MOLECULAR DIET ANALYSIS OF THE HOUSE-FARM SWIFTLETS (APODIDAE, COLLOCALIINI) IN PERAK, MALAYSIA

CHAN KOK SIM

MASTER OF SCIENCE

FACULTY OF SCIENCE
UNIVERSITI TUNKU ABDUL RAHMAN
JANUARY 2019

MOLECULAR DIET ANALYSIS OF THE HOUSE-FARM SWIFTLETS (APODIDAE, COLLOCALIINI) IN PERAK, MALAYSIA

By

CHAN KOK SIM

A thesis submitted to the Department of Biological Science,
Faculty of Science,
Universiti Tunku Abdul Rahman,
in partial fulfillment of the requirements for the degree of
Master of Science
January 2019

Dedication

Thanks to my dad and mum, Jacky and Jenny

for your unconditional love, comfort, support and trust.

"Trust in the Lord with all your heart

and lean not on your own understanding;

in all your ways submit to him,

and he will make your paths straight."

Proverbs 3:5-6

ABSTRACT

MOLECULAR DIET ANALYSIS OF THE HOUSE-FARM SWIFTLETS (APODIDAE, COLLOCALIINI) IN PERAK, MALAYSIA

Chan Kok Sim

House-farming of the edible white-nest swiftlets has been a lucrative industry in Malaysia since three decades ago. This form of semi-captive farming allows the swiftlets to forage outside the swiftlet houses. While it is generally known that swiftlets are insectivores that feed opportunistically, studies on the feeding biology and the diet profiles of these birds are scarce. The present study aims to assess the diet profiles of the house-farm swiftlets using high-throughput sequencing approach, followed by comparison of the diet profiles in different landscapes in Perak, Malaysia. A preliminary assessment of two sets of metabarcoding mitochondrial COI (Cytochrome-c oxidase subunit I) primers was conducted using DNA cloning on the freshly collected swiftlet faeces. The screening indicated that the mICoIintF/HCO2198 primer-pair (for "mICo" region) showed bias towards feather mites, fungal contaminants and swiftlet DNA. Conversely, the LepF1/MLepF1_Rev primer-pair (for "Lep" region) was able to amplify the partial COI sequence of various arthropods in the swiftlet faeces. A bioinformatic pipeline was developed based on the cloning results and can be summarised into four steps for taxonomic assignment of arthropod COI region. High-throughput sequencing using LepF1/MLepF1_Rev for the 218-bp mitochondrial COI region was then performed for the swiftlet faecal samples

collected from six swiftlet farms that represented three landscape types, i.e., monocrop, urban and mixed-used landscape. A total of 4,852 operational taxonomic units (OTUs) were generated, out of which 266 were arthropod DNA sequences. Following the taxonomic assignment pipeline established in the preliminary assessment, the diet of the house-farm swiftlets comprised of Diptera (62.74%), followed by Hemiptera (18.87%), Coleoptera (12.26%), Lepidoptera (2.36%), Hymenoptera (1.89%), Blattodea (0.94%) and Odonata (0.94%). A total of 20.30% of OTUs showed no genetic affinity with their respective top BLASTn hits and were therefore of uncertain identity. This could possibly be due to the short length of the DNA marker used in distinguishing the arthropod identity. Furthermore, due to the incompleteness of the DNA database for the insects found in Malaysia, the insect taxa identified using molecular method should also be cross-checked with the species distribution record. The urban (Ipoh) and mixed-use (Pantai Remis) landscape had the most diverse arthropod orders (total of five), while the monocrop landscapes (Beruas OP1 and Beruas OP2) were shown to have the least arthropod orders (total of two). It was proposed that some of the habitats (man-made or natural) in the urban landscape of Ipoh could act as an urban green space to support the substantial arthropod diversity which could be fed upon by the house-farm swiftlets. On the other hand, the mosaic and heterogeneous mixed-use landscape could provide continuity of food source to the house-farm swiftlets. In the monocrop landscapes, the low arthropod abundance in the diet of house-farm swiftlets could be due to some of the farming practices such as the application of pesticides. Also, this study suggests that the urban and mixed-used landscape could be a relatively more ideal place for the house-farm swiftlets to persist.

The role of house-farm swiftlets to provide ecological service as insect pest predators in all three landscape types should be further investigated.

ACKNOWLEDGEMENT

I would first like to thank my supervisor Dr Goh Wei Lim. The door to her office was always open whenever I ran into a trouble spot or had a question about my research or writing. She consistently allowed this thesis to be my own work but directed me patiently in the right the direction whenever she thought I needed it. Words will never be enough to be expressed my deepest gratitude to her. Special thanks to my co-supervisor Dr Tan Ji for all his useful opinion and encouragements throughout the study.

