

**DETECTION OF 9-BP DELETION IN
COII/tRNA^{Lys} INTERGENIC REGION OF
MITOCHONDRIAL DNA AMONG MELANAU
ETHNIC GROUP**

by

NUR SYAQIRAH BINTI AHMAD HASANUDDIN

**Dissertation submitted in fulfilment of the requirements
for the degree of
Bachelor of Science (Honours) (Forensic Science)**

February 2026

ACNOWLEDGEMENT

First and foremost, all praise to the Almighty for granting me the opportunity, strength and guidance to complete this research. Alhamdulillah for providing me with clarity and perseverance throughout this journey.

I would like to express my sincere appreciation to my supervisor, Dr Nur Haslindawaty binti Abd Rashid for giving me a chance to undertake this project. I am deeply grateful for her invaluable guidance, continuous support and patiently teaching me throughout this research. I would also like to extend my sincere gratitude to all the lecturers and academic staff of the Bachelor of Forensic Science programme for their guidance and knowledge shared throughout my academic journey.

A special thanks goes to Saiful Ammar for his guidance throughout my experimental work and for his valuable advice. My appreciation is also extended to the laboratory staff and technical officers for their assistance and cooperation during the experimental procedures. I am truly grateful to the individuals who voluntarily participated in this study and contributed the blood samples. Without their participation, this research would not have been possible. Special thanks also extended to my friends and fellow students for their encouragement and assistance throughout the completion of this project.

Finally, I would like to express my heartfelt gratitude to my family for their continuous support, understanding and encouragement throughout my studies. Their unwavering support and motivation, along with the guidance of my supervisor and others who were fully or partially involved in this journey, provided the strength and inspiration that were necessary for me to complete this research successfully.

TABLE OF CONTENTS

CERTIFICATE.....	i
DECLARATION.....	iii
ACNOWLEDGEMENT	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	vii
LIST OF TABLES	viii
LIST OF ABBREVIATIONS	ix
LIST OF SYMBOLS	xi
LIST OF APPENDICES	xii
ABSTRAK.....	xiii
ABSTRACT.....	xiv
CHAPTER 1	1
INTRODUCTION	1
1.1 Introduction.....	1
1.2 Problem Statement	3
1.3 Significance of the Study	3
1.4 Objectives	5
1.4.1 Main Objective.....	5
1.4.2 Specific Objectives	5
CHAPTER 2	6
2.1 Mitochondrial DNA (mtDNA).....	6
2.1.1 Mitochondrial DNA (mtDNA) polymorphism	9
2.1.2 The 9-bp Deletion inCOII/tRNALys Intergenic Region.....	11
2.1.3 Application of mtDNA Variants in Genetic Studies.....	13
2.2 Melanau Ethnic Group	14
2.3 Haplogroup B.....	16
2.4 Polymerase Chain Reaction (PCR).....	17
CHAPTER 3	20
3.1 Materials	20
3.1.1 Chemicals and Reagents	20
3.1.2 Instruments and Apparatus.....	21
3.1.3 Chemical and Reagent Preparation	21

3.1.3.1 Preparation of 10X Tris-Borate-EDTA (TBE)	21
3.1.3.2 Preparation of 0.5X Tris-Borate-EDTA (TBE) Buffer.....	22
3.2 Methodology	22
3.2.1 Sampling Method	22
3.2.2 Sample Collection	23
3.2.3 Genomic DNA Extraction	23
3.2.4 Agarose Gel Electrophoresis of Genomic DNA	25
3.2.5 DNA Quantification	26
3.2.6 Polymerase Chain Reaction (PCR) Amplification.....	27
3.2.7 Agarose Gel Electrophoresis of PCR Product	29
CHAPTER 4	30
4.1 Genomic DNA Extraction	30
4.1.1 Agarose Gel Electrophoresis	30
4.1.2 DNA Quantification	31
4.2 PCR Amplification of mtDNA COII/tRNA ^{Lys} Intergenic Region	32
CHAPTER 5	36
5.1 Genomic DNA Extraction	36
5.1.1 Agarose Gel Electrophoresis of Extracted Genomic DNA	36
5.1.2 DNA quantification	37
5.2 PCR Amplification of COII/tRNA ^{Lys} Intergenic Region	39
CHAPTER 6	41
6.1 Conclusion	41
6.2 Limitations and Future Study.....	42
REFERENCES	43
APPENDICES	51

LIST OF FIGURES

		Page
Figure 2.1	Human mtDNA is double-stranded circular molecule encoded for 37 genes (Penman <i>et al.</i> , 2020) consisting of 13 protein-coding genes (yellow, blue and pink), 2 rRNAs (light blue) and 22 tRNAs (grey).	7
Figure 2.2	Location of the two copies of a 9-bp sequence between COII and tRNA ^{Lys} in the intergenic region in mtDNA (Alcántar-Aguirre <i>et al.</i> , 2025).	12
Figure 2.3	Map of traditional Melanau settlement areas (Rensch, 2012).	15
Figure 2.4	Schematic presentation of the polymerase chain reaction (PCR) principle (Garibyan & Avashia, 2013).	18
Figure 4.1	The extracted genomic DNA from unrelated Melanau individuals. Lane 1: 1 Kbp DNA ladder and Lane M049 - M068: Extracted genomic DNA from unrelated Melanau individuals.	30
Figure 4.2	Figure 4.2: Representative results of PCR amplicon visualised using 3% agarose gel electrophoresis. Lane 100 bp: 100 bp DNA Ladder; Lane NC: negative control sample; Lane M071-M090: Amplified PCR product of Melanau individuals.	33
Figure 4.3	Distribution of the 9-bp deletion among Melanau individuals based on PCR amplification results of the COII/tRNA ^{Lys} region of mtDNA.	34
Figure 4.4	Frequencies of haplogroup B in Melanau with other populations (Schurr & Wallace, 2002).	35

LIST OF TABLES

		Page
Table 2.1	Haplogroup B frequency in Southeast Asia (Schurr & Wallace, 2002).	17
Table 3.1	List of chemicals, reagents and consumables used.	20
Table 3.2	List of instruments and apparatus used in this research.	21
Table 3.3	Primer used for amplification of the COII/tRNA ^{Lys} intergenic region.	27
Table 3.4	Quantities of reagent in PCR mastermix.	28
Table 3.5	PCR cycling parameters for Veriti 96 Well Thermal Cycler.	28
Table 4.1	Representative results of quantification on 19 samples using UV-visible spectrophotometer (DeNovix DS-11, USA).	31

