COMPLETE MITOCHONDRIAL GENOME OF TRAGULUS KANCIL AND TRAGULUS NAPU IN PENINSULAR MALAYSIA

ASIA ABDILLAHI SULEIMAN

UNIVERSITI SAINS MALAYSIA

2023

COMPLETE MITOCHONDRIAL GENOME OF TRAGULUS KANCIL AND TRAGULUS NAPU IN PENINSULAR MALAYSIA

by

ASIA ABDILLAHI SULEIMAN

Thesis submitted in partial fulfilment of the requirements
for the degree of
Master of Science (Forensic Science)

September 2023

ACKNOWLEDGEMENT

In the name of Allah, the Most Merciful and Most Benevolent. First and foremost, I am grateful to Allah for granting me the strength and excellent health necessary to complete my thesis Master of Science. I would like to take this opportunity to thank my supervisor, Assoc. Prof. Dr. Edinur Hisham Atan, for his expert guidance, patience, and encouragement throughout my research and studies. I would also like to express my appreciation to Mr. Frankie Anak (Co-Supervisor) for his invaluable advice and encouragement and Dr. Nurasmat Mohd Shukri, who served as the research coordinator for Research Project (GTF 508), as well as all the instructors in the field of Forensic Science, for their unwavering support and attentive guidance. I sincerely thank the General Director of the Peninsular Malaysia Department of Wildlife and National Parks (PERHILITAN) for granting me permission to pursue my research. In addition, I would like to thank my parents for their unwavering love and support throughout my academic journey. Their encouragement, prayers, and sacrifices have been crucial to my success in achieving my objectives. I would also like to express my gratitude to my devoted husband for his unwavering support, motivation, and patience throughout this process. His unwavering confidence in me has provided me with strength and inspiration. I would like to extend this acknowledgment to my two sons Amar and Abdalla for their love and tolerance. They have been a constant source of happiness and inspiration, and their presence in my life is a blessing for which I am thankful every day. This MSc work was supported by the National Conservation Trust Fund by Natural Resources from the Ministry of Energy and Natural Resources Malaysia (Grant no: 304/PPSK/6150219).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

T.kancil	*Tragulus kancil*
T.napu	*Tragulus Napu*
NGS	Next generation sequence
mtDNA	mitochondrial DNA
ATPS	Adenosine triphosphate
STR	Short tandem repeats
VNTR	Variable tandem repeats
tRNA	Transfer RNA
WGRI	Wildlife Genomic Research laboratory
PCR	Polymerase Chain Reaction
TMM	Tagmentation master mix
PTC	Post tagmentation clean -up
TAG	Tagmentation
TWB	Tagment wash buffer
PLT	bead-linked transposomes
Spb	sample purification beads
EtOH	Ethanol
RSB	Resuspension Buffer
DNA	Deoxyribonucleic acid
TB1	Tagment buffer 1
TSB	Tagment stop buffer
EPM	Enhanced PCR mix
RSB	Resuspension buffer

rRNA	Ribosomal RNA
CRs	Coding Control Region
PCGs	Protein-Coding Region
IUCN	International Union Red list
N-J	Neighbour-joining
T.javanicus	*Tragulus Javanicus*
COI	Cytochrome oxidase

LIST OF APPENDICES

GENOM MITOKONDRIA LENGKAP TRAGULUS KANCIL DAN TRAGULUS NAPU DI SEMENANJUNG MALAYSIA

ABSTRAK

Negara tropika Malaysia mempunyai ekosistem yang unik, terletak di rantau Asia Tenggara dan berkongsi sempadan dengan Thailand,Singapura,Brunei dan Indonesia. Pihak berkuasa pemuliharaan Malaysia telah mengenal pasti keperluan untuk memulihara sumber asli dan alam sekitar mereka termasuk flora dan fauna ini termasuklah *Tragulidae* kumpulan keluarga kecil dengan 3 genera dan 2 spesies. Dua daripada spesies *Tragulidae T.kancil* dan *T.napu* yang berasal dari Semenanjung Malaysia menghadapi ancaman yang serius termasuk kemerosotan habitat, tompokan hutan, pemburuan haram, dan perdagangan. Ancaman ini membawa kepada penurunan saiz populasi *kancil* dan *napu* Kajian ini melibatkan jujukan genomik mtDNA penuh menggunakan platform jujukan generasi seterusnya ke atas *T.kancil* dan *T.napu* dari Semenanjung Malaysia untuk kali pertama. Sebanyak 17 individu *Tragulus* spp. digunakan untuk pengumpalon spesimen tisu. Sampel-sampel ini adalah 10 untuk *T.kancil* dan tiga sampel dari satu individual *T.napu* yang diambil dari Perlis dan Perak. Tambahan enam daripada 17 tersebut adalah sampel *T.kancil* yang dihantar ke makmal forensik hidupan liar dari Selangor, Pahang, Kedah, Perak dan Sarawak sebagai sebahagian daripada operasi rutin.Dalam pokok filogenetik, *T.kancil* berbeza daripada saudara *Tragulus* yang lebih kecil (*T.javanicus*). Tidak seperti *T.napu* di mana perbezaan yang jelas boleh diperhatikan antara sampel *T.napu* dari Malaysia (Perlis). berbanding T.napu dari Thailand, tiada rujukan mtDNA penuh *T.kancil* dari negara lain tersedia untuk perbandingan geografi pada masa ini.