I am also grateful to lab members: Christina, Esther, Wei Chi, Yee Teng and Zoe that never failed to make me laugh and be there for me. Perhaps our daily tea time is always the best motivation for me to do more even after working hour! My thanks also go to Yee Teng for being such a great and caring friend, encouraging me when I am down and inspire me with many creative ideas on writing. Also, thanks to Jun Ray that always providing moral support and cheering me up at tough times. My thanks to all people who have help throughout the study is indescribable.

I must express my very profound gratitude to my parents for providing me with unfailing love and prayers and continuous encouragement throughout my years of study. I would also like to extend my appreciation to my church mates: Stephen and Samuel for their continuous encouragement and prayers. Lastly, this accomplishment would not have been possible without God. I thank to the Almighty God for His comfort and guide throughout my Master studies.

APPROVAL SHEET

This dissertation/thesis entitled "**MOLECULAR DIET ANALYSIS OF THE HOUSE-FARM SWIFTLETS (APODIDAE, COLLOCALIINI) IN PERAK, MALAYSIA"** was prepared by CHAN KOK SIM and submitted as partial fulfillment of the requirements for the degree of Master of Science at Universiti Tunku Abdul Rahman.

Approved by:

_____

(Dr. Goh Wei Lim) Date: ……
Assistant Professor/Supervisor
Department of Biological Science
Faculty of Science
Universiti Tunku Abdul Rahman

_____

(Dr. Tan Ji) Date: ……
Assistant Professor/Co-supervisor
Department of Agricultural and Food Science
Faculty of Science
Universiti Tunku Abdul Rahman

FACULTY OF SCIENCE

UNIVERSITI TUNKU ABDUL RAHMAN

Date: 31 JANUARY 2019

SUBMISSION OF THESIS

It is hereby certified that **CHAN KOK SIM** (ID No: **17 ADM 01170**) has completed this thesis entitled "**MOLECULAR DIET ANALYSIS OF THE HOUSE-FARM SWIFTLETS (APODIDAE, COLLOCALIINI) IN PERAK, MALAYSIA**" under the supervision of Dr. Goh Wei Lim (Supervisor) from the Department of Biological Science, Faculty of Science, and Dr. Tan Ji (Co-Supervisor) from the Department of Agricultural and Food Science, Faculty of Science.

I understand that University will upload softcopy of my thesis in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,

_____

(*CHAN KOK SIM*)

DECLARATION

I hereby declare that the dissertation is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTAR or other institutions.

Name _____

Date _____

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

%	percentage
°C	degree Celsius
mg	milligram
µg	microgram
ng	nanogram
µL	microlitre
µM	micromolar
DNA	Deoxyribonucleic Acid
mL	millilitre
mM	milimolar
PCR	Polymerase Chain Reaction
g	Gravitational acceleration
min	Minute
s	Second
BLASTn	Basic Local Alignment Search Tool for nucleotide
NCBI	National Center for Biotechnology Information
BOLD	Barcode of Life Data System
NJ	Neighbor-joining
-bp	base pair
UV-Vis	Ultraviolet–visible
km	kilometre
GPS	Global Positioning System
ca.	circa (around)
dB	decibel

CHAPTER 1

INTRODUCTION

1.1 Swiftlet farming

House-farm swiftlets refer to the edible nest swiftlets which are attracted to roost and nest in the specifically designed buildings that mimic cave environment. Because of the "white" nest (i.e., made almost entirely of saliva) they produce and the morphological resemblance to the cave swiftlets, the house-farm swiftlets have presumably migrated from the caves and hence are commonly regarded as *Aerodramus fuciphagus*, a name which is applied to the white-nest swiftlets in the Indo-Malayan region by many ornithologists. Recent studies, however, suggested distinction between the house-farm swiftlets and the white-nest swiftlets in the natural habitats. For instance, Ramji et al. (2013) reported roosting and nesting behaviours of the house-farm swiftlets which were different from those of the cave swiftlets. Close examination on the plumage coloration of the rump revealed that the house-farm swiftlets resembled neither of the wild species / subspecies (Cranbrook et al., 2013). Genetic evidence supported the genetic isolation of the house-farm swiftlets from the cave swiftlets and suggested that the house-farm swiftlets are novel domesticates, although the precise origin of the house-farm swiftlets is yet to be confirmed (Cranbrook et al., 2013; Goh et al., 2018).