LIST OF ABBREVIATIONS

A	Adenine
aDNA	Ancient DNA
ATP	Adenosine triphosphate
bp	Base pair
C	Cytosine
COII	Cytochrome C Oxidase Subunit II
ddH ₂ O	Deionized Distilled Water
D-loop	Displacement Loop
DNA	Deoxyribonucleic Acid
dsDNA	Double Stranded DNA
EDTA	Ethylenediaminetetraacetic Acid
EtBR	Ethidium Bromide
ETC	Electro Transport Chain
G	Guanine
gDNA	Genomic DNA
HCC	Hepatocellular Carcinoma
[HG]	Haplogroup
H strand	Heavy strand
Kbp	Kilobase pair
L strand	Light strand
MELAS	Mitochondrial Encephalomyopathy, Lactic Acidosis, and Stroke-like Episodes
MERRF	Myoclonic epilepsy with Ragged-Red Fibers
mtDNA	Mitochondrial DNA
nDNA	Nuclear DNA
OriH	Origin of Replication for H strand
OriL	Origin of Replication for Light strand
OXPHOS	Oxidative Phosphorylation
PCR	Polymerase Chain Reaction
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
rRNA	Ribosomal RNA

SNP	Single Nucleotide Polymorphism
ssDNA	Single-stranded DNA
SSM	Slipped Strand Mismatching
T	Thymine
TBE	Tris-Borate-EDTA
T _m	Melting Temperature
tRNA	Transfer RNA
tRNA ^{Lys}	Lysine tRNA
UV	Ultraviolet
UV-Vis	Ultraviolet-Visible
YBP	Years Before Present

LIST OF SYMBOLS

>	Greater than
%	Percentage
°C	Degree Celcius
A	Absorbance
g	Gram
L	Litre
min	Minutes
mL	Millilitre
μL	Microlitre
n	Number of samples
ng	Nanogram
nm	Nanometre
pmol	Picomoles
V	Volt
w/v	Weight/Volume

LIST OF APPENDICES

Appendix A	JEPeM-USM Ethical Approval
------------	----------------------------

**PENGESANAN PEMADAMAN PADA DNA MITOKONDRIA DI KAWASAN
INTERGENIK COII/tRNALys DALAM KALANGAN KUMPULAN ETNIK
MELANAU**

ABSTRAK

Mitokondria merupakan organel penting dalam organisma yang bertanggungjawab menghasilkan tenaga sel untuk reaksi metabolik. Mitokondria mempunyai informasi genetiknya sendiri yang dikenali sebagai DNA mitokondria (mtDNA), mempunyai ciri-ciri seperti warisan daripada ibu, kadar mutasi yang tinggi, nombor salinan yang tinggi dan kurang penggabungan semula. MtDNA telah digunakan dalam pengenalan forensik untuk menyelesaikan kes forensik dan juga digunakan dalam kajian populasi genetik. Salah satu DNA mitokondria polimorfisme yang digunakan dalam kajian populasi genetik terletak di kawasan antara gen cytochrome c oxidase II (COII) dan gen lysine tRNA (tRNALys) yang mempunyai dua salinan 9 pasangan asas yang berulang (CCCCCTCTA). Pemotongan satu salinan merupakan salah satu ciri untuk DNA mitokondria kumpulan haplo B. Kajian ini bertujuan untuk mengenal pasti pemotongan 9 pasangan asas di kawasan intergenik di antara COII dan tRNALys dalam mitokondria di kalangan individu Melanau melalui amplifikasi PCR dan elektroforesis gel agarosa. DNA genomik telah diekstrak daripada sampel darah yang telah dinilai dari segi kualiti dan kuantiti dengan memastikan DNA mencukupi untuk analisis seterusnya. Amplifikasi PCR untuk kawasan sasaran telah dijalankan menggunakan primer 8215F dan 8385R. Keputusan telah menunjukkan bahawa 24 individu Melanau (20%) mempunyai pemotongan 9 pasangan asas dan pengagihan ini telah di sokong oleh kajian sebelum ini yang menyatakan bahawa pemotongan 9 pasangan asas ini telah kerap dikenal pasti dalam populasi yang terletak di kawasan tanah besar Asia Tenggara, Pulau Asia Tenggara dan Asia Barat Daya.

DETECTION OF 9-BP DELETION IN COII/tRNALys INTERGENIC REGION OF MITOCHONDRIAL DNA AMONG MELANAU ETHNIC GROUP

ABSTRACT

Mitochondria are essential organelles in organism responsible for cellular energy production, which is necessary for metabolic reactions. It has its own genetic material called mitochondrial DNA (mtDNA) and has interesting features such as maternal inheritance, high mutation rate, high copy number and lack of recombination. MtDNA has been applied in forensic identification to solve forensic cases and is also used in population genetics studies. One of the mtDNA polymorphisms used in population genetics studies is located between cytochrome c oxidase II (COII) and lysine tRNA (tRNALys) genes, consisting of the two copies of a 9-bp tandem repeat sequence (CCCCCTCTA). Deletion of one copy in this tandem is used as a characteristic for mtDNA haplogroup B. This study aimed to detect the presence of the 9-bp deletion in the COII/tRNALys intergenic region of mitochondrial DNA among Melanau individuals through PCR amplification and agarose gel electrophoresis. Genomic DNA was extracted from the whole blood samples and assessed for quality and quantity to ensure sufficient DNA for downstream analysis. PCR amplification of the target region was performed using primers 8215F and 8385R. The findings show that 24 Melanau individuals (20%) had 9-bp deletion, and this distribution is supported by the previous studies that 9-bp deletion is frequently observed in the population of mainland southeastern Asia, Island southeast Asia and Southwest Asian.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Mitochondria are intracellular organelles found in eukaryotic cells and are known as the powerhouses of the cell, which are involved production of energy (Habbane *et al.*, 2021). Mitochondrial DNA (mtDNA) located within the mitochondria, is a small, circular and double-stranded molecule. In humans, mtDNA represents about 1% of total cellular DNA with an estimated 1000 to 10,000 copies per cell (Habbane *et al.*, 2021). Most of the mitochondrial proteins are encoded by nuclear genes, except for 13 of the respiratory proteins, 22 transfer RNAs (tRNAs) and 2 ribosomal RNAs (rRNAs), which are encoded by the mitochondrial genome (Giuliano *et al.*, 2024). Unlike nuclear DNA (nDNA), mitochondrial genes lack introns, highly compact and contain very few non-coding regions. The only non-coding region is the control region, also known as the displacement loop (D-loop). This region plays a crucial role in regulating mtDNA replication and transcription.