Secara, keseluruhan data genomik mtDNA yang dikumpul daripada *T.kancil* dan *T.napu* menyokong klasifikasi semasa *Tragulus* dan pembezaan antara spesies *Tragulus* dan *Rusa*.

COMPLETE MITOCHONDRIAL GENOME OF TRAGULUS KANCIL AND TRAGULUS NAPU IN PENINSULAR MALAYSIA

ABSTRACT

The tropical country of Malaysia has a unique ecosystem, located in the southeast Asian region and sharing a border with Thailand,Singapore,Burnei and Indonesia. Malaysian conservation authorities have identified the need to conserve their natural resources and environment including floral and faunal. These include *Tragulidae* which is a small family group with 3 genera and 2 species. Two of *Tragulidae* species are *T.kancil* and *T.napu* that native to Peninsular Malaysia are facing a crucial threat of extinction including habitat degradation, forest fragmentation, illegal hunting, and trading. These threats are leading to decline in their population size. The present study involved full mitochondrial DNA (mtDNA) genomic sequence using next generation sequencer platform of *T.kancil* and *T.napu* from Peninsular Malaysia for the first time. A total of 17 individuals of *Tragulus* spp. were used to collect tissue specimens. There were 10 *T.kancil* individuals and triplet samples of a *T.napu* which were collected from Perlis and Perak. Additional six of those 17 were *T.kancil* samples sent to wildlife forensic laboratory from Selangor, Phang,Kedah,Perak and Sarawak as a part of routine operations.

In phylogenetic tree, *T.kancil* are different from their lesser *Tragulus* relative (*T.javanicus*). Unlike *T.napu* where a clear distiniction can be observed between a *T.napu* samples from Malaysia (Perlis) versus a *T.napu* from Thailand, no full mtDNA reference sequence of *T.kancil* from other country is currently available for geographical comparison.

Overall, the whole mtDNA genomic data collected from *T.kancil* and *T.napu* support the current classification of *Tragulus* and distinication between *Tragulus* and *Rusa* species.

CHAPTER 1

INTRODUCTION

1.1 Background

The tropical country of Malaysia has a unique ecosystem, located in the southeast Asian region and sharing a border with Thailand,Singapore,Brunei and Indonesia (DasGupta & Shaw, 2013). Peninsular Malaysia and Borneo are Malaysia's two distinct geographical regions, with variety of regional habitat types that contributes significantly to species richness and diversity (William-Dee et al.,2019).

Malaysian conservation authorities have identified the need to conserve their natural resources and environment including faunal and floral (Schwabe et al., 2014). In Peninsular Malaysia, 2.9 million hectares of land were reserved for natural resource conservation. Specifically, the 25% lowland forest and 14% hill dipterocarp forest of designated land were notified as protected areas. The surrounding lands have been developed into residential settlements, agricultural plots, and industrial lands (William-Dee et al., 2019).

Mousedeer are small ruminant animal species of the *Tragulidae* family found throughout Peninsular Malaysia. *Tragulus kancil* and *Tragulus napu* are the two species of mousedeer that are native to Peninsular Malaysia (Endo et al., 2004). The *T.kancil* is the smallest mousedeer species also known as lesser mousedeer, weighing between 0.7 and 2.6 kg, whereas *T.napu* is a larger mousedeer and known as greater mousedeer, weighing between 3.5 and 6.0 kg. The former is often resident to below 100m of lowland forest habitat, whereas *T.napu* is resident to higher lowland forest habitat (Bakar et al., 2018).

Tragulus spp. are mainly fed on young leaves, leaf shoots and fallen fruits of native trees (Ng et al., 2011). The Malaysian mousedeer are not listed as endangered species on the global International Union for Conservation of Nature (IUCN) red list, even though their present population size raises severe concerns about extinction due to habitat loss and degradation (Olivieri et al., 2008). Furthermore, those mousedeer species are the source of food for the local hunters. The drastic decline of the mousedeer population will result in the loss of valuable genetic diversity of the country. The Global Wildlife Protection Act 210 has designated mousedeer as protected species worldwide (Aziz et al., 2013). Unfortunately, genetic information on mousedeer in Malaysia is lacking, thus limit conservation and protection efforts (Bakar et al., 2018).

Mitochondria is a vital organelle to regulate apoptosis, aging, immunity and metabolism. (Son & Lee, 2019). It is also a site for oxidative phosphorylation, adenosine triphosphate (ATPs) production, and biochemical functioning (Boore, 1999). The mitochondrial DNA (mtDNA) population data has proven useful for phylogenetic and evolutionary inferences (Muangkram et al., 2018). Nuclear DNA is less frequently studied than mtDNA, as the latter has a greater number per cell and high mutation rate (Sequeira et al., 2015). Thus mtDNA can be used as an effective tool for examining the evolutionary relationships and conservation purpose at inter species or intra species level (Muangkram et al., 2018).