In Malaysia, the swiftlet farming industry started blooming in several towns of Peninsular Malaysia in the 1980s and is now well expanded across the country, including Sarawak and Sabah (Goh et al., 2018). At present, there are approximately 60,000 to 80,000 swiftlet houses that generate an annual income of US$ 300 million (Connolly, 2017; Tan et al., 2018). In the recent decades, decline in the house-farm swiftlet population in the urban areas has probably promoted the establishment of swiftlet farms in rural areas, such as agricultural land (paddy field, palm oil and rubber plantations), and the countryside near mangrove peat swamps and lowland dipterocarp forests. Unlike poultry that are kept in captivity, the house-farm swiftlets are free to roam and forage outside the buildings (Marzuki, 1994; Lim and Cranbrook, 2014). Some farmers believe that building swiftlet houses in the countryside is advantageous as the natural environment gives better support to the insect community and hence provides ample food supply to the house-farm swiftlets, although scientific research on the swiftlet diet is scarce.

1.2 Feeding ecology of house-farm swiftlets

Even though swiftlet farming brings lucrative incomes to the country, little information is known on the feeding ecology of the house-farm swiftlets. As the swiftlet farmers do not feed the swiftlets, they usually perceive the surrounding vegetation as the habitat that supports the dietary insect community.

Diet studies for the swiftlets are scarce, and only three studies involved the nest swiftlets collected in Malaysia (Langham, 1980; Lourie and Tompkins, 2000; Rahman et al., 2016). In their studies, Diptera (true flies), Ephemeroptera (mayflies) and Hymenoptera (mostly flying ants and fig wasps) were found abundantly in the diet of the white-nest swiftlets.

While previously studies were based on morphological observations of the food boluses regurgitated by the swiftlets, advances in molecular techniques has now permitted meta-profiling of the diet based on DNA residues that could be extracted from the stomach contents and even highly degraded faecal samples. Diet analysis using faecal samples were non-invasive and has been widely adopted in monitoring the prey items of the birds (Deagle et al., 2007; Deagle et al., 2010). Molecular meta-profiling for insects has now become feasible with the development of insect-specific DNA metabarcoding markers and the analysis pipelines in high-throughput sequencing approach (Brandon-Mong et al., 2015).

1.3 Objectives of this study

Since most house-farms then existing (71.6%) were constructed in urban areas, Othman et al. (2008) concluded that this habitat was suitable for the purpose. In Peninsular Malaysia, government policy and public pressure have since opposed house-farms in urban areas and encouraged the establishment of new swiftlet house-farms in rural countryside, including agricultural land (paddy fields, oil palm or other plantation crops) and other areas of mixed land-use

(Nurshuhada et al. 2015). With the understanding of the arthropod fed in the diet profiles of house-farm swiftlets from different landscapes, it is possible to find out the arthropods that can improve the nest quality (i.e. sialic acid content etc.). This can improve the edible-bird nest production in each swiftlet farms and boost up the export rate for the edible-bird nest to other countries. Therefore, the present study aims to understand how different landscape types (monocrop, urban and mixed-use landscape) can affect the diet composition of the house-farm swiftlets in the Peninsular Malaysian State of Perak.

Specific objectives are:

(1) To establish a useful taxonomic assignment pipeline for the dietary insects of house-farm swiftlets using the meta-profiling approach;
(2) To investigate the diet composition of the house-farm swiftlets using a high-throughput sequencing approach;
(3) To elucidate the diet profile patterns that correspond to the landscape features of swiftlet farms.

CHAPTER 2

LITERATURE REVIEW

2.1 Swiftlets

2.1.1 Classification of swiftlets

Swiftlets are classified under the family of Apodidae and subfamily of Apodinae. Within the Apodinae, the single genus *Collocalia* was further separated into three different genera, including *Aerodramus* (black, mossy and white nest swiftlet), *Collocalia* (white bellied and glossy swiftlets) and *Hydrochous* (giant waterfall swifts) (Brooke, 1970; Lee et al., 1996; Price et al., 2004; Thomassen et al., 2003; Thomassen et al., 2005). The phylogenetic evidences using NADH dehydrogenase and cytochrome-b mitochondrial genes also showed the monophyletic group of swiftlets, which formed two separated clades between genera *Aerodramus* and *Collocalia* (Price et al., 2004).

2.1.2 Distribution of swiftlets

As shown in Figure 2.1, swiftlets can be found across the Indo-Pacific region of the world. Based on the sight records, the limits of distribution extend to the west of the Madagascar and on the Seychelles islands in the Indian Ocean, and to the east on the Marquesas in the South Pacific. The sight record for the northern region covers Himachal Pradesh, North-east India, and in Sichuan, China. In the South, the sight records spread out from Mauritus in the Indian

Ocean to Queensland, Australia and to New Caledonia in the south-western Pacific Region (Lim and Cranbrook, 2014).

Within this geographical range, some of the regions in the Southeast Asian countries are currently cultivating edible-bird nests to meet the market demands. For instance (as shown in Figure 2.1), Andaman and Nicobar Islands in India, Hainan Island in China, Palawan island in Philippines, the coasts and islands of Vietnam, Sumatra, Java and the Lesser Sunda Islands of the Indonesia archipelago, multination island of Borneo, Cambodia, Thailand, Myanmar, Malaysia, and Singapore (Lim and Cranbrook, 2014).

