-15    | F   | 86   | ไทย       | พิษณุโลก    | ลาวพวน    | สระบุรี          | ไทย       | บางระกำ พิษณุโลก |
| PLBK-16    | F   | 52   | ไทย       | พิษณุโลก    | ไทย       | บางระกำ พิษณุโลก | ไทย       | บางระกำ พิษณุโลก |
| PLBK-17    | F   | 70   | ไทยพวน    | พิษณุโลก    | ไทย       | บางระกำ พิษณุโลก | ลาวพวน    | ภาชี ออยุธยา     |
| PLBK-18    | F   | 50   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | สุพรรณบุรี       | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-19    | F   | 45   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | อุทอง สุพรรณบุรี | ลาวโซ่ง   | บางระกำ พิษณุโลก |

Table A2 (cont.)

| รหัสข้อมูล | เพศ | อายุ | เชื้อชาติ | สถานที่เกิด | มารดา     |                   | บิดา      |                    |
|------------|-----|------|-----------|-------------|-----------|-------------------|-----------|--------------------|
|            |     |      |           |             | เชื้อชาติ | ภูมิลำเนา         | เชื้อชาติ | ภูมิลำเนา          |
| PLBK-20    | M   | 56   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-21    | F   | 57   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-22    | F   | 70   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-23    | F   | 54   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-24    | M   | 68   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-25    | F   | 55   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | วชิรบำรุง พิษณุโลก |
| PLBK-26    | F   | 45   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | นครปฐม            | ลาวโซ่ง   | นครปฐม             |
| PLBK-27    | M   | 70   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | กำแพงแสน นครปฐม   | ลาวโซ่ง   | กำแพงแสน นครปฐม    |
| PLBK-28    | F   | 42   | ไทย       | พิษณุโลก    | ไทย       | กำแพงดิน พิษณุโลก | ลาวโซ่ง   | เพชรบุรี           |
| PLBK-29    | F   | 58   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บ้านโป่ง เพชรบุรี  |
| PLBK-30    | F   | 44   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | สุพรรณบุรี         |
| PLBK-31    | F   | 72   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-32    | F   | 53   | ลาวโซ่ง   | นครสวรรค์   | ลาวโซ่ง   | นครปฐม            | ลาวโซ่ง   | นครปฐม             |
| PLBK-33    | M   | 73   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | นครปฐม            | ลาวโซ่ง   | นครปฐม             |
| PLBK-34    | F   | 70   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-35    | M   | 56   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-36    | F   | 60   | ลาวโซ่ง   | สุพรรณบุรี  | ลาวโซ่ง   | สุพรรณบุรี        | ลาวโซ่ง   | สุพรรณบุรี         |
| PLBK-37    | F   | 72   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-38    | F   | 67   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK-39    | M   | 62   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | นครปฐม            | ลาวโซ่ง   | นครปฐม             |

Table A2 (cont.)

| รหัสข้อมูล | เพศ | อายุ | เชื้อชาติ | สถานที่เกิด | มารดา     |                       | บิดา      |                  |
|------------|-----|------|-----------|-------------|-----------|-----------------------|-----------|------------------|
|            |     |      |           |             | เชื้อชาติ | ภูมิลำเนา             | เชื้อชาติ | ภูมิลำเนา        |
| PLBK-40    | F   | 50   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก      | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-41    | M   | 52   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก      | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-42    | F   | 63   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก      | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-43    | F   | 50   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก      | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-44    | F   | 53   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | นครปฐม                | ลาวโซ่ง   | นครสวรรค์        |
| PLBK-45    | F   | 47   | ลาวโซ่ง   | สุโขทัย     | ลาวโซ่ง   | บางระกำ พิษณุโลก      | จีน       | -                |
| PLBK-46    | M   | 51   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก      | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-47    | M   | 80   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | ราชบุรี               | ลาวโซ่ง   | ราชบุรี          |
| PLBK-48    | F   | 32   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | พิษณุโลก              | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-49    | F   | 73   | ลาวโซ่ง   | สุพรรณบุรี  | ลาวโซ่ง   | สุพรรณบุรี            | ลาวโซ่ง   | สุพรรณบุรี       |
| PLBK-50    | F   | 58   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | นครปฐม                | ลาวโซ่ง   | นครปฐม           |
| PLBK-51    | M   | 74   | ลาวโซ่ง   | พิจิตร      | ลาวโซ่ง   | ไพร่ประทับช้าง พิจิตร | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-52    | F   | 52   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | พิษณุโลก              | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-53    | F   | 53   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | ราชบุรี               | ลาวโซ่ง   | ราชบุรี          |
| PLBK-54    | F   | 51   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก      | ลาวโซ่ง   | วชิรบำรุง พิจิตร |
| PLBK-55    | F   | 57   | ไทย-ลาว   | พิจิตร      | ไทย       | สามง่าม พิจิตร        | ลาวโซ่ง   | ราชบุรี          |
| PLBK-56    | M   | 56   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | อุทอง สุพรรณบุรี      | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-57    | F   | 68   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางเลน นครปฐม         | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-58    | F   | 72   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก      | ลาวโซ่ง   | บางระกำ พิษณุโลก |
| PLBK-59    | M   | 67   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | สมุทรสาคร             | ลาวโซ่ง   | สมุทรสาคร        |