Mitochondrial DNA differs from nDNA in its mode of inheritance. MtDNA is one of the lineage markers, and it is inherited exclusively from the maternal line, the mother passes her mtDNA to all children, and only her daughters will pass their genetic code to the next generations. This inheritance pattern is due to the high copy number of mtDNA inside the oocyte of females and the fact that mitochondria in the intermediate region of the sperm are destroyed during the first cell divisions (Habbane *et al.*, 2021). Additionally, mtDNA mutation rate has been reported to be 10 to 100 times higher than the nDNA (Gupta *et al.*, 2023). The inheritance that occurs from generation to generation in a cluster of single-nucleotide

polymorphisms (SNPs) in mtDNA creates a different pattern of haplotypes, which is useful for defining geographic genetic populations (Giuliano *et al.*, 2024).

In addition to the control region, several intergenic regions within the mitochondrial genome display population-specific variability. One such region is the COII/tRNA^{Lys} intergenic region, located between the cytochrome c oxidase subunit II (COII) and the lysine tRNA genes (tRNA^{Lys}). Most human mitochondrial genomes contain a 9-base pair (bp) tandem duplication within this region (Cann and Wilson 1983; Wrischnik *et al.* 1987). However, the occurrence of 9-bp deletions found in this region has attracted attention to be further studied since it showed variation in human mtDNA (Wrischnik *et al.*, 1987). The deletion of one copy of a duplicated sequence occurs in certain populations and has been widely used to infer population affinities and human migration patterns (Betty *et al.*, 1996). Although the 9-bp deletion was initially thought to be unique to individuals of Asian ancestry (Ballinger *et al.*, 1992), subsequent research has identified its presence in other populations as well.

Mitochondrial DNA (mtDNA) has been widely used as a genetic tool for investigating evolutionary relationships among human populations. Due to its distinctive mode of maternal inheritance, mtDNA variation accumulates through the sequential addition of mutations along maternal lineages. This provides valuable insights into human population history and migration patterns. These variations serve as key molecular markers for the identification and classification of mitochondrial haplogroups. In this context, haplogroup B is defined by the presence of a characteristic 9-bp deletion in the COII/tRNA^{Lys} intergenic region of the mitochondrial genome, as described by (Derenko *et al.*, 2012). This 9-bp deletion is most frequently observed among Austronesian populations of the Pacific and Austroasiatic populations of Southeast Asia, suggesting a shared common ancestry within haplogroup B

(Thangaraj *et al.*, 2005). Therefore, studying the distribution of the 9-bp deletion among the Melanau can provide important insights into their maternal genetic ancestry and migration patterns, especially involving the aborigine groups of Sarawak people.

1.2 Problem Statement

Although the occurrence of 9-bp deletion in the COII/tRNALys intergenic region of mitochondrial DNA (mtDNA) is not exclusively restricted to Asian populations and may have arisen independently multiple times within Asia. It serves as an important genetic marker due to the high-level of variation present in mtDNA among individuals. As previously reported, the mitochondrial DNA has been widely applied in human migration and population studies (Al-Zahery *et al.*, 2003; Kashyap *et al.*, 2003). Despite its extensive use in population genetic research, no such study has been conducted in this region focusing on the Melanau individual, suggesting a lack of 9-bp deletion data among the indigenous ethnic groups of Malaysia. Considering the significance of this mtDNA polymorphism, it may provide valuable insight into population structure and genetic relationship within Malaysia. Hence, this study aims to fill the research gap by examining the presence of the 9-bp deletion in the COII/tRNALys intergenic region of mtDNA among the Melanau ethnic group.

1.3 Significance of the Study

The study of 9-bp deletions in the COII/tRNALys intergenic region of mtDNA is significant for advancing the understanding of mtDNA variation within the Melanau population. This deletion is widely used as a genetic marker defining mitochondrial haplogroup B, which occurs at high frequency in Southeast Asia and is broadly distributed among Southeast Asian, Austronesian and Pacific populations (Derenko *et al.*, 2012). The 9-bp deletion has been used as an important marker for the classification of the haplogroup. Derenko *et al.*, (2012) have

demonstrated that the combination of the 9-bp deletion with a mutation at nucleotide position 16189 at the mtDNA hypervariable region, defined a major Asian lineage of haplogroup B. This lineage forms a largely monophyletic group that includes the sub-haplogroups B4 and B5, which are commonly found in East Asia, Southeast Asia, and Island Southeast Asia (Derenko *et al.*, 2012). This phylogenetic pattern is further supported by Thangaraj *et al.* (2005), who also reported the association of the 9-bp deletion in Asian-specific mitochondrial lineages.

Other than that, the recurrent 9-bp deletion has also been reported across diverse global populations and has shown significant association with a wide range of disease conditions. Based on the previous studies involving Taiwanese families have shown this polymorphism is associated with mitochondrial disorders such as mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS), myoclonic epilepsy with ragged-red fibres (MERRF) syndromes (Giuliano *et al.*, 2024). It has also been discovered in patients with heterozygous polycystic ovary syndrome disorder at a higher rate compared to healthy individuals (Zhuo *et al.*, 2010) and significantly higher prevalence among patients with hepatocellular carcinoma (HCC) (Jin *et al.*, 2012). The proximity of this deletion to the tRNA^{Lys} gene is of particular interest, as this gene plays an essential role in mitochondrial protein synthesis. Although this deletion was considered a non-pathogenic polymorphism, the variations that occur near tRNA genes may influence mitochondrial gene expression and energy production, thereby contributing to disease susceptibility. This may suggest that the 9-bp deletion can serve as a predictive marker and play an important role in the future.

1.4 Objectives

1.4.1 Main Objective

To identify a 9-bp deletion in the COII/tRNALys intergenic region of mtDNA among the Melanau ethnic group.

1.4.2 Specific Objectives

- I. To PCR amplify COII/tRNALys intergenic region in the Melanau ethnic group.
- II. To determine the 9-bp deletion in the Melanau ethnic group.
- III. To compare the frequency of 9-bp deletion in the Melanau ethnic group with other populations.

CHAPTER 2

LITERATURE REVIEW

2.1 Mitochondrial DNA (mtDNA)

Mitochondria are semi-autonomously reproductive organelles within eukaryotic cells carrying their own genetic material, called the mitochondrial genome (mtDNA) (Crimi & Rigolio, 2008). The mitochondrial genome is a complete set of genetic material of a mitochondrion, whereas mitochondrial DNA (mtDNA) constitutes only a small fraction of total cellular DNA. Human mtDNA is a double-stranded, circular molecule approximately 16,569 base pairs in length, and unlike the nuclear genome it has no introns (Habbane *et al.*, 2021) (Figure 2.1). The two strands of mtDNA are classified into the heavy strand (H strand) and light strand (L strand). The H strand is rich in purine bases, which consist of guanine (G) and adenine (A) whereas the L strand, rich in pyrimidine bases consists of cytosine (C) and thymine (T). Most of the information found in the H strand (Habbane *et al.*, 2021).