1.2 Problem statement

The *Tragulidae* is a small family group with 3 genera and 2 species of *Tragulidae* species *T.kancil* and *T.napu* that native to Peninsular Malaysia facing a crucial threats including habitat degradation, forest fragmentation, illegal hunting and trading (Laidlaw, 2000; Magintan et al., 2021). These threats are leading to decline in their population size. Previously, Bakar et al. (2018) has tried to study the relationship distance between these two species using part of mtDNA data Cytochrome oxidase I (COI), but sequencing the full mtDNA genomic sequence of *T.kancil* and *T.napu* from Peninsular Malaysia has never been reported. Therefore, full mtDNA sequences of *T.kancil* and *T.napu* are needed for conservation and protection purposes.

1.3 Objectives of the study

1.3.1 General objective

- To obtain the complete mtDNA genomes of *T.kancil* and *T.napu* species in Peninsular Malaysia.

1.3.2 Specific objectives

- To genotype full mtDNA sequence of *T.kancil* and *T.napu* using the next generation sequence platform.
- To conduct a comparative analysis of mtDNA genomes obtained from both *Tragulus spp*.
- To study variations of mtDNA genomes of *T.kancil and T.napu* population in Peninsular Malaysia and those reported from elsewhere.

1.4 Significance of the study

The entire mtDNA genome of *T.kancil* and *T.napu* species can reveal important details about their phylogenetic relationships, genetic diversity, and evolutionary histories. Such genomic data can also be used conservation which include species identification associated with wildlife crimes (William-Dee et al., 2019). Currently full mtDNA of *Tragulus* species in Malaysia are only available for *Tragulus javanicus* but not for other *Tragulus* species include *T.kancil* and *T.napu.*

CHAPTER 2

LITERATURE REVIEW

2.1 Wildlife animal species in Malaysia

Malaysia is home to a broad range of wildlife animal species, including many unique and uncommon animals found nowhere else on the planet. Malaysia has about 300 mammal species, 742 bird species, 242 reptile species, and 150 amphibian species, according to the IUCN Red List (IUCN, 2021).

Among those is *Orangutans* also known as *Pongo pygmaeus*,that can be found in the rainforests of Borneo and Sumatra *Orangutans* are in a state of critical endangerment due to the destruction of their natural habitat and the illegal hunting for their meat. Many conservation efforts and ecotourism endeavors are currently under way to conserve these animals. (WWF, 2017).

Another iconic species in Malaysia is the Malayan tiger (*Panthera tigris jacksoni),* which can only be found in Malay Peninsula. Like orangutans, Malayan tigers are also critically endangered due to habitat loss and poaching (WWF, 2022). Conservation efforts are underway to protect the remaining tigers, including anti-poaching patrols and habitat restoration projects (WWF, 2022). The Bornean pygmy elephant *(Elephas maximus borneensis),* the probosci's monkey (*Nasalis larvatus*), and the Malayan tapir (*Tapirus indicus*) are also endangered state. The Bornean pygmy elephant is a subspecies of the Asian elephant that is only found on the island of Borneo, while the proboscis monkey is known for its large nose and distinctive vocalizations. The Malayan tapir is a large herbivorous mammal with a distinctive black-and-white pattern (IUCN, 2021).

The Illegal Wildlife Trade (IWT) describes wildlife crimes as illegal exploitation, possession, processing, transportation, and sale of protected flora and fauna, as well as the concealment or laundering of financial gains from these operations. However, international IWT regulations are either ineffective or inadequate to handle the gravity of some offenses. IWT used to be considered to be a new threat to ecosystems and the biodiversity they support, but it is now recognized as one of the largest transnationals organized criminal activities, alongside drug trafficking, arms trafficking, and human trafficking, with an estimated annual value of $7 billion to $23 billion US dollars. (Ahlers et al., 2017). Wildlife forensics, or the scientific investigation of wildlife, their products, and derivatives for enforcement reasons, has been developed to assist with enforcement action against wildlife criminals (Jun et al., 2011).

Wildlife crimes showed an increasing trend in Peninsular Malaysia and the threat included *T.kancil* and *T.napu* and the persistent global issue is driven by a rising need for resources related to wildlife (Xin et al., 2022). Overall, Malaysia's wildlife is an important part of the country's natural history, and efforts to conserve and protect these species are essential for their existence and the ecosystem's health.

2.2 Tragulus kancil

The lesser mousedeer (*Tragulus kancil*), also known as the smaller Malay chevrotain or *kancil*, is a species of even-toed ungulate in the *Tragulidae* family. Numerous Southeast Asian countries, including Myanmar, Brunei, Cambodia, China (Southern Yunan), Indonesia (Kalimantan, Sumatra, and many other small islands), Laos, Malaysia (Peninsular Malaysia, Sarawak, and many other small islands) Thailand, and Vietnam, have lesser mousedeer (Siteadmin, 2021). They live on the floor of primary and secondary forests feeding on leaves, shoots, fruits, and sometimes fungi. It is mainly

crepuscular (i.e., active early morning and late afternoon), but is sometimes nocturnal (Meijaard & Sheil, 2007).

Habitat fragmentation caused by human activities (e.g., road and railway, logging, agricultural expansion) is one of the major threats to global biodiversity as it leads to declines in nearly all taxonomic groups including *T.kancil* species. (Hazwan et al., 2022).