Table A2 (cont.)

| รหัสข้อมูล | เพศ | อายุ | เชื้อชาติ | สถานที่เกิด | มารดา     |                   | บิดา      |                    |
|------------|-----|------|-----------|-------------|-----------|-------------------|-----------|--------------------|
|            |     |      |           |             | เชื้อชาติ | ภูมิลำเนา         | เชื้อชาติ | ภูมิลำเนา          |
| PLBK61     | F   | 62   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK62     | F   | 74   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | สมุทรสาคร         | ลาวโซ่ง   | สมุทรสาคร          |
| PLBK63     | F   | 29   | ลาวโซ่ง   | พิจิตร      | ลาวโซ่ง   | พิจิตร            | ลาวโซ่ง   | พิจิตร             |
| PLBK64     | F   | 63   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | เพชรบุรี           |
| PLBK65     | F   | 83   | ลาวโซ่ง   | เพชรบุรี    | ลาวโซ่ง   | หัวสะพาน เพชรบุรี | ลาวโซ่ง   | หัวสะพาน เพชรบุรี  |
| PLBK66     | F   | 54   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK67     | F   | 55   | ลาวโซ่ง   | เพชรบุรี    | ลาวโซ่ง   | หัวสะพาน เพชรบุรี | ลาวโซ่ง   | คลองหยวก นครสวรรค์ |
| PLBK68     | F   | 64   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK69     | F   | 53   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK70     | F   | 52   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | นครสวรรค์          |
| PLBK71     | F   | 69   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | นครปฐม            | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK72     | F   | 90   | ลาวโซ่ง   | นครปฐม      | ลาวโซ่ง   | นครปฐม            | ลาวโซ่ง   | กำแพงแสน นครปฐม    |
| PLBK73     | F   | 57   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | กำแพงแสน นครปฐม   | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK74     | M   | 76   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK75     | M   | 54   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK76     | M   | 53   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | กำแพงดิน พิจิตร   | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK77     | M   | 75   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | ราชบุรี           | ลาวโซ่ง   | ราชบุรี            |
| PLBK78     | M   | 66   | ลาวโซ่ง   | เพชรบุรี    | ลาวโซ่ง   | เพชรบุรี          | ลาวโซ่ง   | เพชรบุรี           |
| PLBK79     | M   | 65   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | บางระกำ พิษณุโลก  | ลาวโซ่ง   | บางระกำ พิษณุโลก   |
| PLBK80     | M   | 63   | ลาวโซ่ง   | พิษณุโลก    | ลาวโซ่ง   | ดงยาง สุพรรณบุรี  | ลาวโซ่ง   | ดงยาง สุพรรณบุรี   |

