A BLOWFLY-DERIVED DNA APPROACH TO ASSESS DIVERSITY OF TROPICAL MAMMALS

LEE PING SHIN

THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2016
UNIVERSITI MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **LEE PING SHIN**

Registration/Matric No: **SHC140001**

Name of Degree: **Doctor of Philosophy (ECOLOGY AND BIODIVERSITY)**

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"): **A BLOWFLY-DERIVED DNA APPROACH TO ASSESS DIVERSITY OF TROPICAL MAMMALS**

Field of Study: **ECOLOGY AND BIODIVERSITY**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature
Date

Subscribed and solemnly declared before,

Witness’s Signature
Date

Name:
Designation:
ABSTRACT

Most tropical mammal species are threatened or data-deficient. Data collection is impeded by traditional monitoring approaches which can be laborious, expensive and struggle to detect cryptic diversity. Monitoring approaches using mammal DNA derived from invertebrates, in particular blowfly-derived DNA, have recently been suggested as alternatives to traditional field methods. As a step towards development of blowfly-derived DNA as an effective method for mammal monitoring in Peninsular Malaysia, my research objectives are: (i) to determine the persistence period of amplifiable mammal mtDNA in blowfly guts; (ii) to design and test primers that can selectively amplify mammal COI DNA mini-barcodes in the presence of high concentrations of blowfly DNA; (iii) to determine the dispersal range of blowflies in forests; (iv) to calibrate the performance of blowfly-derived mammal DNA at generating species inventories and richness estimates against traditional methods: cage trapping, mist netting, hair trapping and scat collection in the field; (v) to calibrate the performance of blowfly-derived DNA at generating species inventories and richness estimates against the most popular method to assess megafaunal diversity - camera trapping in the field.

The persistence period of amplifiable mammal DNA in blowfly guts was in the range of 24 h to 96 h post-feeding indicating the need for retrieving flies within 24 h of capture to detect mammal mtDNA in sufficient quantity and quality. A short (205 bp) mammal COI DNA mini-barcode, suitable for high-throughput sequencing, could distinguish most mammal species (including separating dark taxa). The daily dispersal range of Chrysomya spp. blowflies suggests that individuals contributing blowfly-derived DNA would have been within 3 km the location where blowflies were collected over 24 h period. The field study in Ulu Gombak Forest Reserve that calibrated the performance of blowfly-derived mammal DNA against traditional methods (cage traps, mist nets, hair traps, and scat collection) over 1,440 trap days revealed that blowfly traps and mist
nets detected the joint highest number of wild mammal species (6 species), and only one species was detected by multiple methods. Compared to traditional field methods, blowfly-derived DNA detected both volant and non-volant species from wider body size ranges. In Tembat Forest Reserve with megafauna, calibration of blowfly-derived DNA against camera traps, the most popular method in megafaunal diversity assessment, blowfly-derived DNA detected more species (n=11) than camera traps (n=9) across 1,200 trap days, with only one species detected by both methods. With further calibration, blowfly-derived DNA may join the list of traditional field methods.
ABSTRAK

Kebanyakan spesies mamalia terancam atau tidak diketahui maklumatnya. Pengumpulan maklumat dibuntui oleh pembatasan kaedah pemantauan tradisional yang memerlukan pelaburan tenaga dan masa yang banyak, mahal dan mengalami kesulitan dalam mengesan kepelbagaian samar. Kaedah pemantauan menggunakan DNA mamalia yang diperoleh daripada invertebrat, terutama DNA mamalia perolehan langau, telah dicadangkan sebagai alternatif terhadap kaedah pemantauan tradisional. Sebagai langkah dalam pembangunan DNA mamalia perolehan langau menjadi kaedah pemantauan mamalia yang efektif di Semenanjung Malaysia, objektif kajian saya adalah: i) untuk menentukan jangka pengekalan mtDNA mamalia yang boleh diamplifikasi daripada usus lalat; ii) untuk mereka bentuk dan menguji primer yang boleh mengamplifikasi secara memilih kodbar DNA mamalia dalam DNA langau yang berkepekatan tinggi; iii) untuk menentukan jarak penyebaran langau di hutan; iv) untuk menentukurkan prestasi DNA mamalia perolehan langau di lapangan dalam menjana inventori dan anggaran spesies berbanding dengan kaedah pemantauan tradisional: perangkap sangkar, jaring kabus, perangkap bulu dan pengutipan tahi; v) untuk menentukurkan prestasi DNA mamalia perolehan langau di lapangan dalam menjana inventori dan anggaran spesies berbanding dengan kaedah yang paling popular dalam menilai kepelbagaian megafauna - perangkap kamera. Jangka pengekalan DNA mamalia yang boleh diamplifikasi daripada usus langau adalah antara 24 hingga 96 jam selepas pemakanan. Hasil kajian ini menunjukkan keperluan mendapatkan langau dalam masa 24 jam selepas penangkapan supaya dapat mengesan mtDNA mamalia yang mempunyai kuantiti dan kualiti yang cukup. Kodbar-mini DNA COI mamalia yang pendek (205 bp) adalah sesuai untuk penjujukan bagi daya pemprosesan yang tinggi dan mampu membezakan kebanyakan spesies mamalia (termasuk membezakan taksa samar). Penentuan jarak penyebaran langau Chrysomya spp. mencadangkan bahawa
individu yang menyumbang kepada DNA perolehan langau boleh berada di dalam linkungan 3 km daripada lokasi di mana langau dikumpul dalam jangka masa 24 jam. Kajian di lapangan yang menentukurkan prestasi DNA mamalia perolehan langau di Hutan Simpan Ulu Gombak berbanding dengan kaedah pemantauan tradisional (perangkap sangkar, jaring kabus, perangkap bulu dan pengutipan tahi) daripada jangka perangkap sebanyak 1,440 hari menunjukkan perangkap langau dan jaring kabus mengesan bilangan spesies mamalia liar yang paling banyak (6 spesies), dan hanya satu spesies yang sama dikesan oleh semua kaedah. Penentukuran prestasi DNA mamalia perolehan langau berbanding dengan kaedah yang paling popular dalam menilai kepelbagaian megafauna - perangkap kamera di Hutan Simpan Tembat yang mengandungi megafauna menunjukkan bahawa DNA mamalia perolehan langau mengesan lebih banyak spesies (n=11) daripada perangkap kamera (n=9) daripada jangka perangkap sebanyak 1,200 hari, dan hanya satu spesies yang sama dikesan oleh kedua-dua kaedah. Berbanding dengan kaedah pemantauan tradisional, DNA mamalia perolehan langau mengesan spesies mamalia jenis terbang dan tidak terbang daripada julat saiz badan yang lebih luas. Melalui penentukuran yang lebih mendalam, DNA mamalia perolehan langau mungkin menyertai kaedah-kaedah tradisional di lapangan yang digunakan untuk menilai kepelbagaian mamalia di hutan-hutan tropika.
ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to everyone who provided me possibility to complete this research and dissertation. Several researchers and helpers have made contributions to the chapters of the thesis and it is necessary to recognise their contributions. I am supported by grants from the Nagao Natural Environment Foundation Japan and Ministry of Higher Education Fundamental Research Grant Scheme (FP042-2014A) as well as scholarship from Skim Basiswa University of Malaya to complete the dissertation.

Special gratitude to my supervisor, Dr. John James Wilson for providing me invaluable guidance throughout my study especially in planning, analysis and writing. His patience and advice in my research and writing helped me overcome challenges in completing research and thesis. I would also like to thank my labmates or colleagues (Sing Kong Wah, Khairunnisa Syaripuddin, Kong Wye Lup, Brandon Mong, Jisming Shi, Lim Voon Ching, Cheah Men How, Thary Gazi, Aik Hean, Kai Ren) for their generous help in my fieldwork and labwork.

In addition, I would also like to thank my collaborators for my research - Reuben Clement Gopalasamy and his team of RIMBA (myrimba.org) for their generous technical support and suggestions for the fieldwork and manuscript, and Dr. Gan Han Ming for providing helpful technical support in sequencing services and suggestions for manuscript writing, as well as Dr. Wee Suk Ling and her team for providing support in fieldwork and manuscript writing. Thanks also to the owners of the rubber estate to allow me permission to access for conducting my research.

At last but not least, million thanks to my fieldwork assistants (Lee Yoon Hin, Tee Ah Jee, Lee Kah Hau, Lee Sueh Loong, Lee Sueh Wen, Ng Eain Yi, Min) for their endless support and guidance to complete my thesis.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRAK</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF SYMBOLS AND ABBREVIATIONS</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2: LITERATURE REVIEW</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Mammal diversity</td>
<td>4</td>
</tr>
<tr>
<td>2.2 Traditional methods for mammal monitoring</td>
<td>5</td>
</tr>
<tr>
<td>2.3 Invertebrate-derived DNA as a mammal monitoring tool</td>
<td>8</td>
</tr>
<tr>
<td>2.4 The status of tropical mammal diversity assessment in Peninsular Malaysia</td>
<td>10</td>
</tr>
<tr>
<td>CHAPTER 3: PERSISTENCE PERIOD OF AMPLIFIABLE MAMMAL MTDNA IN BLOWFLY GUTS</td>
<td>13</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>13</td>
</tr>
<tr>
<td>3.2 Materials and methods</td>
<td>14</td>
</tr>
<tr>
<td>3.3 Results</td>
<td>15</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>17</td>
</tr>
<tr>
<td>CHAPTER 4: DESIGN AND TESTING OF PRIMERS FOR AMPLIFICATION OF MAMMAL DNA MINI-BARCODES FROM BLOWFLY GUTS</td>
<td>19</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>19</td>
</tr>
<tr>
<td>4.2 Materials and methods</td>
<td>21</td>
</tr>
<tr>
<td>4.3 Results</td>
<td>26</td>
</tr>
</tbody>
</table>
### CHAPTER 5: BLOWFLY DISPERSAL RANGE AND IMPLICATIONS FOR BLOWFLY-DERIVED DNA IN TROPICAL MAMMAL MONITORING

5.1 Introduction 33  
5.2 Materials and methods 35  
5.3 Results 39  
5.4 Discussion 39  

### CHAPTER 6: FIELD-CALIBRATION OF BLOWFLY-DERIVED DNA AGAINST TRADITIONAL FIELD METHODS

6.1 Introduction 43  
6.2 Materials and methods 44  
6.3 Results 50  
6.4 Discussion 54  

### CHAPTER 7: FIELD-CALIBRATION OF BLOWFLY-DERIVED DNA AGAINST CAMERA TRAPS

7.1 Introduction 59  
7.2 Materials and methods 60  
7.3 Results 64  
7.4 Discussion 67  

### CHAPTER 8: CONCLUSION AND RECOMMENDATION 70

REFERENCES 73  
LIST OF PUBLICATIONS AND PAPERS PRESENTED 97  
APPENDIX 98
LIST OF FIGURES

Figure 3.1 PCR amplification of beef mtDNA from blowfly guts post feeding (8 h – 120 h).

Figure 4.1 Number of 16S rRNA and COI sequences publicly available on GenBank for the 246 mammal species (excluding Homo sapiens) found in Peninsular Malaysia.

Figure 4.2 A) Relative positions of mammal primers on the COI barcode region. B) The binding sites of the primers Uni-Mini-bar F and RonPing (reverse complement).

Figure 4.3 Baited blowfly trap for mammal monitoring.

Figure 4.4 Number of mismatches between the RonPing primer sequence (19 bp) and 1472 mammal COI sequences from GenBank.

Figure 4.5 Comparison of amplification success using primer pairs: Uni-Mini-bar F/ RonPing and Uni-Mini-bar F/ Uni-Mini-bar R for Artiodactyla (n=3), Carnivora (n=2), Chiroptera (n=28), Soricomorpha (n=1), Rodentia (n=5), and Scandentia (n=1).

Figure 4.6 Examples of fragments of mammal sequences amplified using Uni-Mini-bar F/ RonPing primers from mammal/ C. megacephala DNA mixtures.

Figure 5.1 The location of Kalumpang, Selangor showing the experimental design for blowfly recapture.

Figure 5.2 The number of blowflies recaptured based on number of days since release and distance from the release point (1-5 km).

Figure 6.1 The location of Ulu Gombak Forest Reserve showing the location of two transects and the experimental design for comparing five different field methods.

Figure 6.2 Species detected by different field methods at Ulu Gombak Forest Reserve, including species unique to each method and species shared across different methods.

Figure 6.3 Rarefaction curves showing expected species richness (using chao1 with number of randomizations = 100) and sampling completeness ratios (in brackets) for each field method at Ulu Gombak Forest Reserve.

Figure 7.1 The location of Tembat Forest Reserve showing the locations of twenty trap stations each with a blowfly trap and camera trap.

Figure 7.2 Photographic evidence of nine mammal species detected by camera traps at Tembat Forest Reserve, Terengganu.
Figure 7.3  Species detected by different field methods at Tembat Forest Reserve.

Figure 7.4  Rarefaction curves for expected species richness (using chao1 with number of randomizations = 100) and sampling completeness ratios (in brackets) for each field method at Tembat Forest Reserve.
LIST OF TABLES

Table 2.1  Comparison of different field methods including targeted mammal groups, advantages, disadvantages and cost.  6

Table 4.1  List of 41 mammal species from Malaysia for comparison of amplification success between primer pairs, Uni-Mini-bar F/ RonPing and Uni-Mini-bar F/ Uni-Mini-bar R and BOLD Process ID numbers for those sequenced after troubleshooting.  22

Table 5.1  Summary of MRR studies on adult fly dispersal including the marking techniques, study locations, and dispersal ranges.  34
LIST OF SYMBOLS AND ABBREVIATIONS

~  Approximate
\geq  Greater than or equal to
<  Less than
>  More than
%  Percentage
\pm  Plus-minus
BOLD  Barcode of Life Datasystems
COI  *Cytochrome c oxidase subunit I*
ddH$_2$O  Double distilled water
DNA  Deoxyribonucleic acid
e.g.  Latin phrase *exempli gratia* (for example)
et al.  Latin phrase *et alia* (and other)
g  Gram
h  Hour
ha  Hectare
iDNA  Invertebrate-derived DNA
i.e.  Latin phrase *id est* (that is)
m  Meter
mL  Millilitre
mtDNA  Mitochondrial deoxyribonucleic acid
MRR  Mark-release-recapture
PCR  Polymerase chain reaction
SD  Standard deviation
sp.  Species (singular)
spp.  Species (plural)
UMKL  University of Malaya, Kuala Lumpur
vs.  Versus
CHAPTER 1: INTRODUCTION

Most tropical mammals remain data-deficient or disproportionately threatened according to the Global Mammal Assessment (Schipper et al., 2008). Data collection is impeded by traditional monitoring approaches which can be laborious (Campbell et al., 2011), expensive (Trolle & Kery, 2003; Rowcliffe et al., 2008), and struggle to detect cryptic species which are not always morphologically distinct and more often recognised by molecular techniques (Bickford et al., 2006; Ceballos & Ehrlich, 2008). Considering that the current monitoring approaches are challenged by ethics (e.g., lack of guidelines or consideration on the effects of research and trapping on animals; Powell & Proulx, 2003), precision, and accuracy, new approaches are urgently needed.