2.3 Tragulus napu

T.napu is a small, solitary ungulate that is commonly known as the "greater mousedeer." The species is currently in the list of concerns in Malaysia due to habitat loss and hunting pressure (Wongloet et al.,2023). *T. napu* belongs to the family *Tragulidae* and is classified under the order *Artiodactyla*. Its scientific name is *Tragulus napu* (Mousedeer or chevrotain). Therefore, *T.napu* is found in Southeast Asia, including Malaysia, Indonesia, Thailand, and the Philippines. The species inhabits a variety of habitats, including tropical rainforests, mangrove forests, and grasslands. (Conjeaud, 2017). *T.napu* is primarily herbivorous, feeding on a variety of plants, including leaves, fruits, flowers, and shoots. The species has been observed feeding on more than 100 plant species, including the leaves of various shrubs and trees such as *Diospyros* sp., *Melastoma malabathricum*, and *Artocarpus* sp (Chandrapatya et al., 2016).

T.napu can be distinguished from the closely related lesser mouse deer, *T.kancil*, through several observable external characteristics. These distinctions include the number of throat stripes, with *T.napu* having five and *T.kancil* having three. Additionally, *T. napu* lacks the brown stripes on the belly that are typically found in *T.kancil*. They can also be distinguished by examining the presence of the nape line and the pattern of mottling on their underparts (Chua et al., 2009).

2.4 Animal mtDNA

Mitochondrial DNA (mtDNA) is important for studying genetic relationships, but its analysis can be challenging due to the presence of nuclear copies of mitochondrial genes (numts) and heteroplasmy (Xue et al., 2023). Obtaining pure mtDNA is crucial to ensure accurate data interpretation. A study conducted by Ibarguchi et al. (2006), demonstrated the ease of obtaining semi-pure mtDNA from wildlife tissues preserved under field conditions. They compare the success of different extraction methods and highlight the use of unfertilized or undeveloped eggs as a pure source of mtDNA. The study emphasizes the importance of confirming the mitochondrial origin of reference sequences and warns against relying solely on long-template PCR, cloning, and degenerate primers for reliable results. Purified mtDNA combined with these techniques can effectively distinguish numts from true heteroplasmy (Ibarguchi,et al., 2006).

Animal mtDNA genome maintains a consistent size and gene content, but exhibits flexibility through various gene rearrangements, potentially mediated by transfer RNA (tRNA) genes (Logan, 2017). While it was believed that mtDNA structure remained constant within a phylum, recent studies show transpositions occur even among closely related organisms (Saccone et al., 1999).The nucleotide composition of mtDNA is highly variable, with low GC content and a skewed distribution of complementary bases, especially in vertebrates. In mammals, the replication mechanism of mtDNA influences coding strand base composition and gene conservation. The evolutionary rate of mtDNA components differs, with gene-specific non-synonymous rates similar to nuclear genes but 22-fold higher for synonymous mutations (Meyer et al., 2012). Even though tRNA genes are highly conserved they evolve 100 times faster than their nuclear counterparts (Chen et al., 2008).

mtDNA sequence data have numerous applications, but current methods require expensive and sophisticated laboratory equipment's which limit field testing advances in microfluidic chips, paper-based methods, and isothermal amplification offer potential for on-site mtDNA detection. Developing integrated field test platforms for sample preparation and detection is the future direction (Dhar,et al 2022).

Morin,et al(2010) examined mtDNA genome variations in marine samples using high throughput NGS sequencing machine. They identified distinct ecotypes with divergent evolutionary histories and suggested reclassifying some as separate species. The study highlights the importance of whole mtDNA genome for species characterization and evolutionary studies obtained accurate mutation rate estimates.

2.5 Animal whole mtDNA data

Using the shotgun genome skimming method, the complete mtDNA genome of Bubo bubo was sequenced. The mtDNA genome (mitogenome) is 18,956 bp long and has a base composition of 29.6% A, 22.5% T, 33.8% C, and 14.1% G. It includes 13 protein-coding genes (PCGs), 22 tRNA genes, 2 rRNA genes, and 2 non-coding control regions (CRs). Most PCGs start with the ATG codon, except for ND3 and ND5, which begin with ATA. Seven PCGs end with the TAA codon, while three (ND2, ND4, and COX3) end with a single T.ND1 and COX1 terminate with AGG and ND6 stops with TAG. Phylogenetic analysis based on the 13 PCGs showed that Bubo species form a monophyletic group, which is closely related to the genus Strix within the Strigidae family (Ren et al., 2012).

Zhang,et al(2012) sequenced the complete mitogenome of the Scimitar-horned oryx for the first time. It has a length of 16,756 bp and follows the typical structure found in mammals, including 22 tRNA genes, 13 protein-coding genes, and 2 ribosomal RNA (rRNA) genes. The genome's base composition is 33.5% A, 26.8% T, 26.2% C, and 13.5% G. The protein-coding genes use ATG as the start codon, and most end with TAA. The tRNA genes form cloverleaf-shaped structures, except for tRNA-Ser (*AGY*).