Table A3 Basis data of Hmong from Nakhon Thai district, Phitsanulok province

| รหัสข้อมูล | เพศ | อายุ | เชื้อชาติ | ที่เกิด  | มารดา     |                  | บิดา      |                  |
|------------|-----|------|-----------|----------|-----------|------------------|-----------|------------------|
|            |     |      |           |          | เชื้อชาติ | ภูมิลำเนา        | เชื้อชาติ | ภูมิลำเนา        |
| PLNT-01    | F   | 17   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-02    | F   | 17   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-03    | F   | 40   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-04    | M   | 35   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-05    | F   | 28   | ม้ง       | ตาก      | ม้ง       | เขาค้อ เพชรบูรณ์ | ม้ง       | เขาค้อ เพชรบูรณ์ |
| PLNT-06    | M   | 20   | ม้ง       | พิษณุโลก | ม้ง       | ด่านซ้าย เลย     | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-07    | M   | 23   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-08    | F   | 42   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-09    | M   | 24   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-10    | F   | 21   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-11    | M   | 30   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-12    | M   | 45   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-13    | M   | 28   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ไทย       | พะเยา            |
| PLNT-14    | F   | 48   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-15    | M   | 21   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-16    | F   | 47   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-17    | M   | 39   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-18    | F   | 66   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |
| PLNT-19    | F   | 21   | ม้ง       | พิษณุโลก | ม้ง       | ทุ่งช้าง น่าน    | ม้ง       | ทุ่งช้าง น่าน    |
| PLNT-20    | F   | 43   | ม้ง       | พิษณุโลก | ม้ง       | นครไทย พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก  |



Table A3 (cont.)

| รหัสข้อมูล | เพศ | อายุ | เชื้อชาติ | ที่เกิด   | มารดา     |                    | บิดา      |                    |
|------------|-----|------|-----------|-----------|-----------|--------------------|-----------|--------------------|
|            |     |      |           |           | เชื้อชาติ | ภูมิลำเนา          | เชื้อชาติ | ภูมิลำเนา          |
| PLNT-21    | M   | 54   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-22    | F   | 47   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-23    | F   | 60   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-24    | F   | 47   | ม้ง       | เพชรบูรณ์ | ม้ง       | หล่มเก่า เพชรบูรณ์ | ม้ง       | หล่มเก่า เพชรบูรณ์ |
| PLNT-25    | F   | 27   | ม้ง       | เพชรบูรณ์ | ม้ง       | เขาค้อ เพชรบูรณ์   | ม้ง       | เขาค้อ เพชรบูรณ์   |
| PLNT-26    | M   | 71   | ม้ง       | เชียงราย  | ม้ง       | เชียงคำ เชียงราย   | ม้ง       | เชียงคำ เชียงราย   |
| PLNT-27    | F   | 44   | ม้ง       | พิษณุโลก  | ม้ง       | ด่านซ้าย เลย       | ม้ง       | ด่านซ้าย เลย       |
| PLNT-28    | F   | 77   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-29    | M   | 48   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-30    | M   | 29   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-31    | F   | 60   | ม้ง       | เพชรบูรณ์ | ม้ง       | หล่มเก่า เพชรบูรณ์ | ม้ง       | หล่มเก่า เพชรบูรณ์ |
| PLNT-32    | F   | 34   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-33    | F   | 33   | ม้ง       | เพชรบูรณ์ | ม้ง       | เขาค้อ เพชรบูรณ์   | ม้ง       | เขาค้อ เพชรบูรณ์   |
| PLNT-34    | F   | 56   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-35    | F   | 90   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-36    | M   | 54   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-37    | F   | 53   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-38    | F   | 51   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |
| PLNT-39    | F   | 54   | ม้ง       | เพชรบูรณ์ | ม้ง       | เขาค้อ เพชรบูรณ์   | ม้ง       | เขาค้อ เพชรบูรณ์   |
| PLNT-40    | F   | 51   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก    |