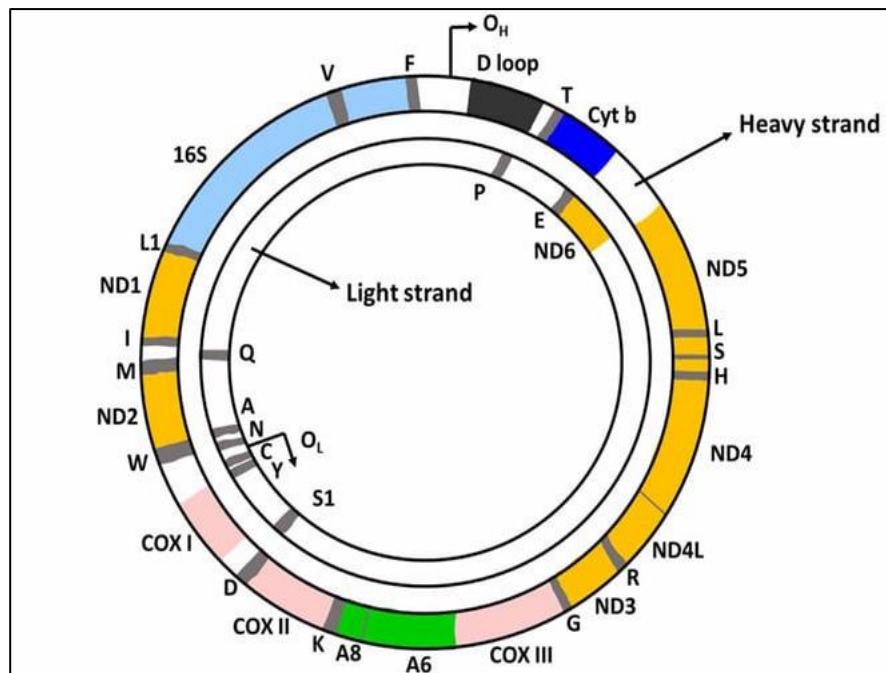


Figure 2.1: Human mtDNA is double-stranded circular molecule encoded for 37 genes (Penman *et al.*, 2020) consisting of 13 protein-coding genes (yellow, blue and pink), 2 rRNAs (light blue) and 22 tRNAs (grey).

The mitochondrial genome encodes 37 genes, including 13 protein-coding genes that are essential components of the electron transport chain (ETC) involved in oxidative phosphorylation (OXPHOS). These include seven genes encoding for subunits within complex I (MT-ND1, MT-ND2, MT-ND3, MT-ND4, MT-ND4L, MT-ND5, MT-ND6), one for complex III (MT-CYB), three for complex IV (MT-CO1, MT-CO2, MT-CO3) and two for complex V (MT-ATP6, MT-ATP8), in addition to two ribosomal RNAs (MT-RNR1 and MT-RNR2) and 22 transfer RNAs (Betty *et al.*, 1996). These genes are essential for normal mitochondrial function and mitochondrial protein synthesis.

Mitochondrial DNA is highly compact and contains very few non-coding regions. The largest non-coding segment is the displacement loop (D-loop), also referred to as the control region, which spans approximately 1.1 Kbp of the mtDNA genome (Giuliano *et al.*, 2024). This region

plays a central role in mtDNA replication as it contains the origin of replication for the H strand (OriH) (Anderson *et al.*, 1981) and promoters for transcription of both the H and L strands (Cantatore & Attardi, 1980). Due to its non-coding nature and involvement in replication processes, it accumulates mutations more rapidly than coding regions, contributing to its high variability. Aside from the D-loop, the remaining non-coding sequences of mtDNA are minimal and limited to short intergenic sequences that often consist of few or non-coding bases between them (Giuliano *et al.*, 2024). This region can be characterized as a tiny gap between genes and these genes are directly adjacent or even overlapping. Another small non-coding feature is the origin of light strand replication (OriL), which also contributes to the regulation of mtDNA replication.

One of the defining features of mtDNA, it is inherited exclusively through the maternal lineage. Although both sperm and oocytes contain mitochondria, paternal mitochondria are typically eliminated shortly after fertilisation through enzymatic degradation processes (Habbane *et al.*, 2021). Consequently, mtDNA is transmitted vertically from mothers to offspring without contribution from the paternal genome. If the paternal mitochondria are not eliminated and kept in the oocyte, it may lead to mitochondrial disease (Habbane *et al.*, 2021). Additionally, mtDNA is present in exceptionally high copy numbers within cells compared to nuclear DNA. In humans, there can be over 1,000 mitochondria per cell. Each mitochondrion harbours approximately 5 to 10 copies of mtDNA (Spelbrink, 2010), resulting in an estimated 1,000 to 10,000 copies per cell. This high copy number enhances the likelihood of successful DNA recovery from degraded, ancient or low quality biological samples.

Furthermore, the mutation rate of mtDNA is substantially higher than that of nuclear DNA, exceeding by at least an order of magnitude (Kivisild, 2015). The mutation rate is 10 to 20 times higher than the nuclear genome (Árnadóttir *et al.*, 2024; Habbane *et al.*, 2021). This elevated mutation rate and the presence of multiple copies of mtDNA within a cell, allow for the coexistence of more than one mtDNA variant in a single individual. When mutant and wild-type mtDNA molecules are present simultaneously, the condition is referred to as heteroplasmy. In contrast, when only a single mtDNA variant is present, the condition is known as homoplasmy (Habbane *et al.*, 2021).

2.1.1 Mitochondrial DNA (mtDNA) polymorphism

Mitochondrial DNA (mtDNA) exhibits a high degree of genetic variation. This phenomenon results from several biological factors that contributed to the sequence variation. One of the major factors is the central role of mitochondria in oxidative phosphorylation, which contribute from biochemical pathway that continuously generates reactive oxygen species (ROS) creates mtDNA more vulnerable to both exogenous and endogenous exposure that can cause mtDNA damage (King & Copeland, 2025). This vulnerability is particularly pronounced during replication, as mtDNA remains transiently single-stranded, increasing its sensitivity to radical attacks (Habbane *et al.*, 2021). Besides that, the lack of histone protein and higher-order chromatin structures that protect nuclear DNA also makes more exposed to oxidative insults. It has also been reported that mtDNA lacks DNA repair capacity compared to nuclear DNA resulting in less efficiency in correcting replication errors and oxidative DNA lesions (King & Copeland, 2025).