2.1.3 Description and ecology of swiftlets

Generally, swiftlets can be found in inaccessible caves (Lim and Cranbrook, 2014). They can be distinguished with their rump colouration. The *Collocalia* have a white underbelly, with blue or green glossy plumage on the upper parts of the body (Lim and Cranbrook, 2014). All species of *Aerodramus* have dull, dark blackish or brown upperparts, in some case a pale-grey or white bar across the rump, contrasting with the grey-brown underparts (Lim and Cranbrook, 2014).

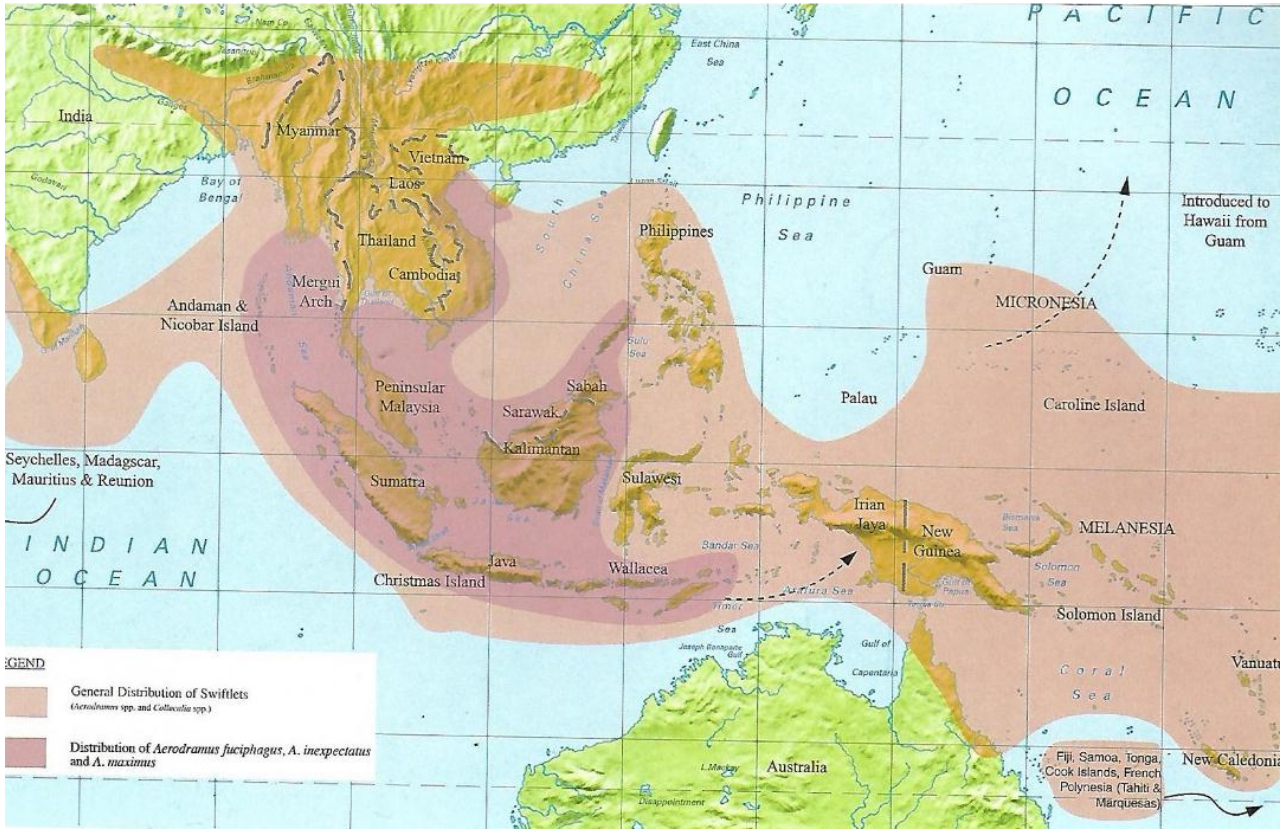


Figure 2.1 General distribution of the *Aerodramus* spp. and *Collocalia* spp. and the foraging range of the white -nest swiftlets (*A. fuciphagus* and *A. inexpectatus*) and black-nest swiftlets (*A. maximus*). Adapted from Lim and Cranbrook (2014).

Among the three genera (*Aerodramus*, *Collocalia* and *Hydrochous*) of swiftlets, *Aerodramus* was suggested to use echolocation as an orientation mechanism in the dark or poor light conditions (Cranbrook and Medway, 1965; Thomassen et al., 2005). The low frequency of the echolocation by swiftlets is insufficient to detect small objects like insect prey (Cranbrook and Medway, 1965). Previously, it was thought that only the genus *Aerodramus* can echolocate, thus, presenting a good criterion to distinguish between the genus *Aerodramus* (echolocating swiftlets) and *Collocalia* (non-echolocating swiftlets) (Brooke, 1970). However, the discovery of the echolocation ability in the *Collocalia* species (pygmy swiftlets *C. troglodytes*), showed that the echolocation ability can no longer be separated among the swiftlet genera (Price et al., 2004).