| รหัสข้อมูล | เพศ | อายุ | เชื้อชาติ | ที่เกิด   | มารดา     |                    | บิดา      |                     |
|------------|-----|------|-----------|-----------|-----------|--------------------|-----------|---------------------|
|            |     |      |           |           | เชื้อชาติ | ภูมิลำเนา          | เชื้อชาติ | ภูมิลำเนา           |
| PLNT-43    | M   | 27   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-44    | M   | 57   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-45    | M   | 17   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-46    | F   | 83   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-47    | M   | 50   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-48    | F   | 27   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | ชาติตระการ พิษณุโลก |
| PLNT-49    | F   | 22   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-50    | F   | 36   | ม้ง       | เพชรบูรณ์ | ม้ง       | เขาค้อ เพชรบูรณ์   | ม้ง       | เขาค้อ เพชรบูรณ์    |
| PLNT-51    | M   | 34   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-52    | M   | 81   | ม้ง       | พิษณุโลก  | ม้ง       | บ้านนา             | ม้ง       | บ้านนา              |
| PLNT-53    | M   | 72   | ม้ง       | เขียงราย  | ม้ง       | ญวน                | ม้ง       | ญวน                 |
| PLNT-54    | M   | 56   | ม้ง       | เขียงราย  | ม้ง       | เขียงคำ เขียงราย   | ม้ง       | เขียงคำ เขียงราย    |
| PLNT-55    | M   | 25   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-56    | M   | 28   | ม้ง       | พิษณุโลก  | ม้ง       | เขาค้อ เพชรบูรณ์   | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-57    | M   | 26   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก    | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-58    | M   | 62   | ม้ง       | เพชรบูรณ์ | ม้ง       | เขาค้อ เพชรบูรณ์   | ม้ง       | เขาค้อ เพชรบูรณ์    |
| PLNT-59    | M   | 22   | ม้ง       | พิษณุโลก  | ม้ง       | ด่านซ้าย เลย       | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-60    | M   | 55   | ม้ง       | เพชรบูรณ์ | ม้ง       | จีน                | ม้ง       | จีน                 |
| PLNT-61    | M   | 48   | ม้ง       | เพชรบูรณ์ | ม้ง       | หล่มเก่า เพชรบูรณ์ | ม้ง       | หล่มเก่า เพชรบูรณ์  |
| PLNT-62    | F   | 18   | ม้ง       | เพชรบูรณ์ | ม้ง       | หล่มเก่า เพชรบูรณ์ | ม้ง       | หล่มเก่า เพชรบูรณ์  |

| รหัสข้อมูล | เพศ | อายุ | เชื้อชาติ | ที่เกิด   | มารดา     |                     | บิดา      |                     |
|------------|-----|------|-----------|-----------|-----------|---------------------|-----------|---------------------|
|            |     |      |           |           | เชื้อชาติ | ภูมิลำเนา           | เชื้อชาติ | ภูมิลำเนา           |
| PLNT-63    | F   | 26   | ม้ง       | เพชรบูรณ์ | ม้ง       | หล่มเก่า เพชรบูรณ์  | ม้ง       | หล่มเก่า เพชรบูรณ์  |
| PLNT-64    | F   | 27   | ม้ง       | เพชรบูรณ์ | ม้ง       | บ้านนา              | ม้ง       | บ้านนา              |
| PLNT-65    | F   | 56   | ม้ง       | เพชรบูรณ์ | ม้ง       | ยุรนาน              | ม้ง       | ยุรนาน              |
| PLNT-66    | F   | 24   | ม้ง       | พิษณุโลก  | ม้ง       | ชาติตระการ พิษณุโลก | ม้ง       | ชาติตระการ พิษณุโลก |
| PLNT-67    | F   | 40   | ม้ง       | เชียงราย  | ม้ง       | เชียงของ เชียงราย   | ม้ง       | เชียงของ เชียงราย   |
| PLNT-68    | F   | 40   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก     | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-69    | F   | 45   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก     | ม้ง       | นครไทย พิษณุโลก     |
| PLNT-70    | F   | 54   | ม้ง       | พิษณุโลก  | ม้ง       | นครไทย พิษณุโลก     | ม้ง       | นครไทย พิษณุโลก     |