The complete DNA sequence of the bovine mitogenome, consisting of 16,338 nucleotides, is highly similar to the human mtDNA (Anderson et al., 1982). This includes gene organization and protein gene sequences which show 63% to 79% homology with their human counterparts. Most nucleotide differences occur in the third positions of codons. The bovine mtDNA and rRNA genes show conserved features consistent with proposed secondary structure models. However, the bovine D-loop region, which differs in length, is only slightly homologous to the corresponding region

in the human mitochondrial genome and accounts for the major size difference between the mtDNA genomes of human and bovine (Anderson et al., 1982).

2.6 MtDNA studies of *Tragulus* spp.

Several attempts were made to study the mtDNA of *Tragulus* spp. for conservation and to study the phylogenetic relation of different species under *Tragulida*e species. Endo et al. (2004) reported the phylogeographical distributions and variations of *T.napu* and *T.javanica* using cytochrome b gene of mtDNA. In contrast only 0.145 and 0.28 gene variation were observed at D-loop and cytochrome oxidase- I (COI) of *T.kancil* and *T.napu,* respectively Bakar et al. (2018).

The first whole mtDNA genome for *Tragulidae* spp were obtained for the Indain mousedeer (*Moschiola indica*) (Sarvani et al., 2018). Genome analysis was achieved using 22 pairs of oligonucleotide polymerase chain reaction (PCR) primers. The mitogenome of *M. indica*, spanning 16,444bp, exhibited a high degree of similarity to the mtDNA genomes of most vertebrates in terms of organization. It contained 13 protein-coding genes, 22 transfer RNA, 2 ribosomal RNA, and 1A+T-rich region. Comparative analysis with over 52 mitogenomes of the *Artiodactyla* order revealed a conserved nature in terms of gene organization, codon usage, gene orientation, and evolutionary rates of proteins. Notably, *Moschiola indica* (*M. indica*) was found to possess an additional copy of trnF gene. The complete mitogenome and protein-coding genes of *M. indica* exhibited a strong bias towards 1A+T content. Among the protein-coding genes, ATP8 showed the highest rate of evolution, while COX3 exhibited the lowest. Additionally, the study observed a higher purifying selection pressure on the *Tragulidae* family compared to Bovidae and Cervidae. phylogenetic analysis placed *M. indica* as the sister-group to all other ruminants, consistent with previous analyses, and

surprisingly, *Moschiola* was found to be the sister-group to the other two *tragulid* genera, *Tragulus* (Asia) and *Hyemoschus* (Africa). This finding deviates from the conventional notion that Asian species form a monophyletic group, highlighting the need for further genome studies on uncharacterized species this include for *T.kancil* and *T.napu* where whole mtDNA yet to be reported (Sarvani et al., 2018).

CHAPTER 3

METHODOLOGY

3.1 Experimental design

Figure 3.1 shows the workflow of the present study. All work related to this study was conducted internally at the Wildlife Genomic Research Laboratory (WGRL), Department of Wildlife and National Parks Peninsular Malaysia (PERHILITAN) Cheras, Kuala Lumpur.