Approaches using mammal DNA derived from invertebrates (iDNA) have recently been suggested as alternatives for large-scale mammal monitoring. Early studies have suggested that such approaches have potential for detecting and identifying rare, cryptic, endangered and new mammal species (Calvignac-Spencer et al., 2012; Schnell et al., 2012; Calvignac-Spencer et al., 2013; Schnell et al., 2015). Yet, certain key parameters still require evaluation to determine the feasibility of iDNA as an effective monitoring tool.

Establishing the persistence period of mammal DNA in invertebrate guts is essential for designing standardised trapping or sampling protocols to maximise iDNA detection. However, the post-feeding persistence period of mammal DNA in the guts of many invertebrates has never been investigated (Calvignac-Spencer et al., 2013; Schnell et al., 2015).

“Universal” primers have been designed for amplification by PCR of mammal cytochrome c oxidase subunit I (COI) DNA barcodes (Ivanova et al., 2012) but a short target of <250 bp is thought to be required for ingested DNA, which is likely to be partially degraded and difficult to amplify (Calvignac-Spencer et al., 2013) (but see
Gariepy et al., 2012; Lassen et al., 2012). Chapter 3 and Chapter 4 have been published together in the article: Ping-Shin Lee, Kong-Wah Sing, & John-James Wilson. 2015. Reading mammal diversity from flies: the persistence period of amplifiable mammal mtDNA in blowfly guts (Chrysomya megacephala) and a new DNA mini-barcode target. PLOS ONE, 10(4): e0123871.

The dispersal and range sizes of invertebrates have also been suggested to contribute towards uncertainties over the use of iDNA (Calvignac-Spencer et al., 2013; Schnell et al., 2015). Leeches exhibit little movement and can consequently provide reliable information regarding the location and habitat preferences of target species (i.e., nearby the location where the leech was collected) (Schnell et al., 2015). Conversely, blowflies presumably have wider range sizes resulting in greater uncertainties about the actual location where target species DNA was sampled by the blowflies (Schnell et al., 2015). Knowledge of blowfly dispersal distances is essential in addressing uncertainties regarding a blowfly-derived DNA approach and providing estimates of the location a mammal species was sampled, relative to the location where the blowflies were collected (Calvignac-Spencer et al., 2013; Schnell et al., 2015). Yet, there is no data available on blowfly dispersal ranges in the tropics. Chapter 5 is being prepared for submission as: Ping-Shin Lee, Suk-Ling Wee, & John-James Wilson. 2016. Blowfly dispersal range and implication for blowfly-derived DNA in mammal monitoring. Tropical Ecology.

Despite the increasing interest in iDNA as a tool for mammal diversity assessments, field calibrations of the performance of iDNA against traditional methods have yet to be conducted (Schnell et al., 2015). A wide variety of field methods have been used to assess mammal diversity in tropical forests, against which comparisons should be performed. Frequently used methods include: live trapping, including both cage traps (Hanif-Ridzuan et al., 2010; Madinah et al., 2011), mist nets and harp traps
(Kingston et al., 2003; Sing et al., 2013); camera traps (Clements, 2013; Hedges et al., 2015a); indirect signs such as tracks or scat (Daim, 2002), interviews with local communities (Sharma et al., 2005); direct observations by researchers (Syakirah et al., 2000; Jayaraj et al., 2013); and hair traps (Castro-Arellano et al., 2008; Hedges et al., 2015b). Chapter 6 and Chapter 7 have been published together in the article: Ping-Shin Lee, Han-Ming Gan, Gopalasamy-Reuben Clements, & John-James Wilson. 2016. Field calibration of blowfly-derived DNA against traditional methods for assessing mammal diversity in tropical forests. Genome, 10.1139/gen-2015-0193.

Blowfly-derived DNA has promising potential as iDNA for mammal diversity assessments (Calvignac-Spencer et al., 2012; Calvignac-Spencer et al., 2013) because blowflies are distributed in all habitats (Norris, 1965) and exhibit broad host preferences (Calvignac-Spencer et al., 2012; Azwandi et al., 2013; Calvignac-Spencer et al., 2013; Schnell et al., 2015). As a step towards development of blowfly-derived DNA as an effective method for mammal monitoring in biodiversity hotspots of Peninsular Malaysia, I will address the following objectives in this thesis:

i. To determine the persistence period of PCR amplifiable mammal DNA in blowflies.

ii. To design and test primers that can selectively amplify mammal DNA mini-barcodes in the presence of high concentrations of blowfly DNA.

iii. To determine the dispersal range of blowflies in forests.

iv. To calibrate the performance of blowfly-derived DNA at generating species inventories and richness estimates against traditional methods in the field.

v. To calibrate the performance of blowfly-derived DNA at generating species inventories and richness estimates against the most popular method to assess megafaunal diversity - camera trapping - in the field.
CHAPTER 2: LITERATURE REVIEW

2.1 Mammal diversity

A diversity of mammal species is crucial for sustaining normal ecosystem function. From direct roles in seed dispersal (Ostfelda & Keesing, 2000), nutrient recycling (Rudinow Saet nan, 2000), and plant biomass control (Keesing, 2000), mammal species have recently been associated with maintaining the capacity of tropical forests to store carbon in the face of global climate threats (Brodie & Gibbs, 2009; Jansen et al., 2010). Seed dispersal, especially by large mammals, plays a key role in the spatial pattern of tree recruitment and survival in tropical forests (Effiorn et al., 2013). Small mammals, such as bats (Jones et al., 2009) and deer mouse (Leis et al., 2008), are also particularly useful indicators of ecosystem health as their responses to both long- and short-term disturbance can be measured in a predictable manner (Olson & Brewer, 2003; Leis et al., 2008).

The global total of wild mammal species stands at 5,502 according to the International Union for Conservation of Nature (IUCN) (IUCN, 2016) and includes 1,098 mammal species from Southeast Asia; 20% of the global total (IUCN, 2016). Two hundred and twenty-two mammal species have been recorded from the Southeast Asian region of Peninsular Malaysia alone (DWNP, 2010).

The class Mammalia is divided into two extant subclasses - Prototheria and Theria, representing 29 orders and 153 families (Wilson & Reeder, 2005). Some orders include many families (e.g., Primates is an order containing 15 families), while other orders include only a single family (e.g., Proboscidae is an order containing only one extant family, Elephantidae) (Wilson & Reeder, 2005). Mammals are distributed globally, showing diversity in size, form, locomotion, and diet, as well as high adaptability to their surroundings (Feldhamer et al., 2007). For a group of relatively
large, charismatic animals, there are still surprisingly high rates of species discovery, especially in South America and Southeast Asia (Reeder et al., 2007), and most tropical mammal species remain data deficient or disproportionately threatened according to the Global Mammal Assessment (Schipper et al., 2008).

2.2 Traditional methods for mammal monitoring

Monitoring is crucial for conservation efforts, providing data on species distribution patterns and population size estimates (Thomsen & Willerslev, 2015). Terrestrial mammal diversity has traditionally been assessed using: live traps, including cage traps (Hanif-Ridzuan et al., 2010; Madinah et al., 2011), mist nets and harp traps (Kingston et al., 2003; Sing et al., 2013); killing traps, including pitfall traps (Bury & Corn, 1987; Handley & Kalko, 1993) and snap traps (Powell & Brooks, 1981; Batzli et al., 1983); camera traps (Clements, 2013; Hedges et al., 2015a); indirect signs such as tracks or scat (Daim, 2002); interviews (Sharma et al., 2005); direct observations (Syakirah et al., 2000; Jayaraj et al., 2013); and hair traps (Castro-Arellano et al., 2008; Hedges et al., 2015b). Each field method listed above differs in terms of the targeted mammal groups and costs, with advantages and disadvantages for each method (summarised in Table 2.1).

Species monitoring has traditionally relied on physical identification of species based on distinct morphological characters observed during visual inspection of “captured” individuals (Thomsen & Willerslev, 2015). However, diversity assessment is hampered by limitations of existing monitoring approaches. Field trapping techniques vary in efficiency (Lambert et al., 2005) and are often a biased representation of diversity (Fontúrbel, 2010; Torre et al., 2010). Identification of animal signs is laborious, requiring the input of specialists over an extended time period (Campbell et al., 2011), and can be imprecise (Davison et al., 2006; Mumma et al., 2014). Expensive
### Table 2.1: Comparison of different field methods including targeted mammal groups, advantages, disadvantages and cost.

<table>
<thead>
<tr>
<th>Field method</th>
<th>Targeted mammal groups</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cage traps</td>
<td>Rodents, insectivores (Catling et al., 1997)</td>
<td>Operates 24 h without supervision (Catling et al., 1997)</td>
<td>Difficult to standardize (varying dimensions, baited/un-baited and deployment of location effects) (Fontarbel, 2010; Torre et al., 2010)</td>
<td>Medium (Garden et al., 2007)</td>
</tr>
<tr>
<td>Harp traps</td>
<td>Insectivorous bats (Kingston et al., 2003)</td>
<td>Does not require constant monitoring (Tidemann &amp; Woodside, 1978)</td>
<td>Bulky and not easy to transport (Tidemann &amp; Woodside 1978)</td>
<td>High (Tidemann &amp; Woodside 1978)</td>
</tr>
<tr>
<td>Pitfall traps</td>
<td>Small-bodied mammals (e.g., shrews; Spencer &amp; Pettus, 1966)</td>
<td>Multiple mammals may be trapped in a single pitfall trap (Powell &amp; Proulx, 2003)</td>
<td>Trapped mammals may be subject to predation by larger mammals or may be killed by each other when trapped together; Must be checked multiple times daily to maximize survival of mammals (Powell &amp; Proulx, 2003)</td>
<td>Low (Garden et al., 2007)</td>
</tr>
<tr>
<td>Snap traps</td>
<td>Small to medium-sized mammals (Powell &amp; Proulx, 2003)</td>
<td>Performance criteria setting is easier than that for restraining traps because unconsciousness and death are relatively easy to define objectively, compared to the injury, anxiety, and hardship that may be experienced by restrained animals (Powell &amp; Proulx, 2003)</td>
<td>Trap placement may greatly affect probabilities of trapping non-targeted animals that are too large to be quickly killed by snap traps (Powell &amp; Proulx, 2003)</td>
<td>Medium (Perry et al., 1996)</td>
</tr>
<tr>
<td>Camera traps</td>
<td>Medium to large-bodied mammals (Bernard et al., 2013)</td>
<td>Effective in detecting species rarely recorded from live traps or direct observations (e.g., Hose’s civet, <em>Diplogale hosei</em>; Bernard et al., 2013)</td>
<td>May under-represent species with specific habitats and unable to distinguish closely related species (e.g., muntjac and mouse-deers; Bernard et al., 2013)</td>
<td>High (Sanderson &amp; Trolle, 2005)</td>
</tr>
</tbody>
</table>
Table 2.1, continued: Comparison of different field methods including targeted mammal groups, advantages, disadvantages and cost.

<table>
<thead>
<tr>
<th>Field method</th>
<th>Targeted mammal groups</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect signs</td>
<td>Medium to large-bodied ground dwelling mammals (Catling et al., 1997)</td>
<td>Effective in detecting species inhabiting open areas (e.g., otters and ungulates; Catling et al., 1997)</td>
<td>Imprecise in species identification (Davison et al., 2006; Mumma et al., 2014) Accuracy and precision are dependent on field conditions and expertise of identifiers (Silveira et al., 2003)</td>
<td>Low (Garden et al., 2007)</td>
</tr>
<tr>
<td>Interviews</td>
<td>Medium to large-bodied mammals (Mohd-Azlan et al., 2013)</td>
<td>No equipment required and can provide important collateral data (Mohd-Azlan et al., 2013)</td>
<td>Less reliable due to varying survey design and bias produced by respondents such as different recalling ability of respondents for different periods of time (Meijaard et al., 2011) (e.g., most villagers interviewed could not recall how many tigers have been killed due to livestock depredation by tigers; Sharma et al., 2005)</td>
<td>Low (Mohd-Azlan et al., 2013)</td>
</tr>
<tr>
<td>Direct observations</td>
<td>Medium to large-bodied mammals that are readily observable (Giese et al., 2001)</td>
<td>Can confirm species presence directly and effectively detect rare and endangered species (e.g., black-footed ferrets, Mustela nigripes; Giese et al., 2001)</td>
<td>The detection success relies much on the protocols used and expertise of identifiers (Hoppe-Dominik et al., 2011; Roberts, 2011)</td>
<td>Low (Garden et al., 2007)</td>
</tr>
<tr>
<td>Hair traps</td>
<td>All ground-dwelling mammals (Stanton &amp; Anderson, 1998; Castro-Arellano et al., 2008)</td>
<td>Portable (Castro-Arellano et al., 2008); Reliable in identifying species through genetic analysis (Beja Pereira et al., 2009; Mullins et al., 2009)</td>
<td>Different trap types used could influence the detection success (Castro-Arellano et al., 2008) such as the use of hair-traps with mechanical devices like barb wire or attached glue-like substances that rarely collect high proportions of hairs with roots (Valderrama et al., 1999; Ebert et al., 2010).</td>
<td>Low (Castro-Arellano et al., 2008)</td>
</tr>
</tbody>
</table>
camera traps cannot identify individuals to species lacking easily observed diagnostic markings (Trolle & Kery, 2003; Rowcliffe et al., 2008). These methods also face taxonomic challenges posed by cryptic species, which are not always morphologically distinct and are more easily recognised by molecular techniques (Bickford et al., 2006; Ceballos & Ehrlich, 2008).

The “best” field methods should be unbiased, precise (replicates produce similar findings), and efficient - a relatively small number of individuals or samples is needed for accurate estimations of “true” species richness or other biodiversity indices (Gotelli & Colwell, 2011). Whichever field methods are used, they rarely produce representative or consistent samples due to observer error (Nichols et al., 2000), species rarity (Dettmers et al., 1999), or varying environmental conditions (Gu & Swihart, 2004). The use of non-parametric species richness estimators such as Jackknife (Burnham & Overton, 1979) and Chao (Chao, 1984; Chao, 1987) has been recommended to enable more accurate estimation as they correct for biases due to sampling effort and species rarity (Iknayan et al., 2014).

2.3 Invertebrate-derived DNA as a mammal monitoring tool

Invertebrate-derived DNA (iDNA) - vertebrate genetic material ingested by invertebrates, is emerging as a powerful tool for mammal monitoring (Calvignac-Spencer et al., 2012; Schnell et al., 2012; Calvignac-Spencer et al., 2013; Schnell et al., 2015). The potential of invertebrates that feed on vertebrate hosts, for example, leeches (Schnell et al., 2012) and blowflies (Calvignac-Spencer et al., 2012), as diversity assessment tools have been demonstrated through opportunistic sampling (Calvignac-Spencer et al., 2013; Schnell et al., 2015). Such approaches have the potential for detecting and identifying rare, cryptic, endangered and newly discovered mammal species (Calvignac-Spencer et al., 2012; Schnell et al., 2012; Calvignac-Spencer et al.,
2013; Schnell et al., 2015). Two rare species - Annamite striped rabbit \textit{(Nesolagus timminsi)} and Truong Son muntjac \textit{(Muntiacus truongsonensis)} - as well as a “cryptic” species, small-toothed ferret-badger \textit{(Melogale moschata)}, were detected from mammal blood recovered from leeches in a Vietnamese rainforest (Schnell et al., 2012).