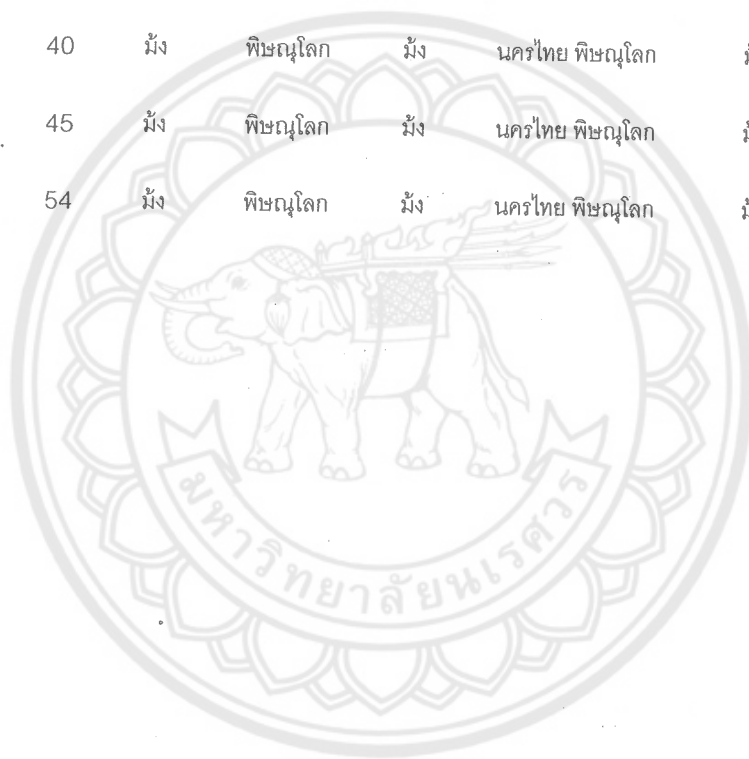


Table A4 The frequency distribution of mutation at each variation site of three ethnic populations

|             |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| rCRS_D-loop | 16162 | 16158 | 16157 | 16153 | 16148 | 16145 | 16140 | 16136 | 16129 | 16126 | 16124 | 16111 | 16108 | 16093 | 16092 | 16086 | 16051 | 16042 |
| Hmong       | 0.019 | 0.000 | 0.019 | 0.000 | 0.038 | 0.000 | 0.113 | 0.019 | 0.245 | 0.000 | 0.038 | 0.189 | 0.189 | 0.113 | 0.038 | 0.057 | 0.000 | 0.019 |
| Lao Song    | 0.034 | 0.000 | 0.034 | 0.034 | 0.000 | 0.000 | 0.069 | 0.000 | 0.310 | 0.000 | 0.034 | 0.000 | 0.000 | 0.034 | 0.000 | 0.000 | 0.034 | 0.034 |
| Thai -Siam  | 0.000 | 0.000 | 0.037 | 0.000 | 0.000 | 0.111 | 0.037 | 0.037 | 0.259 | 0.037 | 0.000 | 0.037 | 0.037 | 0.111 | 0.037 | 0.037 | 0.000 | 0.000 |
| overall     | 0.018 | 0.000 | 0.028 | 0.009 | 0.018 | 0.028 | 0.083 | 0.018 | 0.266 | 0.009 | 0.028 | 0.101 | 0.101 | 0.092 | 0.028 | 0.037 | 0.009 | 0.018 |

|             |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| rCRS_D-loop | 16239 | 16234 | 16231 | 16223 | 16218 | 16217 | 16213 | 16209 | 16192 | 16189 | 16184 | 16183 | 16182 | 16181 | 16172 | 16169 | 16168 | 16167 |
| Hmong       | 0.000 | 0.075 | 0.000 | 0.377 | 0.000 | 0.094 | 0.000 | 0.019 | 0.000 | 0.283 | 0.019 | 0.245 | 0.113 | 0.000 | 0.208 | 0.038 | 0.038 | 0.000 |
| Lao Song    | 0.000 | 0.034 | 0.034 | 0.586 | 0.000 | 0.034 | 0.034 | 0.069 | 0.241 | 0.276 | 0.000 | 0.103 | 0.034 | 0.034 | 0.069 | 0.034 | 0.000 | 0.000 |
| Thai -Siam  | 0.000 | 0.074 | 0.037 | 0.630 | 0.037 | 0.111 | 0.000 | 0.074 | 0.111 | 0.222 | 0.037 | 0.148 | 0.074 | 0.074 | 0.074 | 0.000 | 0.074 | 0.037 |
| overall     | 0.009 | 0.064 | 0.018 | 0.495 | 0.009 | 0.083 | 0.009 | 0.046 | 0.092 | 0.266 | 0.018 | 0.183 | 0.083 | 0.028 | 0.138 | 0.028 | 0.037 | 0.009 |