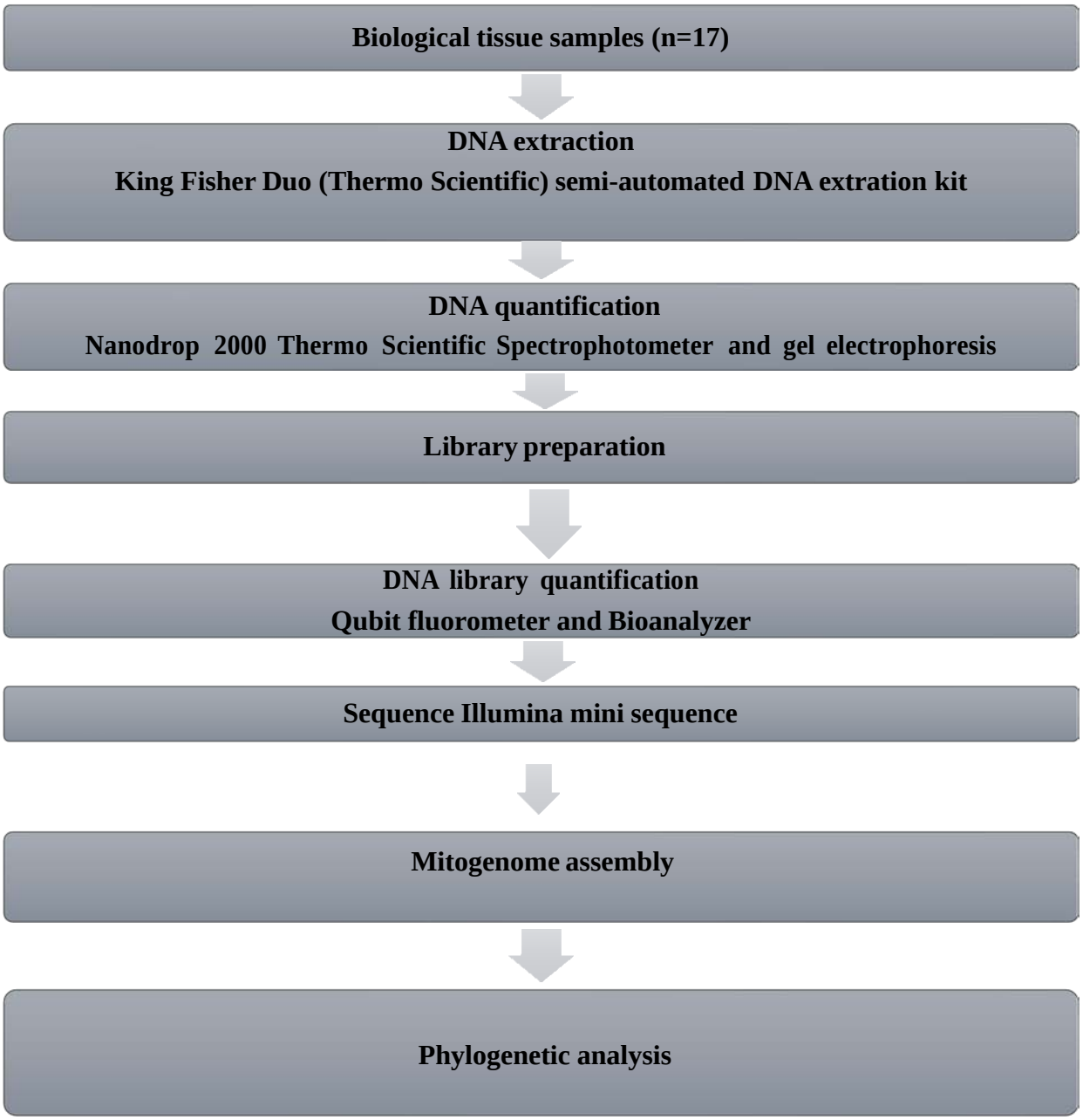

Figure 3.1: The workflow of full mtDNA genotyping of two *Tragulus* species in Peninsular Malysia.

3.2 Biological samples

A total of 17 individuals of *Tragulus spp*. were used to collect tissue specimens. They were 10 under *T.kancil* individuals and sample of one *T.napu* individual were collected from Perlis and Perak (Table 3.1). The sampling was done after physical check and morphological measurements of postmortem individuals. These measurements are body size, weight, height, length, and head measurements.

Additional, six samples of those 17 were from Selangor (TK-Sub01), Pahang (TK-Sub02), Kedah (TK-Sub03), Perak (TK-Sub04), Sarawak (Tk-Sub05) and Kedah (TK-Sub06) all the six were received by the National Wildlife Forensic Laboratory as part of its routine operation. The details of each sample are listed in Table 3.1.

Table 3.1 Detail of sampling location and samples size for two species of *Tragulidae* family.

***Tragulidae* family**	**Location of collection of samples**	**n (17)**	**Sample ID**
T.kancil	PKHL Sugai Batu Pahat Perak	10	TK-01, TK02, TK03, TK-04, TK-05, TK-06, TK-07, TK-08, TK-09, TK-10
T.napu	PKHL Sungkai Perlis	1	T.Napu-A
T.kancil	NationalWildlife ForensicLaboratory (DWNP).	6	TK-Sub-01(Selongor), TKSub-02, (Phang) TKSub-03, (Perak)TKSub-04(Kedah), TKSub-05 (Sarawak), TKSub-06(Kedah).

Figure 3.2: This figure shows the *T.napu* species used in this study.

Figure 3.3: This figure shows *T.kancil* species used in this study.

3.3 Materials

List of consumables, reagents, kits, and instruments used in present study are shown in Table 3.2 and Table 3.3.

Table 3.2: List of Chemical and Commercial kit.

Analysis step	Chemical/ reagent/ Commercial kit	Supplier/Manufacturer
DNA extraction	Isoparopanol, Qiagen (DNeasy, blood, and tissue Kit)	Merck, Germany Qiagen, Hilden, Germany
DNA qualification	Nanodrop No reagent was needed. Qubit,Qubit4.0 dsDNA HS Assay kit	Thermofisher Scientific
Gel electrophoresis	Agarose Powder, LE, Analytical grade,Low TE Buffer,Nucleic acid dye ,DNA Ladder (1bp)	Promega,USA,GbioscienUSA, SigmaAldrich(M)CleaverScientific
Library Preparation	Illumina Library preparation Kit.	Illumina,Inc,San Diego,USA
Sequencing	Illumina -minni sequence Kit	Illumina,Inc,San Diego,USA

Table 3.3: List of consumables

Consumables	Supplier/Manufacturer
Latex examination Gloves	Repfon Glamor Sdn., Malaysia
MicroAmp® PCR reaction tubes	Applied Biosystems ™, USA
Microcentrifuge tubes (1.5ml)	Axygen® Scientific, USA
Pipettes tip (10,20,200 and 1000µl	Bioline, London, UK
BLT(yellow Cap)	Illumine, Inc,USA
TB1	Illumine, Inc,USA
Nuclease free water	Applied Biosystems ™, USA.
PCR plate and seal(micro Biord)	Applied Biosystems ™, USA.
1.5mlcentrifuge tube	Repfon Glamor Sdn., Malaysia
TSB	Illumine, Inc,USA
TWB	Illumine, Inc,USA
96 well plate magnet	Applied Biosystems ™, USA.
Nucleas free water	Applied Biosystems ™, USA.
EPM	Illumine, Inc,USA
Index adapter's	Illumine, Inc,USA
SPB	Illumine, Inc,USA
RSB	Illumine, Inc,USA
80% EtOH	Repfon Glamor Sdn., Malaysia
Deep well plate(kingfisher plate)	Applied Biosystems ™, USA.
PCR tubes	Applied Biosystems ™, USA.