In most terrestrial environments, the carcasses, droppings, sweat, and urine of vertebrates can be energy or nutrient sources for invertebrates. The ingested DNA is likely to be partially degraded and difficult to amplify (Calvignac-Spencer et al., 2013). Consequently, it is necessary to target a short DNA fragment for successful iDNA detection, such seems sensible to target multi-copy mtDNA fragments (Calvignac-Spencer et al., 2013). Previous studies have used \textit{16S rRNA} and \textit{12S rRNA} for detection of mammal mtDNA from blowflies (Calvignac-Spencer et al., 2012), \textit{16S rRNA} for detection of mammal mtDNA from leeches (Schnell et al., 2012), \textit{cytochrome b} for detection of vertebrate mtDNA from mosquitoes and sandflies to determine feeding patterns or blood sources (Abbasi et al., 2009; Bataille et al., 2012), and \textit{cytochrome c oxidase I (COI)} for detection of vertebrate mtDNA from ticks and tsetse flies to study their feeding preferences (Muturi et al., 2011; Gariepy et al., 2012).

Blowflies (Order: Diptera; Family: Calliphoridae) have shown promising potential as sources of mammal DNA, in part due to their presence in all habitats (Norris, 1965) and broad host preferences as saprophagous and coprophagous generalists (Calvignac-Spencer et al., 2012; Azwandi et al., 2013; Calvignac-Spencer et al., 2013; Schnell et al., 2015). \textit{Chrysomya bezziana} has been reported feeding on mammal species from the orders Artiodactyla (7 spp.), Carnivora (6 spp.), Diprotodontia (2 spp.), Perissodactyla (4 spp.), Proboscidea (1 sp.), and Primates (1 sp.) at a zoo in Malaysia (Spradbery & Vanniasingham, 1980). \textit{Lucilia sericata} is found in Asia, North America, and Europe (James, 1947; Picard & Wells, 2010) and has been reported to feed on wounds of sheep, cats, a dog, and a horse in Israel (Schnur et al.,
Blowfly-derived DNA may present advantages over sources of iDNA, such as from leeches, which are habitat restricted (Schnell et al., 2012), or from mosquitoes and tsetse flies, which have narrow host preferences (Lyimo & Ferguson et al., 2009; Muturi et al., 2011). Blowflies also potentially concentrate faecal DNA while feeding coprophagously (Stenglein et al., 2010; Calvignac-Spencer et al., 2013).

Calvignac-Spencer and colleagues (2012) first demonstrated the potential of DNA sequencing of blowfly-derived DNA by sampling flies opportunistically upon immediate arrival at bait. However, there are uncertainties regarding the application of blowfly-derived DNA as a standardised method in the field (Calvignac-Spencer et al., 2013; Schnell et al., 2015). For example, the interval between a blowfly feeding and looking for its next meal could affect the success of mammal DNA detection from blowfly guts (Calvignac-Spencer et al., 2013). How far blowflies carry mammal DNA away from the spot they sampled mammal tissues or faeces is unknown. The few studies of blowfly dispersal capabilities (from temperate and subtropical regions - Braack & Retief, 1986; Smith & Wall, 1998; Tsuda et al., 2009) suggested individuals travel 100-2,400 m per day, but there is no data available on blowfly dispersal ranges in the tropics.

2.4 The status of tropical mammal diversity assessment in Peninsular Malaysia

Peninsular Malaysia is a global biodiversity hotspot and part of a megadiverse country with a high number of endemic mammal species (Davison & Zubaid, 2007; Davidson et al., 2014). Studies of mammal diversity in Peninsular Malaysia have been conducted for 113 years (Corbet & Hill 1992; Francis 2008), yet the total number of mammal species reported from Peninsular Malaysia is still increasing with 22 species added to the records during the last 30 years (Davison & Zubaid, 2007). This is especially the case for several taxonomically difficult groups (e.g., Crocidura;
Cynopterus; Glyphotes; Haeromys; Myotis; Petaurillus; Davison & Zubaid, 2007). Habitat loss, fragmentation and poaching have put many of the mammal species in Peninsular Malaysia at risk of local extinction (Sodhi et al., 2008). This includes the Asian Elephant (Elephas maximus) and Malayan Tapir (Tapirus indicus), with high rates of displacement in the 1990s mainly due to habitat loss (Zahari et al., 2001). These examples further highlight the need to study mammal diversity in global biodiversity hotspots such as Peninsular Malaysia, where the current knowledge on mammal diversity remains relatively poor and mammal species are at risk of extinction.

Two hundred and twenty-two mammal species have been recorded in Peninsular Malaysia (DWNP, 2010), but only 31% of the 222 species recorded for Peninsular Malaysia have been assessed for conservation risk status (DWNP, 2010). The main criteria used in the IUCN assessment are the population size, the extent of occurrence, and the area of occurrence (IUCN, 2016). In Peninsular Malaysia, only five species have been assessed for conservation risk status using population estimates: Sumatran Rhinoceros (Dicerorhinus sumatrensis), Asian Elephant (E. maximus), Malayan Tapir (T. indicus), Tiger (Panthera tigris), and Gaur (Bos gaurus) (DWNP, 2010). That such data is lacking for the majority of mammal species highlights the limitations in the current means of data collection and the need for more efficient mammal monitoring approaches in Peninsular Malaysia.

In the tropical forests of Peninsular Malaysia, mammal diversity has been commonly assessed using: i) live trapping, including both cage traps (Hanif-Ridzuan et al., 2010; Madinah et al., 2011), and mist nets and harp traps (Kingston, 2003; Sing et al., 2013), ii) camera traps (Kawanishi, 2002; Azlan & Sharma 2003; Numata et al., 2005; Clements, 2013; Hedges et al., 2015a), iii) indirect signs such as tracks (Flynn & Abdullah, 1984) or scat (e.g. species identification of Asiatic golden cat and dhole based on mtDNA analysis; Kawanishi & Sunquist, 2008), iv) interviews with local community
members, such as farmers and villagers (Topani, 1990; Sharma et al., 2005), and v) direct observations by researchers (Syakirah et al., 2000; Jayaraj et al., 2013). There are no published studies of hair traps being employed in Peninsular Malaysia but this method has been used extensively in other regions (e.g., Foran et al., 1997; Woods et al., 1999; Beier et al., 2005; Mowat, 2006; Schmidt & Kowalczyk, 2006; Bremner-Harrison et al., 2006; Harrison, 2006; Castro-Arellano et al., 2008). Recently a few studies in Peninsular Malaysia have simultaneously tested the effectiveness of different field methods in detecting different groups of mammal species (Syakirah et al., 2000; Jayaraj et al., 2012; Tingga et al., 2012; Jayaraj et al., 2013).
CHAPTER 3: PERSISTENCE PERIOD OF AMPLIFIABLE MAMMAL MTDNA IN BLOWFLY GUTS

3.1 Introduction

Monitoring approaches using mammal DNA derived from invertebrates are emerging as cost- and time-effective alternatives to traditional methods (Calvignac-Spencer et al., 2012; Schnell et al., 2012; Calvignac-Spencer et al., 2013; Schnell et al., 2015). These methods can provide accurate identification of mammal species, require the least ecological and taxonomic expertise, and yet have the potential to detect rare and cryptic species (Calvignac-Spencer et al., 2012; Schnell et al., 2012). However, certain key parameters still require evaluation to assess the feasibility of blowfly-derived DNA as a standardised, practical monitoring tool.

One key factor affecting successful detection of ingested DNA is the digestion efficiency of the hematophagous, coprophagous, or saprophagous feeder (Calvignac-Spencer et al., 2013). Knowledge of the persistence period of mammal DNA in invertebrate guts is essential for designing standardised trapping methods for large-scale mammal monitoring via this approach. Typically, for ecological studies, blowflies are captured at baited traps, and may remain alive in the trap for several hours (Amat, 2010) or days (Akbarzadeh et al., 2012) before collection by the researcher.

Mammal mtDNA can be detected for as long as several months in leeches and ticks (Kent, 2009; Schnell et al., 2012) while it is usually completely digested after one to a few days in dipterans (Boakye et al., 1999; Mukabana et al., 2002; Kent & Norris, 2005; Kent, 2009). Boakye and colleagues (1999) detected human mtDNA in the bloodmeals of blackflies up to 72 h post-feeding. Human and other mammal (e.g., cow, dog, pig, and goat) mtDNA has been detected from bloodmeals of mosquitoes between 24 h and 48 h post-feeding (Mukabana et al., 2002; Kent & Norris, 2005; Kent, 2009).
Human and black rat mtDNA has been detected from bloodmeals of sandflies between 1 day up to 5 days post-feeding (Abbasi et al., 2009; da Silva Sales et al., 2015). Forensic studies have determined the post-feeding detection period of human mtDNA in the guts of blowfly maggots (Calliphora vicina) to be between 24 h and 48 h (Campobasso et al., 2005), but the post-feeding persistence period of mammal DNA in the guts of adult blowflies still lacks any reliable data.

As a step towards development of blowfly-derived DNA as an effective method for mammal monitoring, the objective of this chapter was to determine the persistence period of PCR amplifiable mammal DNA in adult blowflies.

3.2 Materials and methods

To obtain a sample of blowflies of known age, physiological state, and feeding history a laboratory culture of Chrysomya megacephala was established. A rotting fish carcass was obtained from a local supermarket and placed outside the Museum of Zoology, University of Malaya, Kuala Lumpur, to encourage egg deposition by wild blowflies. The fish was then moved into a 39 x 25 x 33 cm sealed plastic container and the hatching larvae were provided with more fish until pupation. Once all pupae had emerged, C. megacephala adults were selected out (based on the morphological characters of the species), and sorted into three containers of approximately 100 blowflies each.

The adult blowflies were then starved for 24 h to allow digestion of any food taken and to adjust the flies to similar hunger levels. After the starvation period pieces of market-supplied beef liver (Bos taurus) were placed into the containers for 4 h (06:00-10:00). The blowflies were observed to feed almost immediately upon the provision of food. After the beef liver was removed, sugar water was provided as the only food source. At 8, 16, 24, 48, 72, 96, 120, 144 h following removal of the beef
liver, 9 adult blowflies were selected arbitrarily (3 from each container) and frozen at -20°C. The blowfly’s guts were then dissected out with sterile implements for DNA extraction using a NucleoSpin Tissue kit (Macherey-Nagel, Germany), following the manufacturer’s instructions. To provide positive and negative controls respectively, DNA was also extracted directly from the market-supplied beef liver and from wild-caught blowfly legs (*C. megacephala*).

PCR was performed using beef-specific primers targeting a 75 bp region of *Bos taurus*, *cytochrome b* mtDNA (BSP F: 5’-CCCGATTCTTCGCTTTCCAT-3’ and BSP R: 5’-CTACGTCTGAGGAAATTCTGGTTG-3’) (Tanabe *et al.*, 2007). FastMix Frenche Hot Start PCR pre-mix (Intron Biotechnology, Korea) was used for all PCR reactions, adding 0.5 µL of each primer and 1 µL of DNA extract. The thermal cycling conditions were 94°C for 3 min, followed by 30 cycles of 98°C for 10 s, 55°C for 30 s, 72°C for 30s and a final extension at 72°C for 5 min. PCR products were visualised on a 1.5% agarose gel stained with 1 x GelRed (Biotium, USA). Beef/blowfly DNA mixtures were also created from the control DNA extracts, and PCR was performed to determine relative concentrations of beef DNA based on visual comparison of the electrophoresis images.

### 3.3 Results

All blowfly (*Chrysomya megacephala*) guts (10:05) sampled at 24 h post-feeding, or earlier, contained amplifiable beef mtDNA (Figure 3.1). At 48 h post-feeding 78% of blowfly guts had amplifiable beef DNA, but this dropped sharply to 44% at 72 h and 22% at 96 h. At these later time intervals, when amplification was successful, the bands were fainter indicating lower concentrations of DNA. There was no successful amplification from guts sampled at 120 h post-feeding and later.
Figure 3.1: PCR amplification of beef mtDNA from blowfly guts post feeding (8 h - 120 h). The top row 1-9 shows amplification of beef DNA from mixed blowfly/beef DNA extracts at different concentrations. The amount of beef DNA template in the PCR was: 1.25 ng, 0.83 ng, 0.50 ng, 0.28 ng, 0.15 ng, 0.08 ng, 0.04 ng, 0.02 ng and 0.01 ng for bands 1-9 respectively. The -ve control is DNA extracted from a wild-caught blowfly leg and the +ve control is DNA extracted from beef liver directly.
3.4 Discussion

Calvignac-Spencer and colleagues (2012) demonstrated the potential of sequencing iDNA for mammal monitoring and suggested the advantage of blowflies over other invertebrates (Calvignac-Spencer et al., 2013). However, certain questions remained regarding the field-application in large-scale mammal monitoring. Blowflies have been trapped in a variety of ways depending on the purpose of sampling (e.g., for veterinary purposes; Scholtz et al., 2000; Akbarzadeh et al., 2012). Calvignac-Spencer and colleagues (2012) sampled flies opportunistically immediately upon arrival at bait. A probable field-application scenario for large-scale mammal monitoring would see a large number of baited traps, where the trapped blowfly is unable to contact with the bait and is kept alive until collection (e.g., Whitworth, 2010). The traps would be set for a number of hours before the researcher is able to return to empty them.

In previous field experiments, *C. megacephala* usually arrived to fresh carrion within 24 h of exposure (Azwandi et al., 2013) meaning blowflies could potentially be in a trap for a number of hours digesting any mammal DNA present in their guts. My results indicate amplifiable mammal mtDNA persists in the guts of adult *Chrysomya megacephala* for 24-96 h post feeding (89% at 48 h) which is consistent with that determined for blowflies maggots (Campobasso et al., 2005) and other dipterans (mosquitoes, tsetse flies) (Kent, 2009). In contrast, mammal DNA can persist in ticks and leeches for several months (Kent, 2009; Schnell et al., 2012). This may seem like an advantage to using these ticks and leeches for mammal iDNA but means it is more difficult to determine when the detected mammal species was present at the sampling site (Schnell et al., 2015). Based on my findings, I suggest that blowflies will need to be retrieved from traps and processed at least every 24 h to maximise the chance of amplifying mammal DNA from their previous meal. The interval between a blowfly
feeding and looking for its next meal is unknown, but would also be a factor affecting the successful detection of ingested DNA.