|             |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| rCRS_D-loop | 16290 | 16288 | 16278 | 16275 | 16274 | 16272 | 16269 | 16266 | 16262 | 16261 | 16260 | 16259 | 16257 | 16256 | 16252 | 16249 | 16245 | 16242 |
| Hmong       | 0.075 | 0.075 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.132 | 0.038 | 0.132 | 0.132 | 0.000 | 0.057 | 0.019 | 0.019 | 0.000 | 0.000 | 0.000 |
| Lao Song    | 0.000 | 0.138 | 0.034 | 0.000 | 0.103 | 0.000 | 0.034 | 0.103 | 0.000 | 0.069 | 0.000 | 0.138 | 0.034 | 0.000 | 0.000 | 0.172 | 0.034 | 0.034 |
| Thai -Siam  | 0.037 | 0.000 | 0.000 | 0.037 | 0.037 | 0.074 | 0.000 | 0.037 | 0.000 | 0.037 | 0.000 | 0.000 | 0.037 | 0.037 | 0.000 | 0.111 | 0.074 | 0.000 |
| overall     | 0.046 | 0.073 | 0.009 | 0.009 | 0.037 | 0.018 | 0.009 | 0.101 | 0.018 | 0.092 | 0.064 | 0.037 | 0.046 | 0.018 | 0.009 | 0.073 | 0.028 | 0.009 |

Table A4 (cont.)

|             |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| rCRS_D-loop | 16354 | 16343 | 16335 | 16327 | 16325 | 16320 | 16319 | 16318 | 16311 | 16309 | 16304 | 16302 | 16301 | 16300 | 16298 | 16297 | 16292 | 16291 |
| Hmong       | 0.019 | 0.000 | 0.019 | 0.094 | 0.000 | 0.000 | 0.019 | 0.000 | 0.151 | 0.057 | 0.208 | 0.000 | 0.000 | 0.075 | 0.245 | 0.000 | 0.000 | 0.019 |
| Lao Song    | 0.000 | 0.000 | 0.000 | 0.034 | 0.034 | 0.000 | 0.000 | 0.000 | 0.103 | 0.034 | 0.241 | 0.000 | 0.172 | 0.034 | 0.034 | 0.241 | 0.069 | 0.000 |
| Thai -Siam  | 0.111 | 0.000 | 0.037 | 0.000 | 0.000 | 0.037 | 0.037 | 0.037 | 0.111 | 0.000 | 0.296 | 0.037 | 0.000 | 0.000 | 0.000 | 0.037 | 0.037 | 0.111 |
| overall     | 0.037 | 0.009 | 0.018 | 0.055 | 0.009 | 0.009 | 0.018 | 0.009 | 0.128 | 0.037 | 0.239 | 0.009 | 0.046 | 0.046 | 0.128 | 0.073 | 0.028 | 0.037 |
| rCRS_D-loop | 183   | 182   | 177   | 173   | 153   | 152   | 151   | 150   | 146   | 106   | 103   | 94    | 93    | 73    | 16362 | 16357 | 16356 | 16355 |
| Hmong       | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.189 | 0.057 | 0.019 | 0.000 | 0.057 | 0.000 | 1.000 | 0.302 | 0.075 | 0.038 | 0.132 |
| Lao Song    | 0.034 | 0.000 | 0.000 | 0.000 | 0.000 | 0.034 | 0.000 | 0.345 | 0.034 | 0.000 | 0.000 | 0.034 | 0.034 | 1.000 | 0.207 | 0.034 | 0.034 | 0.000 |
| Thai -Siam  | 0.000 | 0.074 | 0.074 | 0.037 | 0.074 | 0.370 | 0.037 | 0.185 | 0.111 | 0.000 | 0.037 | 0.037 | 0.000 | 1.000 | 0.370 | 0.000 | 0.037 | 0.000 |
| overall     | 0.009 | 0.018 | 0.018 | 0.009 | 0.018 | 0.101 | 0.009 | 0.229 | 0.064 | 0.009 | 0.009 | 0.046 | 0.009 | 1.000 | 0.294 | 0.046 | 0.037 | 0.064 |
| rCRS_D-loop | 279   | 263   | 249   | 239   | 236   | 235   | 228   | 225   | 215   | 210   | 207   | 204   | 199   | 195   | 194   | 189   | 186   | 185   |
| Hmong       | 0.038 | 0.981 | 0.396 | 0.000 | 0.019 | 0.000 | 0.000 | 0.000 | 0.057 | 0.113 | 0.170 | 0.132 | 0.000 | 0.113 | 0.019 | 0.075 | 0.000 | 0.075 |
| Lao Song    | 0.000 | 1.000 | 0.103 | 0.034 | 0.000 | 0.034 | 0.000 | 0.000 | 0.000 | 0.069 | 0.000 | 0.034 | 0.276 | 0.172 | 0.000 | 0.000 | 0.000 | 0.000 |
| Thai -Siam  | 0.000 | 1.000 | 0.185 | 0.000 | 0.000 | 0.037 | 0.037 | 0.037 | 0.000 | 0.074 | 0.037 | 0.000 | 0.074 | 0.074 | 0.000 | 0.037 | 0.037 | 0.037 |
| overall     | 0.018 | 0.991 | 0.266 | 0.009 | 0.009 | 0.018 | 0.009 | 0.009 | 0.028 | 0.092 | 0.092 | 0.073 | 0.092 | 0.119 | 0.009 | 0.046 | 0.009 | 0.046 |