2.2 Edible-birds' nest industry

2.2.1 "White-nests" and "black-nests"

Of the 24 species of swiftlets in the world, only few species produce nests of high commercial value (Lim and Cranbrook, 2014). The edible-bird nests that were sold in high price mainly come from three cave swiftlet species namely White-nest swiftlets (*Aerodramus fuciphagus* and *A. inexpectatus*) and Black-nest swiftlets (*A. maximus*) (Lim and Cranbrook, 2014).

The nests of white-nest swiftlets are higher in value, as the nests are built from entirely pure hardened salivary nest cement. The nests of the white-nest swiftlets have few or small contour feathers from swiftlet plumage which are incorporated among the laminae of salivary nest cement. In contrast, the nests

from black-nest swiftlets have more contour and flight feathers between the laminae of the half-bowl shaped nests. Thus, the price of the collected black nest is five to six times less than the higher quality white nest (Lim and Cranbrook, 2014).

Nevertheless, edible-bird nests are known as one of the most expensive animal products, sometimes referred to "the caviar of the East". The edible-bird nest consists of 38.7% carbohydrate, 32.3% protein, 20% of inorganic ash, and 9% of moisture. Some of the bioactive compounds and glycoprotein have been extracted for screening of various medicinal properties, such as anti-aging, anti-tumour, inhibition of influenza virus, immunity enhancement etc. (Ng et al., 1986; Kong et al., 1987; Guo et al., 2006; Vimala et al., 2012).

A recent study also used edible-nest extract to investigate the neuroprotective effects of edible-bird nests in treating Parkinson's disease mice model (Yew et al., 2018). It was also shown to have anti-aging effect by increasing the activity of the antioxidant enzymes in *Drosophila melanogaster* (Hu et al., 2016). However, long term research is necessary to reveal the significance of these changes before it can be concluded that edible-bird nests-based supplements is beneficial to human health.

2.2.2 House-farm swiftlets and their origins

"Swiftlet farming" can be defined as process of obtaining edible-bird nests from the swiftlets that roost and nest in the purposely designed man-made buildings or "swiftlet houses" (Lim and Cranbrook, 2014). This process is like apiculture,

but, instead of beehives, the swiftlet house mimics the cave-like conditions, thus providing a roosting site for the swiftlets. The swiftlets are free to forage for food during the daytime, only returning to roost at night.

The practice of swiftlet farming has been long established in Java, Indonesia (Lim and Cranbrook, 2014). The first swiftlet farm was believed to be initiated at Sedayu in East Java since 1880 (Lim and Cranbrook, 2014). Little effort was carried out to attract more swiftlets into the swiftlet houses, since the colonisation of swiftlets in the houses was based on pure luck. The swiftlet houses in Java was first colonised by Linchi Swiftlets (*Collocalia linchi*) that build mossy nests. However, these houses were converted to "white-nest" swiftlet houses through the cross-fostering technique (Marzuki, 1994; Lim and Cranbrook, 2014). The breeders swapped the eggs of the white-nest swiftlet into the nests of the surrogate Linchi swiftlets.

In Peninsular Malaysia, colonies of white-nest swiftlets were sporadically set up in old shop houses in towns in the 1950s and 1960s. The towns include Sitiawan, Taiping, Nibong Tebal and Penang (Lim and Cranbrook, 2014; Langham, 1980). However, the owners of the houses had little knowledge on the swiftlet farming, thus no effort was carried out to lure the white-nest swiftlets into their houses. This situation only improved and transformed at the late 20th century. For instance, the owners modified the building structures by installing the ventilation to help in the air ventilation in the swiftlet house.

Till now, there are approximately 60,000 to 80,000 of swiftlet farms in Malaysia that generate an annual income of US$ 300 million (Connolly, 2017; Tan et al., 2018). Malaysia supplies about 75% of the edible-bird nests (around 3750 tonnes) to satisfy the global demand every year, thus becoming the second largest exporter of edible-bird nests after Indonesia (Thorburn, 2015; Connolly, 2017). Swiftlet farming is one of the 16 entry point projects that promote the establishment of market driven, industrial scale and integrated agriculture-related businesses under the National Key Economic Areas (Nurshuhada et al. 2015).