This was the first experimental indication that successful detection of mammal DNA from blowflies is due to mammal DNA in their guts as opposed to mammal DNA being carried on their exoskeleton as a result of landing on mammal tissues or faeces. This was not addressed directly in previous studies (Calvignac-Spencer et al., 2012). The gradual decline in amplifiable mammal DNA could indicate a lack of severe enzymatic breakdown of ingested DNA in the blowfly fore-gut. For blowflies, primary digestion is achieved by secretion of salivary enzymes onto food before it is ingested orally (Hobson, 1932; Campobasso et al., 2005). The persistence period may be different depending on the length of the target fragment chosen for amplification. It is worth to note that my experiment to assess mammal DNA persistence is likely to provide an upper estimate as the substrate the blowflies were fed was of high quality (beef liver) and energy expenditure was limited (flies were kept in boxes).
CHAPTER 4: DESIGN AND TESTING OF PRIMERS FOR AMPLIFICATION OF MAMMAL DNA MINI-BARCODES FROM BLOWFLY GUTS

4.1 Introduction

iDNA is emerging as a powerful tool for large-scale mammal monitoring (Calvignac-Spencer et al., 2012; Schnell et al., 2012; Calvignac-Spencer et al., 2013; Schnell et al., 2015). However, certain key parameters still require evaluation to assess the feasibility of blowfly-derived DNA as a standardised, practical monitoring tool. One key factor for the success of iDNA approaches is selection of the target DNA region. The region must be easy to PCR amplify from taxonomically unknown samples, must be variable among taxa to permit species identification, and reference sequences of known species origin must exist in order to match the amplified fragment. Previous studies have used 16S rRNA and 12S rRNA for detection of mammal mtDNA from blowflies (Calvignac-Spencer et al., 2012), 16S rRNA from leeches (Schnell et al., 2012), cytochrome b from mosquitoes and sandflies (Abbasi et al., 2009; Bataille et al., 2012), and cytochrome c oxidase I (COI) from ticks and tsetse flies (Muturi et al., 2011; Gariepy et al., 2012). A fragment of COI is a preferred target for a number of reasons. Variation in COI has been used successfully to discriminate and identify mammals in Southeast Asia (Francis et al., 2010). There are more COI sequences than 16S rRNA sequences on GenBank (Benson et al., 2015) for mammal species (after excluding Homo sapiens sequences) (Figure 4.1) and this includes BARCODE standard records (Hanner, 2009). Therefore the chance of accurately assigning an unknown mammal to a species is suggested to be higher for a fragment of COI than for other gene regions.

Broad “universal” primers have been designed for amplification of mammal COI barcodes (Ivanova et al., 2012) but a short target is required for iDNA, which is likely to be partially degraded and difficult to amplify (Calvignac-Spencer et al., 2013).
Figure 4.1: Number of 16S rRNA and COI sequences publicly available on GenBank for the 246 mammal species (excluding Homo sapiens) found in Peninsular Malaysia.
Considering that field-application will likely involve pooling a large number of blowflies for cost-effective high-throughput sequencing (Stein et al., 2014), a 100 bp - 300 bp fragment is preferred (Shokralla et al., 2014). A “universal DNA mini-barcode for biodiversity analysis” has been published previously (Meusnier et al., 2008) but the primers have variable success amplifying mammal species (77% as reported by Meusnier et al., 2008; 0% as determined in silico by Ficetola et al., 2010; 80% as reported by Arif et al., 2011) and also will amplify blowfly DNA, which is likely to be present at higher concentration.

As a step towards development of blowfly-derived DNA as an effective method for mammal monitoring in biodiversity hotspots of Peninsular Malaysia, the objective of this chapter was to design and test primers that can selectively amplify mammal DNA mini-barcodes in the presence of high concentrations of blowfly DNA.

4.2 Materials and methods

A test dataset of 41 DNA extracts from 41 mammal species (16% of the mammal species found in Peninsular Malaysia) was assembled from a collection obtained during previous field sampling in Peninsular Malaysia (Sing et al., 2013; Syaripuddin et al., 2014; Wilson et al., 2014) and from collection at Ulu Gombak Forest Reserve Selangor, Gerik Perak, and nearby the Museum of Zoology, University of Malaya, Kuala Lumpur (Table 4.1). Based on the exploration of PCR amplification of COI with this dataset and other mammal samples, Uni-Mini-bar F (Meusnier et al., 2008) and RonM (Ivanova et al., 2012) have good success as a forward primers, but LepF1 (Hebert et al., 2004), HCO2198 (Folmer et al., 1994), and VF1d (Ivanova et al., 2012) have lower success. Uni-Mini-bar R (Meusnier et al., 2008) has low success as a reverse primer, but VR1d (Ivanova et al., 2012) has good success. Given the high
Table 4.1: List of 41 mammal species from Malaysia for comparison of amplification success between primer pairs, Uni-Mini-bar F/ RonPing and Uni-Mini-bar F/ Uni-Mini-bar R and BOLD Process ID numbers for those sequenced after troubleshooting.

Troubleshooting was performed using only the primer pair, Uni-Mini-bar F/ RonPing.

<table>
<thead>
<tr>
<th>Specimen Number</th>
<th>Species</th>
<th>Uni-Mini-bar F/ Uni-Mini-bar R 1st Pass</th>
<th>Uni-Mini-bar F/ RonPing 1st Pass</th>
<th>Troubleshooting</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT31B</td>
<td>Balionycteris maculata</td>
<td>/</td>
<td>/</td>
<td>RONP010-14</td>
</tr>
<tr>
<td>P16B</td>
<td>Bandicota indica</td>
<td>/</td>
<td>/</td>
<td>RONP029-14</td>
</tr>
<tr>
<td>BF5A</td>
<td>Bos taurus</td>
<td>×</td>
<td>×</td>
<td>RONP001-14</td>
</tr>
<tr>
<td>DOGB</td>
<td>Canis lupus familiaris</td>
<td>×</td>
<td>×</td>
<td>RONP031-14</td>
</tr>
<tr>
<td>GOATB</td>
<td>Capra hircus</td>
<td>/</td>
<td>/</td>
<td>RONP050-14</td>
</tr>
<tr>
<td>BGM19A</td>
<td>Chironax melanocephalus</td>
<td>×</td>
<td>/</td>
<td>RONP002-14</td>
</tr>
<tr>
<td>BK28B</td>
<td>Cynopterus brachyotis</td>
<td>/</td>
<td>/</td>
<td>RONP049-14</td>
</tr>
<tr>
<td>BT4B</td>
<td>Cynopterus horsfieldii</td>
<td>×</td>
<td>/</td>
<td>RONP040-14</td>
</tr>
<tr>
<td>BK4B</td>
<td>Cynopterus JLE sp. A</td>
<td>×</td>
<td>/</td>
<td>RONP012-14</td>
</tr>
<tr>
<td>BT73B</td>
<td>Eonycteris spelaea</td>
<td>/</td>
<td>/</td>
<td>RONP025-14</td>
</tr>
<tr>
<td>CATB</td>
<td>Felis catus</td>
<td>/</td>
<td>/</td>
<td>RONP027-14</td>
</tr>
<tr>
<td>BT64A</td>
<td>Glischropus tylopus</td>
<td>/</td>
<td>/</td>
<td>RONP009-14</td>
</tr>
<tr>
<td>TI13B</td>
<td>Hipposideros bicolor131</td>
<td>/</td>
<td>×</td>
<td>RONP033-14</td>
</tr>
<tr>
<td>T25B</td>
<td>Hipposideros cervinus</td>
<td>×</td>
<td>/</td>
<td>RONP026-14</td>
</tr>
<tr>
<td>T27B</td>
<td>Hipposideros cf. larvatus</td>
<td>/</td>
<td>/</td>
<td>RONP015-14</td>
</tr>
<tr>
<td>BK1B</td>
<td>Hipposideros diadema</td>
<td>×</td>
<td>×</td>
<td>RONP034-14</td>
</tr>
<tr>
<td>BT30B</td>
<td>Hipposideros doriae</td>
<td>/</td>
<td>/</td>
<td>RONP042-14</td>
</tr>
<tr>
<td>BK13B</td>
<td>Kerivoula cf. hardwickii</td>
<td>/</td>
<td>/</td>
<td>RONP019-14</td>
</tr>
<tr>
<td>BT25B</td>
<td>Kerivoula minuta</td>
<td>/</td>
<td>/</td>
<td>RONP039-14</td>
</tr>
<tr>
<td>BT44B</td>
<td>Kerivoula papillosa</td>
<td>/</td>
<td>/</td>
<td>RONP014-14</td>
</tr>
<tr>
<td>BT17A</td>
<td>Kerivoula pellucida</td>
<td>/</td>
<td>/</td>
<td>RONP004-14</td>
</tr>
<tr>
<td>PING14B</td>
<td>Leopoldamys sabanus</td>
<td>/</td>
<td>/</td>
<td>RONP035-14</td>
</tr>
<tr>
<td>BK12B</td>
<td>Macroglossus sobrinus</td>
<td>/</td>
<td>/</td>
<td>RONP018-14</td>
</tr>
<tr>
<td>PING15B</td>
<td>Maxomys surifer</td>
<td>/</td>
<td>/</td>
<td>RONP036-14</td>
</tr>
<tr>
<td>BK22A</td>
<td>Megaderma lyra</td>
<td>/</td>
<td>/</td>
<td>RONP005-14</td>
</tr>
<tr>
<td>BT59B</td>
<td>Megaderma spasma</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>BT50B</td>
<td>Megaerops ecaudatus</td>
<td>/</td>
<td>/</td>
<td>RONP041-14</td>
</tr>
<tr>
<td>T74B</td>
<td>Murina aenea</td>
<td>/</td>
<td>/</td>
<td>RONP016-14</td>
</tr>
<tr>
<td>BT10B</td>
<td>Murina cyclotis</td>
<td>×</td>
<td>/</td>
<td>RONP013-14</td>
</tr>
<tr>
<td>BK32B</td>
<td>Murina suilla</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>AF1B</td>
<td>Mus musculus</td>
<td>/</td>
<td>/</td>
<td>RONP051-14</td>
</tr>
<tr>
<td>BT2B</td>
<td>Myotis muricola</td>
<td>/</td>
<td>/</td>
<td>RONP037-14</td>
</tr>
<tr>
<td>PING18A</td>
<td>Rattus andamanensis</td>
<td>/</td>
<td>/</td>
<td>RONP006-14</td>
</tr>
<tr>
<td>PING3B</td>
<td>Rattus tiananicus</td>
<td>/</td>
<td>/</td>
<td>RONP038-14</td>
</tr>
<tr>
<td>BT1B</td>
<td>Rhinolophus acuminatus</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>BK49B</td>
<td>Rhinolophus affinis</td>
<td>/</td>
<td>/</td>
<td>RONP021-14</td>
</tr>
<tr>
<td>BT34B</td>
<td>Rhinolophus tricolorius</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>BK67B</td>
<td>Rhinolophus yunanensis</td>
<td>/</td>
<td>/</td>
<td>RONP011-14</td>
</tr>
<tr>
<td>Z2B</td>
<td>Suncus murinus</td>
<td>/</td>
<td>/</td>
<td>RONP032-14</td>
</tr>
<tr>
<td>WB8B</td>
<td>Sus scrofa</td>
<td>/</td>
<td>/</td>
<td>RONP008-14</td>
</tr>
<tr>
<td>PING11B</td>
<td>Tupaia glis</td>
<td>/</td>
<td>×</td>
<td></td>
</tr>
</tbody>
</table>

Total: 29/41 (71%) 32/41 (78%) 36/41 (89%)
success of Uni-Mini-bar F and RonM, and the 210 bp distance between these primers, a reverse primer was designed by targeting the RonM binding site.

Mammal COI sequences were retrieved from GenBank (Benson et al., 2015) and the program Primer3 (Rozen & Skaletsky, 2000) was used to select a 19 bp region slightly upstream of RonM, which enabled design of a reverse primer with appropriate physical and structural properties. When used in combination with Uni-Mini-bar F, the new primer RonPing (5’-TATCAGGGGCTCCGATTAT-3’) should amplify a 205 bp fragment at the 5’ end of the COI barcode region (Figure 4.2).

The success of Uni-Mini-bar F/ RonPing and Uni-Mini-bar F/ Uni-Mini-bar R across mammal species was then systematically compared using the 41 species dataset. PCR was performed using FastMix Frenche Hot Start PCR pre-mix (Intron Biotechnology, Korea) and COI Fast thermocycling program (Wilson, 2012) for all reactions, with slight modification to the DNA volume (0.5-2 µL) depending on DNA extraction method.

Next the ability of the Uni-Mini-bar F/ RonPing combination to amplify low concentrations of mammal DNA in the presence of high concentrations of blowfly DNA was tested. The DNA extracts from ten mammal species were mixed with DNA extracts from blowfly legs in a ratio of 1 part mammal (~0.43 ng) to 16 parts blowfly (~7.0 ng). PCR was performed for the mixed DNA samples as described above. Additionally, DNA extracts from the 48 h post-feeding samples from the feeding experiment above were used as templates for PCR with Uni-Mini-bar F/ RonPing.

A further test involved the detection of mammal DNA from wild-caught blowflies. Four baited traps (modified from Calvignac-Spencer et al., 2013) were set at Rimba Ilmu, University of Malaya (Figure 4.3). Rimba Ilmu is an 80 ha botanical garden and a habitat for small mammals such as bats, squirrels, tree shrews and rats. The traps were emptied every 24 h and blowflies were frozen at -20°C. The guts of 30
Figure 4.2: A) Relative positions of mammal primers on the COI barcode region. B)

The binding sites of the primers Uni-Mini-bar F and RonPing (reverse complement).

The binding sites of the primers were in relation to a *Chrysomya megacephala* (blowfly) and *Thomomys atrovarius* (Smooth-toothed pocket gopher) sequence.
Figure 4.3: Baited blowfly trap for mammal monitoring.
collected blowflies were then dissected out with sterile implements for DNA extraction and PCR with Uni-Mini-F/ RonPing as above. PCR products were sequenced in both directions by a local company (MYTACG-Kuala Lumpur, Malaysia). DNA sequences are available on BOLD (Ratnasingham & Hebert, 2007; Ratnasingham & Hebert, 2013) in the dataset DS-RONPING (http://www.boldsystems.org/index.php/Public_SearchTerms?query=DS-RONPING).

To evaluate the success of the 205 bp DNA mini-barcode amplified by Uni-Mini-bar F/ RonPing for species assignment, COI sequences from mammal species found in Malaysia (based on Davison & Zubaid, 2007) were mined from GenBank (Benson et al., 2015). I retrieved COI sequences >600 bp and the sequences were trimmed to the 205 bp DNA mini-barcode target amplified by Uni-Mini-bar F/ RonPing. A neighbor-joining tree based on number of differences was produced from the aligned sequences using MEGA 6.0 (Tamura et al., 2013).

4.3 Results

The newly designed RonPing primer had a low number of mismatches (less than six) with most mammal COI sequences retrieved from GenBank (Benson et al., 2015) (Figure 4.4). RonPing had five mismatches with C. megacephala (Figure 4.2) and did not amplify this species. Canis lupus (Order: Carnivora) also had five mismatches located in approximately the same positions, with one near to 3’end of primer, but Canis lupus COI was amplified with RonPing. Other mammal species showing 3-4 mismatches were also amplified and sequenced successfully.