3.4 Working area, sterilization, and precaution steps

Cross contamination is a major issue in any DNA related works. (Guo et al., 2014) To produce good and reliable genotyping results, prevention of cross contamination should be considered and become a top priority. To achieve this, different working areas were dedicated for several different DNA analytical procedures in the present study. DNA extraction was carried out on the bench work while library preparation pre-PCR setup

and handling of post-PCR products were conducted in a PCR room equipped with a clean laminar flow. These working areas including all pipettes were decontaminated to clear any leftover library products. Laminar flow was wiped with ethanol and exposed to ultraviolet (UV) light for 10 minutes prior to any library preparation related work.

Sterilization process is also crucial in preventing cross contamination. including All tips, micro centrifuge tubes, 96 well plates and PCR micro reaction tubes were autoclaved at 15 psi, 121℃ for 20 minutes and dried up at 65°C in hot air oven before being used in laboratory works. Several precautionary steps were considered during library preparation for Illumina's next-generation sequencing (NGS) due to the highly sensitive nature of the process. Library preparation involves the amplification of tagmented DNA, and even a minute amount of contamination can significantly impact the entire DNA analysis (Conrads & Abdelbary, 2019). The most common types of contamination during library preparation include cross-contamination between index adapters and samples, contamination with exogenous genomic DNA, and carryover of extracted products (Huang&Li,2022). To prevent these potential sources of contamination, several preventive measures recommended for PCR tagmented were adapted and applied to library preparation for Illumina NGS, as discussed in discussion chapter.

3.5 Genomic DNA extraction

DNA was extracted from each of the sample by following the animal muscle tissue spin-column protocol provided along with the kit purchased from Qiagen (Hilden, Germany). Firstly, 25 mg of *Tragulus* muscle tissue samples were cut into small pieces and then placed in a 1.5 mL microcentrifuge tube. A total of 360µL of ATL buffer was added into the 1.5 mL microcentrifuge tube containing the *Tragulus* muscle tissue. This was

followed by the addition of 40µL of *proteinase K* into the same 1.5 mL microcentrifuge tube. The mixture was mixed thoroughly by vortex. After that, the mixture was incubated overnight at 56°C until the *Tragulus* muscle tissues completely lysed. The mixture was vortexed occasionally during the incubation period to disperse the muscle tissues. After the tissues had completely lysed, which was indicated by the clear solution of the mixture, DNA extraction was conducted in The Kingfisher Duo (Thermo Scientific) semi-automated DNA extraction machine. Briefly, 200uL of Buffer AL, 200uL Isoparopanol, 25uL of MagAttract Suspension G were mixed with 200uL of the lysate. Then 500uL of each wash buffers, namely Buffer AW1 and Buffer AW2 were added. Finally, 100uL of nuclease free water is added into elution well. Based on the predefined protocol in the machine's software, the machine performed DNA capture using magnetic beads, then a series of washing steps, drying, and finally elution. The eluted DNA samples were transferred into 1.5ml tube prior to storage at -20 ^{0}C.

3.6 Agarose gel electrophoresis of high molecular weight DNA

The presence of high molecular weight DNA in extracted samples was detected by using agarose gel electrophoresis technique. The gel was prepared by adding 0.5g agarose powder to 50 ml of 0.5X TBE buffer. The mixture was then dissolved by heat using microwave at 60^{0}C. The solution was allowed to cool down and added with 4ul of nucleic acid dye, before being poured into a casting tray.

A total of 2 µl of each extracted DNA was then mixed with loading dye and loaded into the wells on the gel. Size standard (3µl of 1bp DNA ladder) was also loaded into one of the well as a reference. The loaded high molecular weight DNA and reference size standard were then subjected to electrophoresis in 0.5X TBE buffer at 80V for 45

minutes. The presence of high molecular weight DNA was observed using UV transilluminator (UVltec, Cambridge).

3.7 DNA quantification

DNA quantification of 17 biological samples of both *T.kancil* and *T.napu* species was performed using the Nanodrop 2000 spectrophotometer (Thermo fisher Scientific, USA). The instrument was set up according to the manufacturer's instructions, and the sample pedestal was carefully cleaned before each measurement to avoid any potential contamination.

A small volume typically (1-2μL) of the DNA sample was pipetted onto the sample pedestal of the Nanodrop 2000. The instrument was then initialized, and the absorbance measurements were taken at 260nm wavelengths. The Nanodrop 2000 provided immediate results, displaying the DNA concentration directly on the instrument's screen.

The A260/A280 and A260/A230 ratios were also calculated automatically, indicating the purity and potential contamination of the DNA sample, respectively. A high A260/A280 ratio (around 1.8-2.0) indicated pure DNA with minimal protein contamination, while a suitable A260/A230 ratio (above 1.8) indicated a low level of contamination from substances such as salts or organic compounds.