These unique conditions collectively generate a substantial number of point mutations, insertions, deletions and other sequence alterations resulting in extensive mtDNA polymorphism across human populations. The resulting polymorphisms in mtDNA can be broadly classified into several types reflecting distinct mutational events and evolutionary processes. One of the mtDNA polymorphisms includes single nucleotide polymorphisms (SNPs), which represent the most common polymorphism in the human genome (Sukhumsirichart, 2018). This variation involves single base change in a DNA sequence, with a normal alternative of two possible nucleotides at a given position (Sukhumsirichart, 2018). SNPs can be present in various regions of the human genome, including coding regions (exons), non-coding regions (introns) and upstream regions of genes (promoters) (Vallente, 2024). Another form of mtDNA variation is length polymorphism, primarily due to tandem duplication or deletion of short DNA motifs within the hypervariable segments of the non-coding region. This condition often results from the replication slippage within the homopolymeric tracts and tandem repeats.

Since the mtDNA is polymorphic, these variations are often associated with alterations in mitochondrial functions (Russlies *et al.*, 2019). Human mtDNA variations can be categorized into the following three groups: (1) recent maternally inherited deleterious mutations; (2) ancient adaptive polymorphism; and (3) somatic mutations that accumulate during development and tissue with age (Russlies *et al.*, 2019). Mutations in the mtDNA are associated with a wide spectrum of human disease (Gupta *et al.*, 2023), which are heterogeneous and can appear at any age (from antenatal to advanced life) and affect all tissues (Habbane *et al.*, 2021). The disease mainly affects the brain, heart, skeletal muscle, liver and endocrine systems (Habbane *et al.*, 2021), which contribute to diseases like Leber's hereditary optic neuropathy (LHON), mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes

(MELAS), myoclonic epilepsy with ragged-red fibres (MERRF) and neuropathy, ataxia, and retinitis pigmentosa (NARP) (Schon, DiMauro, & Hirano, 2012).

Since it is maternal mode inheritance and lack of recombination in mtDNA molecules (Bermisheva *et al.*, 2003), these limitations are passed clonally from mother to their offspring. Collectively, these molecular features highlight the dynamic and highly mutable nature of mtDNA, underscoring its significance in evolutionary research, maternal lineage tracing and forensic DNA analysis.

2.1.2 The 9-bp Deletion in COII/tRNALys Intergenic Region

The 9-bp deletion in the COII/tRNALys intergenic region of mitochondrial DNA (mtDNA) was first described by Wrischnick *et al.* (1987) and become the most studied structural polymorphism within the human mitochondrial genome (Thomas *et al.*, 1998). This condition typically involves the loss of one of two copies of a 9-bp tandem repeat sequence (CCCCCTCTA) from 2 non-coding region located between the cytochrome c oxidase subunit II (COII) gene and the lysine transfer RNA (tRNALys) genes (Figure 2.2). In the human mtDNA, the COII gene terminates at nucleotide position 8267, while the tRNALys gene initiates at position 8295 (Alcántar-Aguirre *et al.*, 2025). The two tandem repeats 9-bp sequence located at nucleotide position 8271-828 of the mtDNA coding region (Anderson *et al.*, 1981). The commonly reported location of the 9-bp deletion between COII and tRNALys occurs between positions 8281- 8289 within this region.

Initially, it was identified in the East Asian populations (Wrischnick *et al.* 1987), including the Oceanian and Sub-Saharan African populations, but rarely reported in Europe (Giuliano *et al.*, 2024). In addition to deletion, other length variants within this region, such as triplication of

the same 9-bp motif have been observed in populations of Chukchi from Siberia (Shields *et al.*, 1992) and Tharu from Nepal (Passarino *et al.*, 1993).

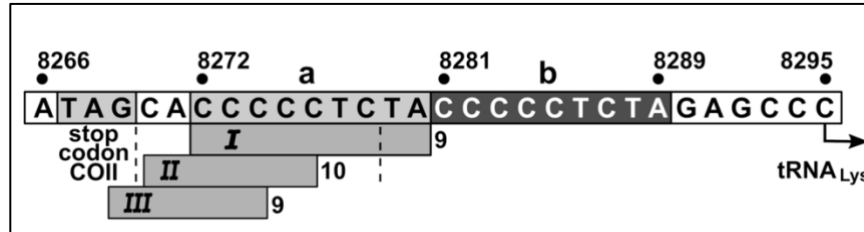


Figure 2.2: Location of the two copies of a 9-bp sequence between COII and tRNA^{Lys} in the intergenic region in mtDNA (Alcántar-Aguirre *et al.*, 2025).

Although this intergenic region does not encode protein or functional RNA molecules, the occurrence of these and other length polymorphisms suggests the existence of sequence length instability in this mtDNA region (Wrischnik *et al.* 1987; Ballinger *et al.* 1992; Redd *et al.* 1995). The repetitive nature of the 9-bp motif could be the result of the slippage-strand mispairing (SSM) during DNA replication (Levinson & Gutman 1987; Richards & Sutherland 1994), predisposing it to length variation and deletion events. The maternal inheritance of mtDNA allows clusters of single nucleotide polymorphism (SNPs) to accumulate over generations, giving rise to distinct haplotypes that are useful for defining geographic genetic populations (Giuliano *et al.*, 2024). Consequently, analysis of the hypervariable segments of the noncoding regions of mtDNA from individuals with the 9-bp deletion provides information on the evolution of this marker and the history of populations carrying (Horai & Hayasaka, 1990).

2.1.3 Application of mtDNA Variants in Genetic Studies

The mitochondrial genome (mtDNA) has become an essential genetic component for eukaryotic species classification and evolutionary studies (Elyasigorji *et al.*, 2023). Early research demonstrated that mtDNA provides crucial genetic information for elucidating phylogenetic relationships among species (Patwardha *et al.*, 2014) due to its conserved structure and mode of inheritance. The unique features of mtDNA make it a powerful marker for tracing maternal lineages and ancestry origins.

Studies of mtDNA variation have played a pivotal role in reconstructing human evolutionary history. For example, analyses of mtDNA lineage enabled population geneticists to trace the ancestors of modern humans back to Africa and to infer subsequent global dispersal (Forster & Matsumura, 2005). Genetic data indicate that Asia and Australia were among the earliest regions settled by modern humans outside Africa, whereas Europe was colonised later, approximately 45,000 years ago (Nesheva, 2014).