House-farm swiftlets are morphologically similar to the white-nest swiftlets in the caves. They weigh around 12.91 g ± 0.68, body length of 11.22 cm ± 0.14. with wing span length of 27.06 cm ± 0.26 (Looi et al., 2015). However, as there are different degrees of grey colouration on their rumps, i.e., they are not identical to either brown rumped-swiftlets or the grey-rumped swiftlets and they are treated as a distinct species from the cave swiftlets (Cranbrook et al., 2013; Goh et al., 2018). As all house-farm swiftlets make white edible nests, they are presumed to have originated from the white-nest swiftlets from the caves, i.e., *A. fuciphagus*. This name has been widely used in many ornithological literatures, until the re-examination of historical museum specimens and the comparison with recent house-farm swiftlets (Cranbrook at al., 2013). Cranbrook at al. (2013) demonstrated morphological and genetic differences between house-farm swiftlets and both wild species.

The house-farm swiftlets have mixture of rumped colour between two white-nest swiftlet species, *A. fuciphagus* (brown-rumped) and *A. inexpectatus* (grey-rumped), respectively (Cranbrook at al., 2013). From the genetic evidence, all house-farm populations are hypothesized to have descended from the same common origin. Recent phylogenetic analyses based on mitochondrial DNA, had shown that house-farm swiftlets in Malaysia are genetically closer to *A. fuciphagus vestitus* that are found in the caves of inland Borneo (Goh et al., 2018). These genetic studies have identified that house-farm swiftlets have genetically diverged into two clades based on the mitochondrial genes (Cranbrook et al., 2013; Goh et al., 2018). However, the genetic evidence is currently insufficient to draw a clear conclusion on the taxonomic assignment of the house-farm swiftlets into either *A. fuciphagus* or *A. inexpectatus*, and therefore, classified under *Aerodramus* spp.

In terms of behavioural change, it was suggested that house-farm swiftlet could be a new form of domestication, as white-nest swiftlets observed in Peninsular Malaysia do not occupy any natural habitats (Cranbrook et al., 2013; Goh et al., 2018). Even in limestone rich areas near Ipoh and Kuala Lumpur (Batu Caves, Selangor) where caves abound, no wild swiftlets colonies were observed (Cranbrook et al., 2013).

2.2.3 Factors for successful swiftlet farming

Factors such as habitat (threat and predators), physical and environmental factors (building structures, temperature and humidity, entry door and artificial sounds) and food resources can possibly affect the swiftlet populations in a

house. The swiftlet population is important as the increase in nest production can raise the income of the swiftlet farmers (Ibrahim et al., 2009).

Firstly, habitats play a crucial role in swiftlet faming. For instance, the swiftlet houses in a rural area was shown to have a higher risk of encountering breaking in by thieves to steal the edible white nests (Ibrahim et al., 2009). The invasion by the intruders not only create "panic" in the adult swiftlets and cause the nestlings to fall from the nests when the nests are detached. The massive intrusion can possibly decrease the swiftlets' colony size in the swiftlet house (Ibrahim et al., 2009). Furthermore, house-farm swiftlets may also encounter attacks from predators. Traps and poisonous fish have been set in the houses to prevent the nestlings from being eaten by the owls. Although the owl species that will attack the house-farm swiftlets remain uncertain, one of the possible predators could be barn owls, *Tyto alba*, as many swiftlet houses are set up near oil palm plantations. The barn owls are previously introduced to reduce rat populations in oil palm plantations (Duckett and Karuppiah, 1990; Wood and Fee, 2003; Basiron, 2007). Some other predators have been observed in natural habitats (caves) included snakes (such as vipers, pythons and cobras), bats, giant crickets and lizards (such as geckos) (Manchi and Sankaran, 2009; Lim and Cranbrook, 2014).

Secondly, past studies on the successful establishment of swiftlet houses have been largely focused on the physical and environmental factors of the buildings. For instance, Ibrahim et al. (2009) had determined the suitable environmental conditions (such as air temperature, surface temperature, relative humidity, air

velocity and light intensity) of building a swiftlet farm. They suggested that ventilation holes, humidifiers, and orientation of swiftlet entrance can provide an ideal environment condition to attract more swiftlets. The swiftlet entrance should be facing the South-North direction to avoid direct sunlight exposure which may increase the internal temperature and light intensity in the house. In addition, Rahman et al. (2018) suggested that high humidity (83.7%), low light intensity (0.16 lux), warm air and surface temperature (30.1 □), with 47dB (internal) and 68dB (external) sounds can ensure higher edible-bird nests production in the house. Also, Ibrahim et al. (2015) proposed a swiftlet house design that could keep the nestlings safe from predators such as owls, snakes, civets, bats and others, and the nests safe from burglary.