The success of primer pairs, Uni-Mini-bar F/ Uni-Mini-bar R (Meusnier et al., 2008) and Uni-Mini-bar F/ RonPing in amplifying COI from 41 mammal species extracts were 71% and 78% respectively (Figure 4.5). A further round of PCR using Uni-Mini-bar F/ RonPing yielded a higher proportion of species amplified (89%).
Figure 4.4: Number of mismatches between the RonPing primer sequence (19 bp) and 1472 mammal COI sequences from GenBank.
Figure 4.5: Comparison of amplification success using primer pairs: Uni-Mini-bar F/ RonPing and Uni-Mini-bar F/ Uni-Mini-bar R for Artiodactyla (n=3), Carnivora (n=2), Chiroptera (n=28), Soricomorpha (n=1), Rodentia (n=5), and Scandentia (n=1).
Mammal sequences amplified from ten mammal/ *C. megacephala* DNA mixtures using Uni-Mini-bar F/ RonPing had high quality peaks and were clear of contamination (Figure 4.6). The Uni-Mini-bar F/ RonPing combination showed successful amplification from 27% of wild-caught *C. megacephala* with sequenced amplicons showing close matches (<95%) to *Rhinolophus* sp., *Bos taurus* and *Gallus gallus*.

Examining the target 205 bp region among mammals from Peninsular Malaysia mined from GenBank (113 species; Benson *et al*., 2015), all species, except for seven, possessed a unique COI sequence or unique sets of COI sequences for the 205 bp DNA mini-barcode (See the neighbor-joining tree here: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0123871). The DNA mini-barcode region was also able to separate 26 “dark” bat taxa recognised in previous DNA barcoding studies (e.g., *Balionycteris maculata*, *Hipposideros armiger*, and *Myotis muricola*) (Francis *et al*., 2010).

### 4.4 Discussion

Calvignac-Spencer and colleagues (2012) used *16S rRNA* in detecting iDNA for mammal monitoring but it was problematic. While the *16S rRNA* fragment is easy to amplify using primers with broad taxonomic coverage across Mammalia, the short “barcode” produced has low species resolution i.e., multiple closely related species can share the same haplotype (Ficetola *et al*., 2010). My results suggest that targeting COI, the animal DNA barcode (Hebert *et al*., 2003), is a very practical option, providing the chance to exploit the identification capacity of BOLD (Ratnasingham & Hebert, 2007; Ratnasingham & Hebert, 2013) and the well-characterised patterns of species level divergence at this region such as from Francis *et al*., (2010). The new primer, RonPing, when used in combination with Uni-Mini-bar F amplifies a 205 bp fragment of COI. This appears to be an optimal length for a DNA mini-barcode allowing amplification.
Figure 4.6: Examples of fragments of mammal sequences amplified using Uni-Mini-bar F/ RonPing primers from mammal/ C. megacephala DNA mixtures.
from degraded samples, such as iDNA, while not suffering a reduction in the ability to distinguish species (Meusnier et al., 2008). This target also falls within the maximum read length for high-throughput sequencing (e.g., 300 bp for the Illumina MiSeq (Shokralla et al., 2014)), including spare length for multiplex identifier (MID) tags to separate multiple samples (Carew et al., 2013).

The new primer combination was able to amplify a higher proportion of mammal species than previously proposed combinations (Meusnier et al., 2008). Another significant advantage of the Uni-Minibar F/ RonPing combination is the fact that it did not appear to amplify C. megacephala COI, even when the ratio of blowfly DNA to mammal DNA is high. This is likely due to RonPing having a double mismatch with C. megacephala within 5 bp of the 3’ end including a purine-purine mismatch (A-G) (Stadhouders et al., 2012). Sixteen mammal species also had an A-G mismatch in the same position, but they only had one mismatch within 5 bp from 3’end of primer not the C. megacephala double mismatch, which may be the main reason for successful amplification from these DNA templates. Lack of C. megacephala amplification negates the need for blocking primers (Vestheim & Jarman, 2008) or, if no blocking probes are used in high-throughput sequencing, prevents significant wastage of sequencing effort due to amplification and sequencing of blowfly DNA (Pinol et al., 2014). The low number and even distribution of mainly pyrimidine-pyrimidine and purine-pyrimidine mismatches between the RonPing primer and GenBank mammal sequences suggest primer bias might not be too severe and consequently the primer may produce a less biased estimate of DNA templates present (Stadhouders et al., 2012). Although the ecoPCR program (Ficetola et al., 2010) predicted that the Uni-Mini-bar F/ RonPing combination could only amplify 19% of mammal species by allowing three mismatches to the whole mitochondrial genomes on GenBank (Benson et al., 2015), results coherent with in vitro PCR can be obtained by allowing a higher number of
mismatches. The success in detecting mammal DNA from wild-caught blowflies (27% of blowflies contained detectable mammal DNA) trapped over a period of 24 hours indicates the potential of a standardised trapping protocol with retrieval of blowflies every 24 h for effective mammal monitoring in the field.

The 205 bp COI fragment was successful in distinguishing nearly all examined mammal species from Peninsular Malaysia and separating “dark” bat taxa, previously recognised species that lack formal taxonomic status (Wilson et al., 2014). This suggested the potential of detecting cryptic taxa overlooked by traditional methods (Francis et al., 2010). The species which shared haplotypes at the 205 bp region such as Rattus tiomanicus and R. rattus included sequences which have previously been identified as problematic and most likely represent misidentifications or contamination rather than shared haplotypes (Shen et al., 2013).
CHAPTER 5: BLOWFLY DISPERSAL RANGE AND IMPLICATIONS FOR BLOWFLY-DERIVED DNA IN TROPICAL MAMMAL MONITORING

5.1 Introduction

Blowflies in the family Calliphoridae, include over 1,100 species worldwide (Mackerras, 1933; Byrd & Castner, 2001) and are among the first insects to colonise animal carcasses (Amendt et al., 2004). Blowflies play a crucial role in the breakdown of vertebrate carcasses and recycling nutrients (Shah et al., 2015). In the tropical forests of Peninsular Malaysia, Chrysomya megacephala is among the first and most abundant species visiting mammal carrion (Azwandi et al., 2013). The ability to arrive in large numbers at animal carcasses and faeces as well as their distribution over all habitats (Norris, 1965) has been recognised as an advantage of the use of blowflies over other invertebrates in iDNA approaches for mammal monitoring (Calvignac-Spencer et al., 2013).

Knowledge of the dispersal range is essential for addressing uncertainties in the use of an iDNA approach i.e., determining the actual or potential location of detected species relative to the location where the invertebrates were collected (Schnell et al., 2015). The range of blowflies, in particular, can result in great uncertainties regarding the precise location of mammal species detected from blowfly-derived DNA (Calvignac-Spencer et al., 2013; Schnell et al., 2015). Blowflies are thought to disperse great distances, relative to other invertebrates such as leeches that exhibit little movements (Schnell et al., 2015). However, data on adult blowfly dispersal capabilities is surprisingly scarce (Braack & Retief, 1986; Amat et al., 2016). A few studies from temperate and subtropical regions suggested daily dispersal capabilities of 100-2,400 m per day (Braack & Retief, 1986; Smith & Wall, 1998; Tsuda et al., 2009; Table 5.1),
Table 5.1: Summary of MRR studies on adult fly dispersal including the marking techniques, study locations, and dispersal ranges.

<table>
<thead>
<tr>
<th>Targeted family</th>
<th>Targeted species</th>
<th>Marking techniques and study locations</th>
<th>Dispersal ranges</th>
<th>Recapture rate</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calliphoridae</td>
<td>Calliphora nigribarbis</td>
<td>Ink marking; Ikumo-Makka, Japan</td>
<td>1.250 – 1.789 km/day</td>
<td>Low; 0.014% - 0.029%</td>
<td>Tsuda et al., 2009</td>
</tr>
<tr>
<td></td>
<td>Chrysomya albiceps; Chrysomya marginalis</td>
<td>Radioactive labeling using $^{32}$P-orthophosphate; Kruger National Park, South Africa</td>
<td>2.20 km/day (for C. albiceps); 2.35 km/day (for C. marginalis)</td>
<td>Low; C. albiceps: 0.1 - 0.45%; C. marginalis: 0.13 - 0.93%</td>
<td>Braack &amp; Retief, 1986</td>
</tr>
<tr>
<td></td>
<td>Lucilia sericata</td>
<td>Fluorescent dust marking; South West England</td>
<td>0.11 - 0.15 km/day</td>
<td>Low; 4-14%</td>
<td>Smith &amp; Wall, 1998</td>
</tr>
<tr>
<td></td>
<td>Phormia regina</td>
<td>Radioactive labeling using $^{32}$P-orthophosphate; West Virginia</td>
<td>Daily dispersal not estimated; Dispersal range: 9-16 km</td>
<td>Low; &lt;1%</td>
<td>Schoof &amp; Mail, 1953</td>
</tr>
<tr>
<td>Muscidae</td>
<td>Musca domestica</td>
<td>Fluorescent dust marking; Selangor, Malaysia</td>
<td>Daily dispersal not estimated; Mean distance of dispersal: 2.05 km</td>
<td>Low; 0.016-0.023%</td>
<td>Nazni et al., 2005</td>
</tr>
<tr>
<td></td>
<td>Musca autumnalis</td>
<td>Immunomarking with egg white; Prosser, Western Australia</td>
<td>Daily dispersal not estimated; Dispersal distance: ≤ 100 m - ≥ 450 m</td>
<td>Low; 16.3%</td>
<td>Peck et al., 2014</td>
</tr>
<tr>
<td>Drosophilidae</td>
<td>Drosophila spp.</td>
<td>Fluorescent dust marking; New Jersey, USA.</td>
<td>Daily dispersal not estimated; 0 - &gt; 60 m range</td>
<td>Low; 10%</td>
<td>Worthen, 1989</td>
</tr>
<tr>
<td>Tephritidae</td>
<td>Anastrepha ludens</td>
<td>Fluorescent dye marking; Nuevo Leon, Mexico</td>
<td>Daily dispersal not estimated; Dispersal range: 0.1-7 km</td>
<td>Low; 0.7-1%</td>
<td>Thomas &amp; Loera-Gallardo, 1998</td>
</tr>
<tr>
<td></td>
<td>Dacus cucurbitae</td>
<td>Enamel painting; Ishigaki Island, Japan</td>
<td>Daily dispersal not estimated; Mean dispersal: ≤ 100 m</td>
<td>Low; 0.26-8.99%</td>
<td>Hamada, 1980</td>
</tr>
</tbody>
</table>
and no data exists on blowfly dispersal ranges in the tropics. The dispersal ranges of blowflies differ between species, with environmental conditions acting as barriers to some species (Macleod & Donnelly, 1960; Tsuda et al., 2009). Previous studies of the dispersal of adult dipterans (including Calliphoridae, Muscidae, Drosophilidae, Tephritidae) including marking techniques, study locations, and dispersal ranges are summarised in Table 5.1.

Considering the implications of the dispersal capabilities of blowflies (Family: Calliphoridae) for the development of blowfly-derived DNA as an effective, standardised mammal monitoring tool, the objective of this chapter was to determine blowfly dispersal ranges in tropical forests.

5.2 Materials and Methods

Adult Chrysomya spp. blowflies were collected using traps baited with rotten fish around the University of Malaya, Kuala Lumpur and Kampong Ulu Dong, Pahang between 17 December 2015 and 26 December 2015 (10 days in total). The wild caught blowflies were transferred into ten cages (39 x 25 x 33 cm) of approximate 300 blowflies each (totalling approximately 3,000 adult blowflies). The blowflies were maintained at room temperature (27°C-33°C) and provided with cotton pads soaked in sugar solution. The blowflies were marked by orange-colored fluorescent dust (Transcend Solutions-Selangor, Malaysia) one day prior to release (following Howard et al., 1989; Nazni et al., 2005). Fluorescent dust remains detectable for the duration of the life of flies or at least 28 days under natural conditions (Pickens et al., 1967; Moth & Barker, 1975; Lillie et al., 1981) and does not affect the survival of the flies (Pickens et al., 1967; Moth & Barker, 1975).

The marked blowflies were released at 10:00 corresponding with peak activity of blowflies on 27 December 2015 in a rubber plantation located at Kalumpang,
Selangor (3°36'N 101°33'E) (Figure 5.1). The plantation is suitable for fly dispersal studies that require sufficient spatial scale of study area (Hassall & Thompson, 2012) as it provides large surface area and easy access for sampling. It has an equatorial climate classified as rainforest climate (Kottek et al., 2006) and is protected from strong winds and direct sunlight by shade trees (Alexander et al., 2002). The annual air temperature in Kalumpang is >24°C with high humidity (80%-90%) and annual rainfall of 2,850 mm with two distinct wet seasons; April-May and September-November (Nieuwolt, 1982).

Recapture of blowflies commenced 24 h after release and continued for 9 consecutive days (following Howard et al., 1989; Chiang et al., 1991; Smith & Wall, 1998). The weather conditions throughout the 9 sampling days were mostly cloudy with slight or no rain, daily temperature ranges from 24°C-34°C, and dominant wind directions of North and Northeast (AccuWeather, 2016). Blowfly traps baited with rotten fish were set 2 m above the ground in five concentric circles around the blowfly release point at an increasing distance from the release point (1 km, 2 km, 3 km, 4 km, and 5 km). 57 traps were set in total with the number of traps per circle increasing with distance of the concentric circle from the release point (Figure 5.1). Blowflies were collected from the traps daily and frozen at 0°C before being examined for fluorescent powder under ultraviolet (UV) light in a darkened room.

Recaptured blowflies were identified to genus using morphological characters following Kurahashi and colleagues (1997). The number of *Chrysomya* spp. blowflies recaptured at different distances from the release point for each sampling day was shown in Figure 5.2.
**Figure 5.1:** The location of Kalumpang, Selangor showing the experimental design for blowfly recapture. X is the release point of blowflies.
Figure 5.2: The number of blowflies recaptured based on number of days since release and distance from the release point (1-5 km). The size of the circle indicates the number of blowflies recaptured.
5.3 Results

Fourty-three *Chrysomya* spp. blowflies were recaptured at a distance of 1-3 km away from the release point during the recapture period of 9 days (Figure 5.2). Of the 43 blowflies recaptured (1.5% of 3,000 released), 34 (79%) were recaptured at a distance of 1 km from release point, eight (18.6%) recaptured at 2 km, and one (2.3%) recaptured at 3 km (Figure 5.2). No marked blowflies were recaptured at 4 km and 5 km away from the release point (Figure 5.2). The densities of recaptured blowflies were 10.82 per km$^2$ (34 out of 3.142 km$^2$) at 1 km from release point, 3.34 per km$^2$ (42 out of 12.566 km$^2$) at 2 km, and 1.52 per km$^2$ (43 out of 28.274 km$^2$) at 3 km.

Of 43 recaptured blowflies, fourteen blowflies were recaptured on the first day after release, 11 on the second day after release, 6 on the third day, 5 on the fourth day, 5 on the fifth day, and 3 on the sixth day. No blowflies were recaptured after six days although trapping of blowflies continued for 9 days following the release (Figure 5.2).