In addition, mtDNA variation is widely used to classify the mtDNA variant into haplogroups (HG) as a fundamental approach in population genetics (Kim *et al.*, 2022). Haplogroups are defined by shared sets of mtDNA polymorphisms, particularly single nucleotide polymorphisms (SNPs), that are inherited from a common maternal ancestor. The most recent common ancestor of all modern humans belongs to haplogroup L, which is predominantly found in Africa. Haplogroup L3 gave rise to the M and N macrohaplogroups, from which all non-African mtDNA lineages are derived. Haplogroup L3 gave rise to the M and N subclades. Asian haplogroups developed from both the N subclade (A, B, F, and Y) and the M subclade (C, D, G, Q, and Z) (Urzua-Traslavina *et al.*, 2014). The distribution of these haplogroups reflects population differentiation and independent evolutionary histories during recent human evolution (Soares *et al.*, 2012).

Moreover, mtDNA variation has been extensively used in reconstructing population history and demographic events, such as bottlenecks, founder effects and admixture through analyses of haplotype and phylogenetic relationships. The global distribution of mtDNA haplogroups is well documented across modern populations, providing a valuable reference framework for ancient DNA (aDNA) studies (Salamon *et al.*, 2024).

2.2 Melanau Ethnic Group

Malaysia is a multi-ethnic and multicultural nation comprising a wide range of indigenous and non-indigenous populations. In Peninsular Malaysia, the indigenous Orang Asli are divided into three main groups, namely the Negrito, Senoi and Proto-Malay (Department of Statistics Malaysia, 2011). In East Malaysia, Sabah has 32 ethnic groups, predominantly Kadazandusun, while Sarawak comprises 27 ethnic groups, and the Iban contribute the largest group (Department of Statistics Malaysia, 2011; Yeo *et al.*, 2024). In Sarawak, the main groups are the Iban, Bidayuh, Melanau, Kenyah, Kayan, Lun Bawang/Murut, Penan, Kelabit, and Punan (Ali *et al.*, 2025). Of the estimated 2.47 million inhabitants of Sarawak, the populations distributions consist of: 28.87% Iban, 22.99% Malay, 8.03% Bidayuh, 4.99% Melanau, 6.33% other indigenous groups, 23.38% Chinese, 0.30% Indian, and 5.11% others (Ting & Rose, 2014).

Among these indigenous communities, the Melanau ethnic group are regarded as one of the earliest settlers in Sarawak and are believed to share ancestral origins with the Orang Ulu populations (Musadat & Bahauddin, 2017). The Melanau traditionally inhabit the lower reaches of the Rajang River and adjacent coastal areas extending from the river mouth to Bintulu (Musadat & Bahauddin, 2017). Historical interactions with the neighbouring indigenous group have led to varying degrees of admixture particularly in Melanau subgroups such as the Sekapan, Rajang, Tanjung, and Kanowit mixed with Dayak migrants, who settled

in the Rajang area (C. Y. Chong, n.d.). Traditional Melanau settlements areas include Rejang, Paloh, Belawai, Daro, Matu, Igan, Oya, Mukah, Bintulu, Dalat, Medong, Kut, Narub, Balingian, Kanowit and Tatau (Rahman, 1987) (Figure 2.3).

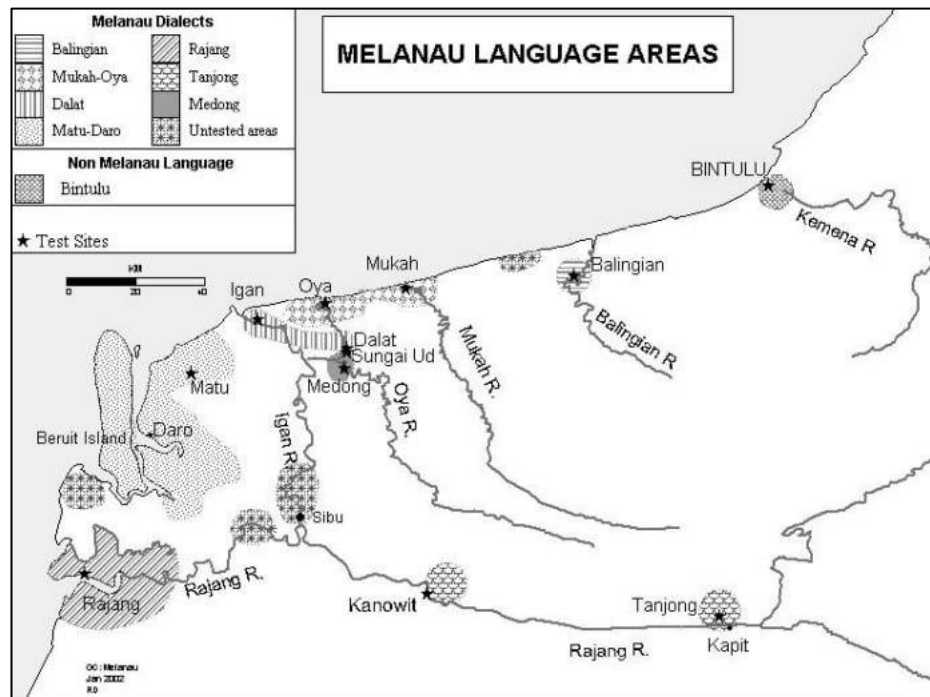


Figure 2.3: Map of traditional Melanau settlement areas (Rensch, 2012).

Historically, Melanau did not identify themselves by the name “Melanau” but instead used the term A-Likou, meaning “people of the river”, reflecting their close relationship with riverine and coastal environments. The name of Melanau is believed to have been assigned by Brunei Malays to describe populations in the coastal swamps and riverbanks of central Sarawak, possibly denoting “coast-dwellers.” (C. Y. Chong, n.d.). Traditionally, Melanau settlements are concentrated near river mouths and coastal plains, where their livelihood centred on fishing, sago palm planting, boat making, carpentry, and blacksmithing (Long & Chik, 2020). As a coastal community, they are closely connected to the river and sea, which are their pillars of economy and culture (Rahman, 1987). From a population genetics perspective, Melanau's long standing settlement patterns, relative geographic continuity and distinct ethnic identity make them an important group in genetic studies in Malaysia.