Thirdly, food resources are vital for the swiftlet population to persist. Some researchers suggested that food sources could be produced based on the previous preliminary studies on the diet of house-farm swiftlets (Kamarudin and Khoo, 2011). Kamarudin and Khoo (2011) proposed that the insect order Diptera can be a good source of food for the house-farm swiftlets due to its short life cycle, size, high nutritional value and versatility to be bred in a wide range of micro-environments. They cultured *Megaselia scalaris* (humpbacked fly) by mass producing them in the laboratory and feeding them to the house-farm swiftlets. Although the feed proportion and requirements are not well established, *M. scalaris* can be considered as a good source of protein as it contained 58% crude protein (Kamarudin and Khoo, 2011).

However, there are some drawbacks from the swiftlet farming. For instance, improper management of the swiftlet farm could also create noise disturbance. The noise created by the amplified recordings and the actual swiftlets' sounds can act as a form of noise pollution and directly affect the life quality of many residents living near the swiftlet farms (Connolly, 2017). Besides, swiftlet farming could cause some disease outbreak in the local community. Some residents also complained that the dried bird droppings can generate fine airborne particles that possibly act as a virus carrier, and potentially cause lung infections as well as other bacteria-related diseases (Sien et al., 2013; Connolly, 2017).

2.3 Diet composition of white-nest swiftlets

To date, the diet composition of swiftlets in Malaysia was investigated only by Harrisson (1974), Langham (1980), Hails and Amirrudin (1981), Waugh and Hails (1983), Lourie and Tompkins (2000) and Rahman et al. (2016). Among previous studies of the diet of swiftlets in Malaysia, three included white-nest swiftlets, both wild *Aerodramus fuciphagus* (Lourie and Tompkins, 2000) and house-farm *Aerodramus* sp. (Langham 1980; Rahman *et al.* 2016).

Langham's (1980) study took an approach to look at the white-nest swiftlets that colonised in an old building at Georgetown, Penang. Hymenoptera (40.8%), Ephemeroptera (26.4%), Homoptera (15.4%), Diptera (7.7%), Psocoptera (3.3%) were the majority (more than 90%) of prey items in the food boluses. Of the Hymenoptera identified, 68.2% belonged to the family Agaonidae (fig

wasps), 14.9% from the family Formicidae (flying ants), and 7.6% of family Torymidae (parasites of fig wasps or gall formers). This study also observed that the white-nest swiftlets would fly from the swiftlet houses before darkness to feed on swarming insects. Thus, the white-nest swiftlets that return to roost after dark could probably consume more nocturnal insects and adult beetles in their food boluses.

Lourie and Tompkins (2000) investigated the diet of white-nest swiftlets that inhabited the Gomantong Cave, Sabah. These swiftlets consumed a wide range of arthropod prey including Hymenoptera, Diptera, Homoptera, Coleoptera, Arachnida and a few other orders. The first five arthropod orders mentioned composed of more than 90% of the total prey items in the food boluses collected. Hymenopteran and Dipteran insects were the top two arthropod orders captured by the white-nest swiftlets, which accounted for 39.2% and 38.6%, respectively.

Rahman et al. (2016) documented the diet of white-nest swiftlets in three different oil palm plantation sites. In their study, 12 arthropod orders were found from the insect survey using the yellow pan trap and malaise trap methods. Among the insect orders surveyed, Diptera (26.53%), Hymenoptera (21.26%), Lepidoptera (15.92%), Coleoptera (9.05%) and Isoptera (1.32%) were the top five insect orders identified. The arthropod prey found in the diets of white-nest swiftlets were Diptera (57.1%), Homoptera (14.3%), Hymenoptera (14.3%), Isoptera and Hemiptera (7.14%, respectively). However, the diet profile of white-nest swiftlets in Rahman et al.'s (2016) study is questionable as they only reported one food bolus of the white-nest swiftlets. Furthermore, the prey items

found in that food bolus is less than the other two studies (Langham, 1980; Lourie and Tompkins, 2000). Lourie and Tompkins (2000) reported 49 to 1104 prey items in the ten food boluses collected, while Langham (2000) reported that each food bolus contained more than 500 prey items.

In general, Diptera (true flies), Ephemeroptera (mayflies) and Hymenoptera (mostly flying ants and fig wasps) are the arthropod orders which are found abundantly in the diet of the white-nest swiftlets. The presence of Hymenoptera and Blattodea (previously Isoptera or termites) in the white-nest swiftlets' diet was suggested to have occurred by chance due to the occurrence of swarming periods in Malaysia (Langham, 1980; Lourie and Tompkins, 2000).

Food resource partition among species that share the same habitat was reported by Lourie and Tompkins (2000). The diet compositions of the white-nest swiftlets, black-nest swiftlets, and mossy-nest swiftlets of the limestone complex of Gomantong Caves, Sabah were different. Within the same habitat, black-nest swiftlets consumed a much greater amount of Hymenoptera (88.5%) and less of Diptera (4.1%) and Coleoptera (1.5%). Besides, mossy-nest swiftlets consumed mainly Hymenoptera (46.0%), Diptera (25.8%) and Coleoptera (5.9%). For the white-nest swiftlets, both Hymenoptera and Diptera were consumed at similar proportions (around 38.0%), followed by Coleoptera (4.7%).