The marked blowflies were recaptured randomly in all directions (North, East, West, and South) at 1 km away from the release point, with majority recaptured in the direction of East (12 out of 34) and others recaptured in the direction of South (9 out of 34), West (8 out of 34), and North (5 out of 34) from the release point. At distance of 2 km away from the release point, the majority of blowflies were recaptured at South (4 out of 8) and Southwest (3 out of 8). At distance of 3 km from the release point, a single blowfly was recaptured in the direction of Southwest from the release point.

5.4 Discussion

Mark-release-recapture (MRR) studies are commonly used to study insect dispersal in the field (Hagler & Jackson, 2001; Nazni *et al.*, 2005), however, this is the first experimental indication of blowfly dispersal in the tropics based on MRR. The
dispersal range of *Chrysomya* spp. blowflies was 1-3 km during the sampling period of 9 days. Most blowflies (79% out of 43) were recaptured at 1 km from the release point throughout the sampling period, whereas a few blowflies (20.9% out of 43) were recaptured 2-3 km away from the release point. No blowflies were recaptured at more than 3 km from the release point. This suggests that *Chrysomya* spp. blowflies did not disperse widely at least over time intervals in the range of a few days. The daily dispersal distance of <3 km recorded for *Chrysomya* spp. blowflies is similar to that reported in studies of *Chrysomya albiceps* and *Chrysomya marginalis* in South Africa with estimated daily dispersal of 2.20 km and 2.35 km respectively (Braack & Retief, 1986).

A 3.34 fold decrease was observed in the density of blowflies recaptured from 1 km to 2 km away from the release point. This was followed by a 2.20 fold decrease in the density of blowflies recaptured from 2 km to 3 km from the release point. The decrease in density of blowflies with increasing area of sampling could be due to insufficient number of blowfly traps employed. Smith *et al.*, 2007 suggested that lower density of animals present per sampling unit could result in lower detection probability of animals, possibly leading to underestimation of animal presence in the area. Consequently, more blowfly traps should be employed relative to the increasing sampling area for higher probabilities of recapturing blowflies.

The maximum blowfly dispersal distance recorded in the present study was 3 km, where a single *Chrysomya* sp. was recaptured 5 days post-release. Maximum flight distance estimates for blowflies varied depending on the species and regions. For example, *Chrysomya albiceps* and *Chrysomya marginalis* were found to disperse up to 37.5 km and 63.5 km respectively after a week of dispersal without disturbance in South Africa (Braack & Retief, 1986), whereas the maximum dispersal distance of *Chrysomya rufifacies* was 16 km over 12 days in Australia (Gurney & Woodhill, 1926). This could
be explained by different blowfly species having distinct dispersal rates and flight capabilities (Macleod & Donnelly, 1960; Tsuda et al., 2009). Another important factor affecting the dispersal of blowflies could be the nature of vegetation in the area (Braack & Retief, 1986).

The recapture rates of blowflies at different distances from the release point were low (0.02-1.1%) throughout the sampling period. This result is similar with the low recapture rates in other blowfly dispersal studies (0.1-0.93% from Braack & Retief, 1986; 0.014-0.029% from Tsuda et al., 2009; see Table 5.1). Fly dispersal studies using MRR are difficult to perform, requiring a relatively large amount of flies to be caught prior to release due to low recapture probabilities (Leak, 1999). The marked blowflies were not detected beyond six days post-release, suggesting the longevity of marked blowflies is around a week in the field. This, however, does not represent the actual longevity of blowflies due to the wild blowflies collected being of varying ages prior to release.

The directional movement by *Chrysomya* spp. blowflies subsequently after release appeared to be random at distance of 1 km away from the release point. However, at 2 km and 3 km away from the release point, the majority marked blowflies dispersed in the direction of South and Southwest away from the release point, where a river is located in the direction of West and Southwest. This could be due to preference of blowflies towards habitats with rivers to fulfill their need for hydration for survival as suggested by Braack & Retief (1986). Although the dominant wind directions during sampling period were North and Northeast, previous blowfly dispersal studies showed no significant effect of wind directions on blowfly directional movement (Braack & Retief, 1986; Smith & Wall, 1998).

My study suggests that within a short (known) period of time (six days which is the period limit of detectable mammal DNA in blowfly guts; see Chapter 3), blowflies
could sample DNA from the tissues and faeces of mammals and travel up to 3 km away from the mammals. This implies that a species detected in blowfly-derived DNA should be present within 3 km of the blowfly sampling location where blowflies were collected daily. One advantage of blowfly-derived DNA as compared to other invertebrates could be short persistence period of mammal DNA in blowfly guts (24-96 h) as this precludes species detected in blowfly-derived DNA from being far from the blowfly sampling location. However, there still remains some uncertainty as blowfly dispersal capability could be affected by environmental factors such as solar radiation, rainfall, temperature, and wind activity (Von Aesch et al., 2003; Tsuda et al., 2009).
CHAPTER 6: FIELD-CALIBRATION OF BLOWFLY-DERIVED DNA AGAINST TRADITIONAL FIELD METHODS

6.1 Introduction

The common field methods used to assess mammal diversity in tropical forests include: live trapping, such as cage traps (Hanif-Ridzuan et al., 2010; Madinah et al., 2011); mist nets and harp traps (Kingston et al., 2003; Sing et al., 2013); indirect signs such as tracks or scat (Daim, 2002); interviews with local communities (Sharma et al., 2005); direct observations by researchers (Syakirah et al., 2000; Jayaraj et al., 2013); and hair traps (Castro-Arellano et al., 2008; Hedges et al., 2015b). Recent additions to the toolbox are secondary sources of mammal DNA. For example, mammal DNA has been detected from owl-pellet bones (Rocha et al., 2015) and iDNA (Calvignac-Spencer et al., 2012; Schnell et al., 2012; Calvignac-Spencer et al., 2013; Schnell et al., 2015). iDNA has been suggested to show promising potential for use in mammal diversity assessments, but to date, there has been no systematic comparison between iDNA and traditional methods in the field (Schnell et al., 2015).

Recently a number of studies from the forests of Peninsular Malaysia, home to 222 mammal species (DWNP, 2010), have simultaneously tested the effectiveness of different field methods in detecting different groups of mammal species (Syakirah et al., 2000; Jayaraj et al., 2012; Tingga et al., 2012; Jayaraj et al., 2013). In Gunung Stong State Park, in the state of Kelantan, a combination of cage traps, mist nets and harp traps yielded low capture rates of non-volant small mammals (Jayaraj et al., 2012); Cage traps did not detect treeshrews and squirrels (probably due to bait incompatibility and poor positioning), while harp traps captured a low diversity of insectivorous bats. In Taman Negara, Pahang, a combination of cage traps, Sherman traps, mist nets, and harp traps demonstrated the efficiency of harp traps for mammal sampling (Tingga et al.,
2012). The harp traps accounted for 84% of the bat species reported and 65% of the total mammal species. In contrast, at Wang Kelian State Park, Perlis, after several years of surveys using a combination of harp traps, cage traps and direct observations, the inclusion of mist nets in the latest sampling period increased the number of species sampled by 33% (from an average of 24 per survey to 36; Jayaraj et al., 2013).

The objective of this chapter was to calibrate the performance of blowfly-derived mammal DNA at generating species inventories and richness estimates against traditional field methods in Peninsular Malaysia. I compared five methods - cage trapping, mist netting, hair trapping, scat collection, and blowfly-derived DNA.

6.2 Materials and methods

A field calibration was conducted in a forest reserve of Peninsular Malaysia - Ulu Gombak Forest Reserve (UGFR) in Selangor. UGFR comprises 17,000 ha of logged dipterocarp forest (3°20'N, 101°45'E; Figure 6.1). UGFR is considered one of the most species-rich localities for Old World bats (Sing et al., 2013) and has been intensively studied since the establishment of Ulu Gombak Field Studies Centre (University of Malaya) 50 years ago (Medway, 1966). I compiled an extensive mammal checklist for UGFR based on previous records (including University of Malaya student theses) of sampled or observed species (Appendix A).

The field survey was conducted between 3 November 2014 and 25 December 2014 for two nights each week (16 nights in total). Two transects of 100 m were established 1500 m apart (Figure 6.1). The survey incorporated (a) baited cage traps, (b) mist nets, (c) baited hair traps, (d) scat collection, and (e) baited blowfly traps.

a) Ten wire mesh cage traps (Figure 6.1) were set per transect at 10 m intervals, monitored and baited with fresh banana daily. Banana is reported among the effective baits for cage traps used for small mammal diversity assessments in tropical forests of
Figure 6.1: The location of Ulu Gombak Forest Reserve showing the location of two transects and the experimental design for comparing five different field methods.
Malaysia (Bernard, 2003; Payne & Francis, 2005; Madinah et al., 2011), and is resistant to removal by invertebrates (Bernard 2003). Cage traps were checked daily and hair samples from any trapped mammals were collected into a 1.5 mL microcentrifuge tube before release.

b) Ten mist nets (9 m x 4 m x 36 mm mesh size) (Figure 6.1) were positioned across potential flight paths of bats (trails or streams). Mist nets were set at a single transect from 1900 and were monitored continuously until 2300 or until it rained. A small wing punch was collected from each captured bat into a 1.5 mL ethanol-filled microcentrifuge tube following AMNH (2013). Scissors and forceps were cleaned with alcohol and sterile tissues between bats to avoid cross-contamination. Mist nets were only used for four weeks (two nights per week) in November.

c) Ten hair traps (PVC pipes covered with reversed duct tape on the inside) (Figure 6.1) were set per transect at 10 m intervals and monitored daily. Hair traps were baited with banana at the start of the 2-night cycle and collected at the end of the cycle. Any hair samples from traps were collected into a 1.5 mL microcentrifuge tube using sterile forceps.

d) Any scat (Figure 6.1) encountered while walking along transects daily was collected into a 1.5 mL microcentrifuge tube.

e) Ten blowfly traps baited with rotten fish (Figure 6.1) were set per transect at 10 m intervals and 2 m above the ground. The traps were emptied every 24 h and collected blowflies (Family: Calliphoridae) were frozen at -20°C within 5 h of trap emptying. The guts of all collected blowflies (=315) were then dissected with sterile implements and pooled (five individual blowfly guts per tube) for DNA extraction (63 extractions in total). The DNA extracts were further pooled by week and transect of sampling, resulting in (8 x 2) 16 pooled DNA extracts for bulk PCR.
The protocol for minimally-invasive collection of mammal DNA samples (hair, wing punches) from mammals trapped in cages or nets has been approved by the University of Malaya Institutional Animal Care and Use Committee (UMIACUC) (Ref. ISB/02/1212013/JJW (R)) and the Department of Wildlife and National Parks, Peninsular Malaysia (Ref. JPHL&TN(IP): 80-4/2 Jld16(24)). DNA was extracted from the collected hair and bat tissue samples using an alkaline lysis extraction method (modified from Ivanova et al., 2012). The samples collected into 1.5 mL microcentrifuge tubes were first cut into smaller fragments using sterile scissors. 35 µL of alkaline buffer (0.1 N NaOH, 0.3 mM EDTA, pH 13.0) was then added into the tube, followed by centrifugation for 1 min at 110,000 g and incubation at 95 °C for 20 min. 65 µL of neutralization buffer (0.1 M Tris-HCl, pH 7.0) was then added to the tubes, followed by vigorous vortexing and centrifugation (110,000 g, 1 min). DNA was extracted from scat using a QIAamp DNA Stool Mini kit (QIAGEN, USA) following the manufacturer’s protocol. DNA was extracted from blowfly guts using a NucleoSpin Tissue kit (Macherey-Nagel, Germany), following the manufacturer’s instructions.

The extracted DNA samples from hair, bat tissue and scat samples were used for PCR in standard protocols following Wilson (2012) and Wilson et al., (2014). Primers VF1d_t1/ VR1d_t1 (Ivanova et al., 2012) were used for a first pass and RonM/ VR1 for a second pass (Ivanova et al., 2012). PCR was performed using EconoTaq PLUS 2X mastermix (Lucigen, USA) and COI Fast thermocycling program (Wilson, 2012), with slight modification of the amount of DNA template added (0.5-2 µL) depending on DNA extraction method. PCR products were then Sanger-sequenced by a local company (MYTACG-Kuala Lumpur, Malaysia) using the reverse PCR primer. Each sequence generated from individual specimens was trimmed of primers and edited using CodonCode Aligner (CondonCode Corp., USA).
For blowfly-derived DNA, a 205 bp COI amplicon was generated using a two-step PCR that utilises universal mammal primers, which are Uni-Mini-bar F and RonPing R with partial Illumina adaptor sequences, TCGTGGCGAGTCAGATGTGTAAGAAGACAG and GTCTCGTGGGCTCGGAGATGTGTAAGAAGACAG incorporated into 5’ end of forward and reverse primers, respectively. EconoTaq PLUS 2X Master Mix (Lucigen, USA) and COI Fast thermocycling program (Wilson, 2012) were used to generate amplicons from 1 µL of DNA extract (five replicate PCRs were performed for each pooled DNA extract, together with a negative control). PCR products were visualised on a 2% agarose gel stained with 1 x GelRed (Biotium, USA) and gels were extracted and purified using a NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Germany), following the manufacturer’s instructions. Of the 16 pooled DNA extracts, seven produced PCR products and were selected for high-throughput (next-generation) DNA sequencing (HTS). The purified PCR products were used as templates for a second round of PCR reaction to generate amplicons containing Illumina adaptors and unique dual-index MID tags. Briefly, 6 uL of the purified PCR product was mixed with 10 µL of NEBNext® High-Fidelity 2X Master Mix (New England Biolabs, USA) and 2 µL of each Nextera XT N70X and N50Y dual-index primers (Illumina, San Diego, CA) to make up a total of 20 uL reaction volume. The thermal cycling conditions were 98°C for 30s, followed by 6 cycles of 98°C for 10 s, 65°C for 30 s, 72°C for 30 s and a final extension at 72°C for 5 min. The PCR products were subsequently purified and size-selected using 0.8x volume ratio of AMPure XT® beads (Beckman Coulter Inc., USA). The final amplicons from all samples were quantified using a KAPA Library Quantification kit (KAPA Biosystems, CapeTown, South Africa), normalised, pooled and then sequenced on the Illumina MiSeq Sequencer (model number: M02133; 2 x 250 bp paired-end read setting) located at the Monash University Malaysia Genomics
Facility. The sequencing run produced 5,541,198 paired end reads corresponding to 2,403 megabases data output. The MiSeq outputs (FASTQ) were submitted to the NCBI Sequence Read Archive under project accession number SRP064503. The Illumina reads were grouped were demultiplexed and trimmed of adaptors onboard the Miseq using the Miseq Reporter Software. The ‘raw’ output of paired-end reads was quality filtered using the online PRINSEQ (Schmieder & Edwards, 2011) by sequence length (75-350 bp), minimum mean quality scores (>25), and GC content (20-50%). Reads with low complexity (threshold using Entropy=80) and characters other than A, C, T, G were excluded. Redundant reads (exact sequence duplicates, 5’ sequence duplicates, 3’ sequence duplicates, reverse complement exact sequence duplicates, reverse complement 5’/3’ sequence duplicates) were dereplicated. Reads with mean scores of less than 5 (sliding window size =5, step size =5) or minimum tail length of poly-A/T tails of five at 5’end and five at 3’end were trimmed of low-quality bases at 5’end (1 bp) and 3’end (1 bp). The output data (in FASTA format) was then further filtered and trimmed of primer sequences with quality control and filtering steps performed manually in CodonCode Aligner (CondonCode Corp., USA) following Brandon-Mong et al., (2015).