2.3 Haplogroup B

The strict maternal inheritance of mitochondrial DNA (mtDNA) results in an accumulation of sequence variations along maternal lineage, leading to the formation of monophyletic clusters known as mitochondrial haplogroups. Individuals belonging to the same haplogroup carry a specific sequence motif as a consequence of sharing a common ancestor (Kloss-Brandstätter *et al.*, 2011) making the mtDNA a valuable marker for tracing maternal ancestry and population history. The strong phylogeographic structure of mtDNA allows the identification of genetic lineages that are characteristic of specific geographic regions including Southeast Asia.

Asian haplogroups belong to both N and M macrohaplogroups (haplogroups A, B, F, and N9 belong to macrohaplogroup N, and haplogroups M7a, M7b, M8, D, and G belong to macrohaplogroup M) (Miyamoto-Mikami & Fuku, 2019). Haplogroup B is found at relatively high frequencies in Mainland southeastern Asia (20.6%), Island southeastern Asia (15.5%). Haplogroup B is identified by the presence of a 9-bp deletion in the COII/tRNALys intergenic region of mtDNA (Derenko *et al.*, 2012). The ancestral haplogroup B4 and its subgroups are found mainly in central and Southeast Asia, particularly in coastal regions (Wood *et al.*, 2019), reflecting early human dispersal. Based on the maximum likelihood estimates of haplogroup diversity presented in Ballinger *et al.* (1992), haplogroup B arose between 33,500-16,750 YBP (Schurr & Wallace, 2002).

The 9-bp deletion between the COII and tRNALys genes is **very common** in Asians and populations of Asian ancestry (e.g., Polynesians and Amerinds), an often considered an ethnic specific polymorphism for these populations (Horai and Matsunaga 1986; Hertzberg *et al.*, 1989; Schurr *et al.* 1990). Studies suggest that this deletion most likely arose once in a common

ancestor and subsequently transmitted through generations, forming the basis of mitochondrial haplogroup B (Torroni *et al.*, 1992; Torroni *et al.*, 1993).

Table 2.1: Haplogroup B frequency in Southeast Asia (Schurr & Wallace, 2002).

Population	Sample size	Frequencies
Taiwanese Han	20	7.1
Malaysian Han	14	32.0
Vietnamese	28	17.9
Orang Asli	31	3.1
Malay	14	14.3
Sabah Aborigines	32	18.6
Papua New Guinea	119	41.8
Korea	13	15.4
Tibetans	54	11.1
Nepalese	50	4.0

Although a small number of mtDNA sequences with 9-bp deletion have also been found in other haplogroups, suggests these deletions can occasionally occur independently in different populations. This indicates that, the event defining the B haplogroup, the 9-bp deletion can occur multiple times during human evolution.

2.4 Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) is a scientific molecular biology technique used to amplify a single or a few pieces of DNA across several orders of magnitude to generate millions to billions of copies of a particular DNA sequence (Joshi & Deshpande, 2010). This amplification allows the target DNA fragment to be analyzed in greater detail. PCR was developed in 1984 by the American biochemist Kary Mullis. However, the basic principle of replicating a piece of DNA using two primers had already been described by Gobind Khorana in 1971 (Joshi & Deshpande, 2010). This technique is most widely applied because it is quick,

inexpensive and simple. PCR is capable of amplifying specific DNA fragments from minute quantities of DNA material, even when the source DNA is of relatively poor quality (Erich, 1989).

PCR take advantages of an enzyme that uses a defined segment in a strand of DNA as a template for assembling a complementary strand (Schochetman *et al.*, 1988). Each PCR reaction requires several essential components, including template DNA, primers, nucleotides, and DNA polymerase (Garibyan & Avashia, 2013). The principle of the PCR is simple, requiring a repeated three-step thermal cycling process: (1) denaturation of double-stranded DNA, (2) annealing of primers, and (3) primer extension (Figure 2.4).

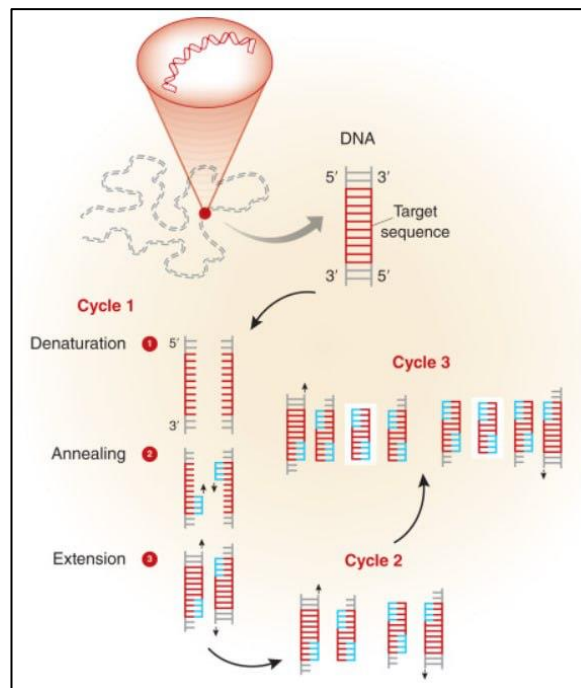


Figure 2.4: Schematic presentation of the polymerase chain reaction (PCR) principle (Garibyan & Avashia, 2013).

During the denaturation step, the double-stranded DNA is held together in the duplex by hydrogen bonds and is separated into a single strand by heating at high temperature, typically 90-97°C (Schochetman *et al.*, 1988). Annealing phase happens at a lower temperature usually

between 50-60°C, allowing the primers to bind to their complementary sequences on single-stranded DNA templates. During the extension step, the thermostable DNA polymerase synthesises new DNA strands at approximately 72°C by incorporating nucleotides complementary to the template strand, starting from the annealed primers. The number of DNA strands doubles upon completion of each cycle. After 30 cycles, a single copy of DNA can be increased up to 1 000 000 copies (Schochetman *et al.*, 1988).

CHAPTER 3

METHODOLOGY

3.1 Materials

3.1.1 Chemicals and Reagents

All the chemicals, reagents and consumables used in this research are listed in Table 3.1 below:

Table 3.1: List of chemicals, reagents and consumables used.

Chemical/Reagents	Company
Tris Base	Vivantis Technologies
Boric Acid	EMSURE
Ethyleneiaminetetraacetate (EDTA)	SIGMA
Agarose Powder	MATRIOUX
Ethidium Bromide (EtBr)	SIGMA
Orange G Dye	Sigma-Aldrich
PrimeWay Genomic DNA Extraction Kit	1 st Base
Absolute Ethanol	Progressive Scientific
PCR reaction mix (One Taq 2X Master Mix with Standard Buffer)	New England biolabs
Primer	Sigma-Aldrich
1 Kbp DNA ladder	Gene DireX
100 bp ladder	ThermoFisher Scientific

Consumables	Company
Pipette tips	Eppendorf
0.5 mL PCR tube	Biologix
1.5 mL centrifuge tube	Biologix
Kim wipes	Kimberly-Clark Worldwide Inc.