3.8 DNA concentration of extracted DNA

Extracted DNA samples were done by using simple protocol assay. Standard 1 and Standard 2 were prepared by adding 190ul buffer to each then added standard of 10ul of each dsDNA high sensitive assay kit (HS) standard 1 and 2, then 197ul of buffer plus 3ul of the sample was added in each sample and incubated for 2 minutes after mixed

well and covered with aluminum foil to protect the light and then the machine read each of the samples.

3.9 NGS library preparation

A volume of 30ul containing 250ng of genomic DNA from each sample was sheared to a target size of 600pb using bead-linked transposomes methods, and each sample was tagged with an index adapter for identification in the multiplex reaction. DNA samples were prepared as required input amount and volume and placed into a PCR plate. A combination of 10ul BLT +10ul TB1 for each sample was prepared as a cocktail (tag mentation master mix (TMM). After that 20ul of TMM was added to each sample on the PCR plate (Using a 200ul pipette). And mixed well by pipetting up and down the samples. Subsequently, the samples were performed using a PCR machine and ran the Tagmentation Procedure (TAG) program (duration 20 minutes) at a temperature of 55^0C for 15 minutes.

In addition to that the next step post tagmentation clean-up was performed, step1 PCR plate was added 10ul TSB into each sample, and each sample was mixed slowly by pipetting to resuspend the beads after that the palate of the PCR plate was put back into the PCR machine and post tagmentation clean-up (PTC) program (duration of 20minutes) And the temperature was (37 ^{0}C for 15 minutes). When the PTC run was done, the PCR plate was taken out of the machine and put into the magnetic stand for 3 minutes until the solution was clear and the beads bind to the side of the well. After this step, all the supernatants were discarded from each sample and only the beads remained in the well. This step was done two times to ensure that the solution is clear in the well.

Then the plate was removed from the magnetic stand, and 100ul of TWB was added to each sample in the plate, pipetting directly into the beads, and mixed slowly until the

beads were suspended. Once again, the plate was placed into the magnetic stand for 3 minutes and then the supernatant was discarded. After that, the plate was removed from the magnetic stand and added 100ul TWB, and slowly mixed until all the beads of each sample were suspended after this step was done the plate was sealed and put back into the magnetic stand until the next step of Amplification of Tagmented DNA. This step while the PCR plate is still on the magnetic stand. A 20ul EPM+20uL nuclease free water PCR cocktail was prepared and multiplied by the number of the samples and then vortexed and centrifuged briefly. Next, all the supernatant was discarded from the plate while it was on the magnetic stand then the plate was removed from the magnet. Then 40ul of prepared PCR master mix was added into each sample in the plate and pipetted directly into the beads.

After that, the plate was sealed and centrifuged briefly. Next based on the planned index adapter combination 5ul of each i7 and 5ul of each i5 were added into the respective samples and were mixed well through pipetting after the plate was sealed and centrifuged once again briefly. Next, the plate was placed to run the PCR machine with the bead linked transposomes (BLT) PCR program and the temperature setting was 68^0C for 3 minutes. Then 98^0C for 3 minutes followed by 5 Cycles of 98^0C (45sec), and 62^0C (0.3 seconds). And the terminated reaction was 68^0C for 60 seconds and the final temperature. Once the PCR was done the plate was taken out and placed into chill on ice, until the next step. After step three was done the PCR plate was centrifuged briefly and placed on the magnetic stand for 5 minutes until the liquid was clear and the new palate (Kingfisher plate) was prepared and transferred into 45ul of each library from step 3 PCR plate into the new plate and then added 40ul of nuclease-free water into each sample after that 45ul of SPB (Sample Purification Beads) was added and mixed well using pipetting. The plate of the sample was placed into the magnetic stand

for 5 minutes until the solution was clear, once again another new deep well plate was labelled, and 15ul of SPB was added to the plate and 125ul was transferred to the new plate containing 15ul of SPB and mixed well.

Then samples were incubated at room temperature for 5 minutes. Once this step was completed, the sample plate was placed on a magnetic stand for 5 minutes until the solution became clear. Subsequently, the supernatant was carefully discarded to avoid disturbing the beads. The next step involved washing the sample twice while it remained on the magnetic stand. An 80% ethanol (EtOH) solution was prepared, and 200ul of this prepared EtOH was added to each sample on the plate. After incubating for 30 seconds during each wash, the supernatants were discarded while the plate was still on the magnetic rack. Then waited for 5 minutes until the samples were completely dry.

Following this, the plate was removed from the magnetic stand, and 32ul of Resuspension Buffer (RSB) was added directly to the beads in each sample on the plate. After thorough mixing, the samples were incubated for 2 minutes. The plate was placed back on the magnetic stand until the liquid became clear, which typically took about 2 minutes. While waiting for the samples, new PCR tubes were labeled with the respective sample names and dates. Finally, 30ul of the final library was transferred into the newly labeled PCR tubes."

3.10 DNA concentration of library preparation and alignment Bioanalyzer

This method of DNA concentration library preparation was done same as described in sub-section 3.8.