Table A4 (cont.)

|             |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| rCRS_D-loop | 281   | 298   | 311   | 316   | 317   | 320   | 324   | 329   | 330   | 332   | 445   | 456   | 461   | 466   | 479   | 482   | 485   | 489   |
| Hmong       | 0.000 | 0.000 | 0.019 | 0.000 | 0.019 | 0.038 | 0.075 | 0.000 | 0.019 | 0.000 | 0.000 | 0.019 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.396 |
| Lao Song    | 0.000 | 0.103 | 0.034 | 0.000 | 0.000 | 0.000 | 0.103 | 0.138 | 0.000 | 0.103 | 0.000 | 0.000 | 0.034 | 0.069 | 0.000 | 0.000 | 0.000 | 0.621 |
| Thai -Siam  | 0.037 | 0.000 | 0.000 | 0.037 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.037 | 0.000 | 0.037 | 0.000 | 0.037 | 0.037 | 0.037 | 0.519 |
| overall     | 0.009 | 0.028 | 0.018 | 0.009 | 0.009 | 0.018 | 0.064 | 0.037 | 0.009 | 0.028 | 0.009 | 0.009 | 0.018 | 0.018 | 0.009 | 0.009 | 0.009 | 0.486 |

|             |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| rCRS_D-loop | 499   | 500   | 512   | 513   | 514   | 516   | 518   | 523   | 524   | 525   | 527   | 539   | 548   | 549   |
| Hmong       | 0.038 | 0.019 | 0.000 | 0.000 | 0.019 | 0.038 | 0.019 | 0.491 | 0.491 | 0.491 | 0.019 | 0.000 | 0.019 | 0.000 |
| Lao Song    | 0.000 | 0.000 | 0.138 | 0.138 | 0.034 | 0.000 | 0.000 | 0.448 | 0.448 | 0.448 | 0.000 | 0.000 | 0.000 | 0.000 |
| Thai -Siam  | 0.037 | 0.000 | 0.000 | 0.000 | 0.000 | 0.037 | 0.000 | 0.185 | 0.148 | 0.148 | 0.000 | 0.037 | 0.000 | 0.037 |
| overall     | 0.028 | 0.009 | 0.037 | 0.037 | 0.018 | 0.028 | 0.009 | 0.404 | 0.394 | 0.394 | 0.009 | 0.009 | 0.009 | 0.009 |