3.1.2 Instruments and Apparatus

All the instruments and apparatus used in this research are shown in Table 3.2.

Table 3.2: List of instruments and apparatus used in this research.

Instruments	Brand	Model
PCR Thermal Cycler	Applied Biosystem	Veriti 96-well
Analytical Balance	Shimadzu	ATX224
Hot Plate and Magnetic Stirrer	ERLA	EMS-HP-700
Microwave Oven	Elba	EMO-1706
Spectrophotometer	DeNovix	DS-11
Microcentrifuge	Hettish Uniersa	Micro 22R
DNA Electrophoresis System	Owl	B1
Power Supply	Bio-Ra	Power-Pac 300V
Vortex Mixer	ERLA	EVM 6000
UV Transsiluminator	Uvitec, Cambridge	BTS-26.MS
Pipette (10 μ L, 20 μ L, 100 μ L)	Eppendorf	Research Plus

3.1.3 Chemical and Reagent Preparation

3.1.3.1 Preparation of 10X Tris-Borate-EDTA (TBE)

The Tris-Borate-EDTA (TBE) buffer was prepared by adding a total of 108 g of Tris base, 55 g of boric acid and followed by adding 40 mL of 0.5 M ethylenediaminetetraacetic acid (EDTA). The components were dissolved in approximately 800 mL of deionised water in a 1 L beaker. The solution was stirred continuously using a magnetic stirrer until all the reagents were completely dissolved. The mixture was transferred into a 1 L volumetric flask and deionised water was added until the mark. The prepared buffer was mixed thoroughly, transferred to a 1 L Schott bottle and sterilised by autoclaving.

3.1.3.2 Preparation of 0.5X Tris-Borate-EDTA (TBE) Buffer

The 0.5X Tris-Borate-EDTA (TBE) buffer was prepared by diluting the 10x TBE stock solution with deionised water. A volume of 50 mL 10X TBE stock solution was measured by measuring cylinder and transferred into a 1 L volumetric flask. Deionised water was added up to the calibration mark and was mixed thoroughly by gentle inversion. The prepared 0.5X TBE buffer was transferred into a labelled 1 L Schott and stored at room temperature.

3.2 Methodology

3.2.1 Sampling Method

Participants in this study were selected from individuals of the Melanau ethnic group through a verbal interview conducted before their participation. The interview was carried out to confirm each participant's ethnic background and to ensure they met the inclusion criteria. Individuals who agreed to participate were provided with an informed consent form that should be read and understood before signing the form, based on their voluntary participation in the research.

The inclusion and exclusion criteria of the biological sample are outlined below: Inclusion Criteria:

- i. Participants are unrelated males or females of Melanau ethnicity for at least three generations.
- ii. Participants should be normal or healthy individuals.
- iii. Adult age above 18 years old.

Exclusion Criteria:

- i. Participants with a history of diseases.
- ii. Participants with a family history of mixed marriage.

3.2.2 Sample Collection

The ethical clearance (USM/JEPeM/15100366) for this research involving human participants was obtained from the Ethical Committee for Research. The samples were collected from the Melanau individuals after the ethical clearance was approved. A total of 120 unrelated Melanau individuals were interviewed to confirm their eligibility based on the inclusion and exclusion criteria. Participants who meet all the criteria and have signed the consent form undergo a blood collection process. Approximately 10 mL of blood was collected from each participant, and the blood samples were transferred into ethylenediaminetetraacetic acid (EDTA) tubes. Each tube was labelled according to the participants' records. The blood samples were stored at -80°C for long-term preservation.

3.2.3 Genomic DNA Extraction

Genomic DNA was extracted from whole blood samples collected from Melanau individuals using the PrimeWay Genomic II DNA Extraction Kit (1st BASE, Malaysia), following the manufacturer's instructions with minor optimisation steps. The extraction procedure involved several main steps, including lysis, binding, washing, drying and elution. Prior to the extraction process, the reagents supplied with it were prepared according to the manufacturer's guidelines. Specifically, 5.5 mL of nuclease-free water was added to reconstitute the Proteinase K solution, and 140 mL of absolute ethanol (>99.5%) was added to Wash Buffer G2. All reconstituted reagents were mixed thoroughly and stored at -20°C until use.

Before beginning the DNA extraction, the elution buffer was preheated in the water bath at 60°C. The frozen blood samples stored in the EDTA tubes were thawed at room temperature for approximately 15 minutes and followed by gentle vortexing to ensure homogeneity. For each extraction, 20 µL of Proteinase K solution was dispensed into a sterile 1.5 mL

microcentrifuge tube and 200 μ L of whole blood was added to the tube. The mixture was briefly vortexed and incubated at 60 °C for 15 minutes to promote lysis of cellular and nuclear membranes.

Following the initial lysis, 200 μ L of GL2 buffer was added and the tube was gently mixed. The sample was incubated again at 60 °C for 15 minutes. Subsequently, 4 μ L of RNase A solution was added to degrade the residual RNA followed by vortexing and incubation at room temperature for 15 minutes. To facilitate DNA binding during the subsequent column purification step, 200 μ L of absolute ethanol was added and mixed thoroughly.

Approximately 600 μ L of the lysate was transferred into the PrimeWay Genomic II Column, which was placed into a collection tube and centrifuged at 14000 x g for 1 minute. The flow-through was discarded and the column was placed into a new collection tube. The purification process continued with sequential washing steps. Firstly, a total of 400 μ L of Wash Buffer G1 was added to the column and centrifuged at 14,000 x g for 1 minute. The flow-through was discarded, and the column was returned to the collection tube. Next, 600 μ L of Wash Buffer G2 was added and centrifuged again at 14,000 x g for 1 minute. The flow-through was discarded, and the column was placed back into the same collection tube.

To ensure complete removal of ethanol and residual contaminants, the column was centrifuged at 14,000 x g for 3 minutes to dry the silica membrane. The column was then transferred to a clean 1.5 mL microcentrifuge tube for the DNA elution. A volume of 100 μ L of pre-heated (60°C) elution buffer was applied directly to the center of the membrane and incubated at room temperature for 3 minutes. DNA was eluted by centrifugation at 14,000 x g for 1 minute. The purified genomic DNA was stored at -20°C until further downstream molecular analyses.