The resulting DNA barcodes obtained from all individual specimens collected from cage traps, mist nets, and scat collection, and unique DNA metabarcodes from blowfly-derived DNA were uploaded to the Barcode of Life Datasystems (BOLD) (Ratnasingham & Hebert, 2007; Ratnasingham & Hebert, 2013) and are available in the public dataset DS-BDNGS. Linnaean species names were assigned to our DNA barcodes (individual specimens collected from cage traps, mist nets, and scat collection) and a representative DNA metabarcode (for each Operational Taxonomic Unit (OTU) recovered from blowfly-derived DNA) when they had sequence similarity matches of >98% to DNA barcodes with Linnaean species names (submitted by other BOLD
users) in the full database of BOLD (see Wilson et al., 2014). In the case of conflicts i.e.,
the DNA (meta)barcodes had sequence similarity matches of >98% to reference DNA
barcodes with several different Linnaean species names, I assigned the BIN number
(Ratnasingham & Hebert, 2013) of the matching reference DNA barcodes to the DNA
(meta)barcodes without using a species epithet. When the DNA (meta)barcodes had
sequence similarity matches of <98% but >94% with reference DNA barcodes, I
assigned the genus name of the closest matching reference DNA barcode to the DNA
(meta)barcodes (see Zeale et al., 2011). When the DNA (meta)barcodes had sequence
similarity matches of <94% but >90% with reference DNA barcodes, I assigned the
order name of the closest matching reference DNA barcode to the DNA (meta)barcodes
(see Zeale et al., 2011). DNA (meta)barcodes sharing <90% sequence similarity to the
closest matching DNA barcode in BOLD were discarded from further analyses but are
discussed anecdotally below.

The percentage of potentially detectable species in the UGFR checklist
(Appendix A) detected by each field method was calculated. The expected species
richness (using Chao1) for each field method was computed in EstimateS Version 9.1.0
(Colwell, 2006). Rarefaction curves of expected species richness (using Chao 1) with 95%
confidence intervals were generated based on the cumulative sampling days using
EstimateS Version 9.1.0 (Colwell, 2006). The sampling completeness ratio (observed
species richness/expected species richness; Soberón et al., 2000) was calculated for
each method.

6.3 Results

A total of 14 mammal species from the orders Artiodactyla, Chiroptera, Primates,
and Rodentia (Figure 6.2) were detected from a total effort of 1,440 trap days (320 from
cage traps; 160 from mist nets; 320 from hair traps; 320 from scat collection; 320 from
blowfly traps). Mist nets and blowfly traps contributed the most species-rich samples with six species detected each, followed by cage traps and scat collection that detected two species each; hair traps detected no species (Figure 6.2). Blowfly traps recorded species from four orders - Artiodactyla, Chiroptera, Primates, and Rodentia (Figure 6.2). The other field methods (except hair traps that detected no species) detected only a single order each. Two domesticated mammal species, *Bos* [BOLD:AAA2294] and *Felis* [BOLD:AAC2892], and non-mammal taxa including one bird species, *Gallus gallus* [BOLD:AAA3630], a midge species, and a fish species, were also detected in blowfly-derived DNA but these were excluded from further analyses as my focus was the detection of wild mammal species. Of 20 potentially detectable species of small-bodied rodents and treeshrews recorded in the UGFR checklist (Appendix A), both cage traps and blowfly traps detected two species (10.0%). Of 57 potentially detectable bats recorded in the UGFR checklist (Appendix A), mist nets detected six species (10.5%) and blowfly traps detected two species (3.5%). Of 45 potentially detectable non-volant mammals recorded in the UGFR checklist (Appendix A), blowfly traps detected the highest number of species (n=4; 8.9%). Blowfly traps detected one new record of non-volant mammal for the UGFR checklist - *Trachypithecus obscurus* [BOLD:AAI0540]. Sampling completeness ratios for blowfly traps, cage traps and scat collection were 1.00, and 0.86 for mist nets (The species richness estimated from Chao 1 is similar with the species richness from Chao 2; Chao 1 is suitable for species richness estimation of our sampling sites as it corrects for bias due to sampling effort and species rarity; Chao, 1984; Iknayan et al., 2014; Figure 6.3). The HTS data included reads from five blowfly species - *Chrysomya* [BOLD:AAC4787], [BOLD:ACD5557], and [BOLD:ACF0516], and *Lucilia* [BOLD:AAA6618] and [BOLD:ACQ1337].
Figure 6.2: Species detected by different field methods at Ulu Gombak Forest Reserve, including species unique to each method and species shared across different methods. The number [in square parentheses] represents the BIN number assigned in BOLD to the species detected.
Figure 6.3: Rarefaction curves showing expected species richness (using chao1 with number of randomizations = 100) and sampling completeness ratios (in brackets) for each field method at Ulu Gombak Forest Reserve. Dashed lines showed 95% confidence intervals of expected species richness for each field method: dashed blue lines for cage traps; dashed red lines for mist nets; dashed green lines for scat collection; dashed purple lines for blowfly traps.
6.4 Discussion

Blowfly traps and mist nets detected the highest number of wild mammal species (6 species each). These were followed by cage traps and scat collection, which detected two species each. The hair traps did not detect any species, but the difficulties associated with obtaining mammal DNA samples from hair traps has already been documented in Malaysia (see Hedges et al., 2015b). There was no difference in the number of non-volant small-bodied mammal species detected by blowfly traps and cage traps (2 species). Only one species was detected by multiple field methods (cage traps, scat collection and blowfly traps) – *Rattus* [BOLD:AAB2208]. Blowfly traps were less effective than mist nets in detecting bat species (2 versus 6 of 57 in the checklist). The effectiveness of mist nets in capturing bat species, and consequently making a significant contribution to overall mammal species richness estimates, has also been shown in other studies in Peninsular Malaysia (Jayaraj et al., 2012; Jayaraj et al., 2013). The number of species detected by our mist nets was similar with another study at UGFR that detected seven species using ten mist nets and four harp traps over 9 trap days (Sing et al., 2013). There was no species detected by both mist nets and blowfly traps.

The blowfly traps detected the highest number of orders (4 orders), in contrast to the traditional methods that detected species from only a single mammal order each. Cage traps can potentially detect small non-volant mammals of several orders but in our study only detected Rodentia; probably due to the placement of cage traps on the ground where they can be avoided by arboreal mammals (see Fontúrbel, 2010). Mist nets are restricted in only being able to detect Chiroptera. Scat collection should, theoretically, not be restricted to detecting any specific mammal orders but likely depends on the range size of the species relative to the length of transect patrolled, and in my study only detected Rodentia. The number of orders detected by blowfly traps increases (to 5
orders) when the two domestic mammal species were included - *Bos* [BOLD:AAA2294], from Artiodactyla, and *Felis* [BOLD:AAC2892], from Carnivora.

The detection of a broad diversity of mammals and other vertebrate orders from blowfly traps is consistent with the study by Calvignac-Spencer *et al.*, (2012), who detected six mammal orders, an avian order and an amphibian order from blowfly-derived DNA sampled at Taï National Park, Côte d'Ivoire, and Kirindy Forest, Madagascar. Although these results demonstrate the broad range of blowfly hosts, biases in PCR primer binding affinity may lead to sequences of certain species being amplified less efficiently than others (Deagle *et al.*, 2014; Thomsen & Willerslev, 2015). This may limit the detections to species with higher primer binding affinity, resulting in species diversity being underrepresented (Shokralla *et al.*, 2012).

One further uncertainty in the use of blowfly-derived DNA for mammal diversity assessments is potential blowfly feeding biases and dispersal distances. Surprisingly the blowfly traps did not detect *Macaca fascicularis*, the long-tailed macaque, despite our frequent observation of numerous individuals at the site throughout the sampling period. Nevertheless, the blowfly traps did detect a different primate - *Trachypithecus obscurus* [BOLD:AAI0540], the dusky leaf monkey, which is a near threatened species (IUCN, 2015). This species is known to occur in the vicinity of UGFR, but had not been reported from the reserve itself. It is possible, however, that the blowflies fed on *T. obscurus* tissue or faeces outside the study site before moving into my sampling area. The few studies on blowfly dispersal (from temperate and subtropical regions - Braack & Retief, 1986; Smith & Wall, 1998; Tsuda *et al.*, 2009) suggested individuals travel 100-2,400 m per day, but there is no data available on blowfly travelling distances in the tropics. It is also not known whether blowflies, as opportunistic feeders (Calvignac-Spencer *et al.*, 2013), are more likely to have fed on carcasses, wounds or faeces. If feeding on carnivore faeces, blowflies may yield DNA
of consumed prey that may have present far from sampling location (Schnell et al., 2015). All field methods recovered a low percentage of their potentially detectable species from the UGFR checklist (2-10%), which may be an artefact of the short sampling period and restricted spatial scale of sampling relative to the total size of forest reserve, rather than feeding or primer biases. However, the study site has more than 20 years of logging history (Nor Hashim & Ramli, 2013) and megafauna recorded in 1961 (Medway, 1966), such as *E. maximus*, is highly likely to be extirpated as no records of activities or populations have been published since.

Another current limitation of the blowfly-derived DNA approach is the reliance on public DNA databases to assign detected species with formal species names (Collins & Cruickshank, 2012). Five out of 14 wild mammal species (36%) detected could not be assigned traditional species names. For example, *Rattus* [under BIN - BOLD:AAB2208] detected from blowfly-derived DNA, cage traps and scat collection likely represents *Rattus “sp. R3”* (Pages et al., 2010), a classic “dark taxon” (see Wilson et al., 2014) whereby it has not yet been formally described and is likely to be reported as *R. andamanensis, R. argentiventer* or *R. tanezumi* in studies relying on visual diagnosis (Robins et al., 2007; Pages et al., 2010; Latinne et al., 2013; Li et al., 2015). The BIN associated with this taxon comprises DNA barcode members under the traditional names *Rattus tanezumi, Rattus tiomanicus, Rattus rattus, Rattus sp.,* and *Rattus sp. abtc47994* in BOLD. The two DNA barcodes named *Rattus rattus* (JF459864 and JF459865 in GenBank) had been flagged earlier as problematic and attributed to misidentification or contamination (Shen et al., 2013). Only formal genus name could be assigned to a DNA metabarcode detected from blowfly-derived DNA whose closest match was *Callosciurus notatus* [under BIN - BOLD:AAF8467] (93-94% sequence similarity).
The utility of blowfly-derived DNA in identifying threatened mammal species in tropical forests remains questionable. However, this is also true for other field methods (e.g., cage traps, mist nets, and scat collection used in my study) that also rely on DNA-based identification for accurate assignment of formal species names (Wilson et al., 2014). It also remains to be seen whether blowfly-derived DNA can yield nuclear DNA of appropriate quality for individual identification (Calvignac-Spencer et al., 2013).

The sampling completeness ratio for all field methods exceeded 0.86. One reason for the high values for sampling completeness of the traditional field methods, despite low detection of the known species present, could be the limited distribution of traps over representative habitats in the study site, as some mammal species may be restricted to specific habitat types (Geier & Best, 1980; Goulart et al., 2009). The dispersal of blowflies could in theory overcome the problem of localised trap placement as blowflies will disperse across all habitat types (Norris, 1965). This could explain the detection of *T. obscurus* in blowfly traps as noted above. Alternatively, for all the field methods, except blowfly traps and potentially hair traps and scat collection, it is relatively safe to conclude that the species detected was present at the precise location of the trap during a known (short) time interval.

Deciding on the most cost-effective technique to conduct a mammal diversity assessment in tropical forests depends on many factors. For example, blowfly traps are more prone to damage but can be easily replaced at low cost (< $1 each). During my study, six blowfly traps and four cage traps ($3 each) were lost. Nonetheless, there is high initial cost involved in purchasing reagents and equipment for molecular detection (DNA extraction, PCR, and next generation sequencing), as well as personnel costs for analysing samples. However, species identification of bats detected from mist nets and harp traps, and rodents detected from cage traps (as exemplified by the case of *Rattus* [BOLD:AAB2208]) with an acceptable degree of accuracy also requires DNA
barcoding (Wilson et al., 2014), together with the associated costs. High-throughput-sequencing can now be outsourced to private companies and/or research institutes at increasingly competitive rates (see Brandon-Mong et al., 2015).
CHAPTER 7: FIELD-CALIBRATION OF BLOWFLY-DERIVED DNA AGAINST CAMERA TRAPS

7.1 Introduction

Camera traps have been used in wildlife monitoring for more than 40 years (Karanth & Nichols, 1998; Cutler & Swann, 1999), and have become an increasingly popular method in mammal surveillance across the globe, particularly during the past two decades (Rowcliffe & Carbone, 2008; Kucera & Barrett, 2011). Camera traps are effective in detecting medium to large mammal species, and can detect species that are rarely recorded from live traps or direct observation (Bernard et al., 2013) (e.g., Diplogale hosei, known as Hose’s civet; Bernard et al., 2013). The popularity of camera traps is partly due to the non-invasive requirements and the cost-effectiveness of camera traps compared to other survey methods such as track counts or scat collection (Balme et al., 2009; Kucera & Barrett, 2011; Foster & Harmsen, 2012; Mann et al., 2014). In addition, the functionality of camera traps is seldom restricted by environmental factors (e.g., topography, substrate or climatic conditions) known to limit the effectiveness of other survey methods (Funston et al., 2010).

In tropical forests of Peninsular Malaysia, camera traps have emerged as a popular mammal monitoring tool (e.g., Mohd-Azlan & Sharma, 2003; Numata et al., 2005; Mohd-Azlan, 2009), and have been widely used in population estimates (Hedges et al., 2015a; Mohamad et al., 2015), species distribution assessments (Sharma, 2003; Clements et al., 2012), determining activity patterns of mammals (Laidlaw & Shaharuddin, 1999; Mohd-Azlan & Sharma, 2006), and assessing environmental impacts of road-building or logging on mammals (Clements, 2013; Clements et al., 2014; Yamada et al., 2014). However, most camera trapping studies in Peninsular Malaysia have been restricted to “mega” projects involving flagship and charismatic...
mammals due to the expense of purchasing imported camera traps (Mohd-Azlan, 2009). Camera trap-derived estimates may underrepresent mammal species that occupy a specific habitat type unless cameras are distributed over all representative habitats (Bernard et al., 2013). Rapid assessment of mammal diversity using camera traps as suggested by Silveira et al., (2003) for flat grassland habitats is not feasible for tropical forests of Peninsular Malaysia (Mohd-Azlan, 2009), as exemplified by the camera trapping study of Mohd-Azlan & Sharma (2006) that needed 16 months with approximately 4600 effective trap nights to reach a plateau in the species accumulation curve in Jerangau Forest Reserve.

iDNA from blowflies has been suggested as a rapid mammal diversity assessment tool (Calvignac-Spencer et al., 2012; Calvignac-Spencer et al., 2013; Schnell et al., 2015). Blowfly-derived DNA may even overcome the ecological challenges of camera traps as blowflies are not restricted to any habitat (Norris, 1965), and potentially target all mammals being generalist feeders (Calvignac-Spencer et al., 2012; Schnell et al., 2012; Azwandi et al., 2013; Calvignac-Spencer et al., 2013). However, field-calibration of blowfly-derived DNA against camera traps has not been conducted to date.

The objective of this chapter was to calibrate the performance of blowfly-derived mammal DNA at generating species inventories and richness estimates against the most popular method to assess megafaunal diversity - camera trapping in a tropical forest reserve with megafauna in Peninsular Malaysia.

7.2 Materials and methods

A field calibration was conducted in a forest reserve of Peninsular Malaysia - Tembat Forest Reserve (TFR) in Terengganu. TFR, comprising logged dipterocarp forest (5°11'N, 102°41'E), is one of 17 ecological linkages recognised in the Malaysia
Federal Government’s “Central Forest Spine Master Plan for Ecological Linkages” to restore connectivity between four fragmented forest complexes (DTCP & DOF 2012).

A field survey was conducted between 11 April 2015 and 27 May 2015 (30 days in total). A 28 km transect was used incorporating (a) baited blowfly traps and (b) camera traps (Figure 7.1).

a) Twenty baited blowfly traps were set along the transect (Figure 7.1), 2 m above the ground, 500 m apart. As above, the traps were emptied every 24 h and collected blowflies (Family: Calliphoridae) were frozen at -20 °C within 5 h of trap emptying. The guts of all collected blowflies (1,345) were then dissected with sterile implements and pooled (five individual blowfly guts per tube) for DNA extraction (269 extractions in total). The DNA extracts were further pooled by sampling day resulting in 30 pooled DNA extracts for bulk PCR.

b) Twenty Panthera V4 camera traps with passive infrared motion sensor were deployed along the transect (Figure 7.1), attached to tree trunks approximately 50 cm above ground level. Camera traps were placed on linear features known to have high detection probabilities for large mammals (e.g., animal trail, ridge or old logging road; Harmsen et al., 2010). No attractants were employed around the cameras. The cameras were triggered by motion, and only photocaptures that were obtained during the sampling period of the blowfly traps (30 days) were used for analysis and were catalogued using Camera Base version 1.4 software (Tobler, 2012).

Methods for the processing of blowfly samples for mammal species detection followed those from Chapter 6.

Identification of species from the photocaptures (Figure 7.2) was carried out by LPS and GRC based on morphological characters with the aid of Francis (2008). Photocaptures of mammals that could not be identified with certainty due to poor lighting or blurred images were excluded.
Figure 7.1: The location of Tembat Forest Reserve showing the locations of twenty trap stations each with a blowfly trap and camera trap.
Figure 7.2: Photographic evidence of nine mammal species detected by camera traps at Tembat Forest Reserve, Terengganu. A) *Elephas maximus*, B) *Helarctos malayanus*, C) *Macaca fascicularis*, D) *Muntiacus muntjak*, E) *Panthera pardus*, F) *Prionailurus bengalensis*, G) *Sus scrofa*, H) *Tapirus indicus*, and I) *Viverra zibertha*. Camera settings: sample rate of passive infrared sensor = 70 times per s; flash brightness = normal (Xenon flash system); LED stealth mode = off; daylight images per trigger = 3 with delay period between images = 1 s; flash images per trigger during night time = 1 with delay period between images = 20 s; image resolution: 3 megapixels.
7.3 Results

A total of 20 mammal species from the orders Artiodactyla, Carnivora, Cetartiodactyla, Chiroptera, Perissodactyla, Primates, Proboscidea, and Rodentia were detected from a total effort of 1,200 trap days (600 from camera traps; 600 from blowfly traps; see Figure 7.3). Blowfly traps detected a higher number of species (n=11) than camera traps (n=9), including five orders - Artiodactyla, Carnivora, Chiroptera, Primates, and Rodentia, while camera traps detected six orders - Artiodactyla, Carnivora, Cetartiodactyla, Perissodactyla, Primates, and Proboscidea. While only one potentially domesticated mammal species was detected - *Bos* [BOLD:AAA2294], its identity could also be that of a wild bovid - *Bos gaurus*. Non-mammal taxa detected from blowfly-derived DNA include a bird species - *Gallus gallus* [BOLD:AAA3630], three fish species, a lizard species, a snake species, another squamate, and a turtle species, but were excluded from further analyses. Additional non-mammal taxa detected, but which had sequence similarity matches of <90% with DNA barcodes in BOLD include an algae (76-80% sequence similarity) and another snake (89% sequence similarity); these were also excluded from further analyses. Sampling completeness ratio of blowfly traps was 0.79, and 1.00 for camera traps (The species richness estimated from Chao 1 is similar with the species richness from Chao 2; see Chapter 6; Figure 7.4). The HTS data included reads from five blowfly species - *Chrysomya* [BOLD:AAA5667], [BOLD:ACD5557], [BOLD:ACF0516], and [BOLD:AAB3064], and *Lucilia* [BOLD:ACQ1337].
Figure 7.3: Species detected by different field methods at Tembat Forest Reserve. The number [in square parentheses] represents the BIN number assigned in BOLD to the species detected.
Figure 7.4: Rarefaction curves showing expected species richness (using chao1 with number of randomizations = 100) and sampling completeness ratios (in brackets) for each field method at Tembat Forest Reserve. Dashed lines showed 95% confidence intervals of expected species richness for each field method: dashed orange lines for camera traps; dashed black lines for blowfly traps.
7.4 Discussion

Blowfly traps detected higher species richness than camera traps - 11 species from the blowfly traps versus 9 species (based on visual identification) from the camera traps. Blowfly traps detected more mammal species that were more volant and arboreal in nature than those detected by camera traps. Only one species was potentially detected by both blowfly traps and camera traps – \textit{Sus} [BOLD:AAA3445] which is probably “\textit{Sus scrofa}” from the camera traps. The blowfly traps detected fewer orders compared to the camera traps (5 versus 6 orders). However, the comparison may be less meaningful given the lack of taxonomic equivalency (Bertrand et al., 2006) between the orders of large charismatic megafauna (specifically targeted by the camera traps) and other mammal orders. For example, Perissodactyla and Proboscidea are orders containing only a single extant species in Peninsular Malaysia. Nine vertebrate orders were detected by blowfly traps when the bird species - \textit{G. gallus} [BOLD:AAA3630], three fish species, a lizard species, a snake species, another squamate and a turtle species, whose DNA was also detected in the blowfly guts, were included.

One uncertainty in the use of blowfly-derived DNA approach in assessing mammal diversity is that the assignment of formal species names to mammal species detected from blowfly-derived DNA depends heavily on the public DNA databases. Six out of 20 species (30%) detected could not be assigned traditional species names. For example, \textit{Rhinolophus sp. 25Ne} showed no species-level match (>98%) in BOLD despite many bat DNA barcodes from Peninsular Malaysia being present in BOLD (Lim & Wilson, 2015). \textit{Sus sp. 4Ba}, also showed no species-level match (>98%) in BOLD, despite the presence of 218 DNA barcodes of \textit{Sus}, suggesting a cryptic species. Only formal genus names could be assigned to DNA metabarcodes detected from blowfly-derived DNA whose closest matches were \textit{Callosciurus notatus} [under BIN - BOLD:AAF8467] (93-94% sequence similarity), and \textit{Paradoxurus hermaphroditus}
It remains in question whether blowfly-derived DNA approach could be used to identify threatened species in tropical forests. However, as species boundaries are increasingly being refined and recognised on the basis of molecular data, especially for species-rich groups (Francis et al., 2010; Pages et al., 2010), but even for charismatic megafauna (Wilting et al., 2015), field methods that rely solely on visual diagnosis (e.g., camera traps) will have increasingly limited utility. Currently, camera trapping is the most popular approach to monitor populations of mammal species, but this is restricted to species that have individuals with unique markings (e.g., spotted and striped felids; Henschel & Ray, 2003; Maffei et al., 2005; Hedges et al., 2015a), or to individuals that can be marked prior to camera trapping (Trolle & Kery, 2003; Rowcliffe et al., 2008).

The sampling completeness ratio of camera traps was 1 and blowfly traps was 0.79, suggesting that with longer sampling effort, blowfly traps may continue to produce higher species richness estimates. The high value for sampling completeness of the camera trapping could be due to the camera traps could not be distributed over all representative habitats in the study site, where some mammal species only live in certain types of habitat (Geier & Best, 1980; Goulart et al., 2009). Blowfly dispersal capability could theoretically overcome the challenge of localised camera trap placement as blowflies are distributed in all habitat types (Norris, 1965). However, for camera traps but not blowfly traps, the species detected was assumed to be at the exact location of the trap for a known (short) period, which is an advantage of camera traps over blowfly traps.

One of the main factors to be considered in choosing a suitable mammal diversity assessment tool in tropical forests is cost-effectiveness. Blowfly traps are
cheaper in cost (< $1 each) than low-end camera traps ($250 each) in my study, but blowfly traps pose higher risk of damage than camera traps. Six blowfly traps and two camera traps were lost during my study. Camera traps may be more prone to theft than blowfly traps, and can leave a significant financial dent in the project when that happens (high-end camera traps can cost up to $1000; Rovero et al., 2013), but are less prone to damage. Nonetheless, blowfly-derived DNA approach required molecular detection involving high cost of purchasing reagents and equipment as well as costs of trained personnels in genetic analysis. The costs of sample processing and analysis with the advance of high-throughput sequencing may be reduced through collaboration with research institutes and/or outsourcing to private companies at competitive prices.
Mammal DNA derived from invertebrates has emerged as a promising alternative to traditional field methods but some uncertainties in developing the approach as a standardised mammal monitoring tool need to be addressed.

My first chapter evaluated the potential of blowfly-derived DNA as a mammal monitoring tool by determining the persistence period of amplifiable mammal DNA in blowflies - a key factor for successful detection of ingested DNA (Calvignac-Spencer et al., 2013). The results indicate that amplifiable mammal mtDNA persists in the guts of adult *Chrysomya megacephala* for 24-96 h post feeding. This was also the first experimental indication that successful detection of mammal DNA from blowflies is due to mammal DNA in their guts as opposed to mammal DNA being carried on their exoskeleton as a result of landing on mammal tissues or faeces. The success in detecting mammal DNA from wild-caught blowflies trapped over a period of 24 hours indicates the potential of a standardised blowfly trapping protocol with retrieval of blowflies every 24 h for effective mammal monitoring in the field.

Another key factor for developing successful iDNA approaches is selection of the target DNA region for amplification. My second chapter proposed a new primer pair, Uni-Mini-bar F/ RonPing that amplifies a 205 bp fragment of *COI*, which appears to be an optimal length for a DNA mini-barcode allowing amplification from degraded samples, such as iDNA, while not suffering a reduction in the ability to distinguish species. This target also falls within the maximum read length for high-throughput sequencing. The new primer combination was able to amplify a higher proportion of mammal species than previously proposed primer combinations - Uni-Minibar F/ Uni-Minibar R (Meusnier et al., 2008). Another significant advantage of the Uni-Minibar F/RonPing combination is the fact that it did not amplify *Chrysomya megacephala COI*, even when the ratio of blowfly DNA to mammal DNA was high. The 205 bp *COI*
fragment was successful in distinguishing nearly all examined mammal species from Peninsular Malaysia and separating 26 “dark” bat taxa, previously recognised species that lack formal taxonomic status (Wilson et al., 2014). This suggested the potential of detecting cryptic taxa overlooked by traditional methods (Francis et al., 2010).

My third chapter determined dispersal range of blowflies to address an uncertainty in the use of blowfly-derived DNA approach - the actual or potential location of target species relative to the location where blowflies were collected. The dispersal range of Chrysomya spp. in my study was between 1 km to 3 km during recapture period of 9 days, suggested that the location of mammal species detected from blowfly-derived DNA could be within 3 km range of sampling location where blowflies were collected over 24 h. However, blowfly dispersal capability could be affected by environmental factors such as solar radiation, rainfall, temperature, and wind activity (Von Aesch et al., 2003; Tsuda et al., 2009), resulting in uncertainties of the precise location of target species detected from blowfly-derived DNA.

My fourth and fifth chapters calibrated the performance of blowfly-derived mammal DNA at generating species inventories and richness estimates against five traditional field methods – cage traps, mist nets, hair traps, scat collection, and camera traps in tropical forest reserves of Peninsular Malaysia. After 1,440 trap days in Ulu Gombak Forest Reserve, blowfly traps and mist nets detected the highest number of wild mammal species (6 species each), followed by cage traps and scat collection that detected two species each, and hair traps that detected none. The fourth chapter revealed that blowfly traps detected the highest number of orders (4 orders), in contrast to the traditional methods that detected species from only a single mammal order each. Only one species was detected by multiple methods, and compared to traditional field methods, blowfly-derived DNA detected both volant and non-volant species from wider body size ranges. This is supported by my fifth chapter that compared blowfly-derived
DNA with the most popular method to assess megafauna diversity - camera traps - in another tropical forest reserve of Peninsular Malaysia with megafauna. After 1,200 trap days, blowfly-derived DNA detected more species (n=11) than camera traps (n=9), with only one species detected by both methods. While the rarefaction curves of both fourth and fifth chapters indicated that blowfly-derived DNA would continue to detect more species with greater sampling effort, the lack of overlap in species detected by different field methods suggested employing multiple field methods may be the fastest way to obtain a representative account of species. However, there are some uncertainties in the field-application of blowfly-derived DNA for mammal diversity assessments – potential biases in blowfly dispersal and primer binding affinity that may lead to underrepresentation of mammal diversity, comparison of iDNA with other field methods in terms of cost-effectiveness, and common problem of discordance between molecular and morphological taxonomy in sequence reference libraries.

With further calibration, blowfly-derived DNA may join the list of traditional field methods. Areas for further investigation include blowfly feeding and dispersal biology, biases in primer binding affinity, and the assembly of a comprehensive and taxonomically-consistent DNA barcode reference library.
REFERENCES


DTCP, & DOF (Department of Town and Country Planning & Department of Forestry). (2012). *Central forest spine: Summary of master plan for ecological linkages*. Kuala Lumpur, Malaysia: DTCP.


Hassall, C., & Thompson, D. J. (2012). Study design and mark-recapture estimates of dispersal: a case study with the endangered damselfly *Coenagrion mercuriale*. *Journal of Insect Conservation, 16*(1), 111-120.


LIST OF PUBLICATIONS AND PAPER PRESENTED