Landscape variations have also been shown to affect the diets of the Glossy swiftlets in Sabah (Lourie and Tompkins, 2000). In urban, Diptera was the most consumed arthropod order (70.7%), followed by Coleoptera (11.9%) and Hymenoptera (11.4%). In the countryside at a distance of 16 km from Sandakan, the abundance of dietary Diptera was 57.2%, lower than what was observed for the urban swiftlets, while Hymenoptera was 21.6%, and Coleoptera was 8.5%. In a forest habitat (Ampang Forest Reserve), Hymenoptera was the more preferred prey item (41.8%) by Glossy swiftlets, while Coleoptera and Diptera consumed was 20.8% and 18.9%, respectively. This suggested that the diet composition of the Glossy swiftlets has become adjusted towards the smaller sized Diptera in the urban area.

Besides, landscape variation could also affect the diet of other insectivores' tropical bird species. In Mansor et al.'s (2018c) study, they found out that three tropical insectivorous birds, Green Iora *Aegithina viridissima*, Pin-striped Tit-Babbler *Macronus gularis* and Chestnut-winged Babbler *Cyanoderma erythropterum*, showed different foraging patterns in different landscapes. So, this could lead to different insects there were consumed by the insectivorous tropical bird species due to the different habitat features (e.g. microclimates, vegetation density).

2.4 Identification of dietary arthropods

2.4.1 Physical examination of the prey items

For avian, the main purpose of a diet analysis is to (1) categorise the food items and (2) to tabulate the results in terms of occurrence and frequency (Rosenberg and Cooper, 1990). Previous studies on insectivorous avian diets have relied on morphological identification of the insect remnants in the food boluses, faecal samples, stomach contents. The physical examination of the undigested prey items may give information of the prey consumed by these insectivorous birds. For instance, in Mansor et al.'s (2018b) study, they identified arthropod taxa such as Coleoptera (53%), Hymenoptera (19%), Blattodea (11%), and Araneae (11%) in 15 species of the insectivorous birds, consisting of 12 babblers and three flycatcher-like species.

A shortcoming of the physical examination method is the biasness towards soft-bodied prey. Some of the prey items (such as insect larvae or eggs) are difficult to identify (King at al., 2015). Dillery (1965) found that soft-bodied insects can be fully digested in gizzards within 5 min, thus some critical information can be missed out during morphological identification.

2.4.2 Metabarcoding and Cytochrome Oxidase Subunit I (COI) region

Advances in the next-generation sequencing (NGS) technology at a relatively low cost in the recent decades have paved the way for species diversity investigations on the environmental samples, e.g. faeces, food boluses, soil etc., in which diverse species such as bacteria and arthropods are present (Meusnier et al., 2008; King et al., 2015). Such molecular approach is not invasive, and it

is useful to identify soft-bodied prey items which could not be identified physical examination (Kohn and Wayne, 1997). For instance, caterpillar or larvae that are partially or fully digested in the insectivores' diet could be identified through molecular method (King et al., 2015).

One of the research areas which has benefited from NGS technology is the diet profile analysis of insectivores (Jedlicka et al., 2013; Jedlicka et al., 2016; Crisol-Martínez et al., 2016; Mansor et al., 2018a). Meta-profiling of the arthropod prey is made possible with the polymerase-chain-reaction primers which are universal across a wide range of arthropod orders and flank a variable region that can distinguish different taxa. To fulfil the short read-length requirement of the NGS platform, the size of the DNA marker is limited to about 300-bp. This allows the degraded DNA fragments in the faecal samples to be analysed.

Among the commonly used DNA barcoding markers for insects was the mitochondrial cytochrome-c oxidase I (COI) (Folmer et al., 1994; Hebert et al., 2003). Numerous small-sized (less than 200-bp) barcoding markers have also been developed to detect the COI region of arthropods (Hajibabaei et al., 2006; Meusnier et al., 2008; Zeale et al., 2011; Hajibabaei et al., 2011). Although the DNA material in faecal samples may be highly degraded, the COI can still be amplified and used for taxonomic identification.

The term "DNA metabarcoding" was coined by Yu et al. (2012) with reference to the high throughput sequencing of the bulk mixture of diverse arthropod taxa

collected in the Malaise trap. Brandon-Mong et al. (2015) tested multiple sets of metabarcoding primers and designed the bioinformatics pipelines for 11 arthropod orders, i.e. Coleoptera, Diptera, Hymenoptera, Lepidoptera etc., collected in Malaysia. When the primer sets mlCOlintF/HCO2198 (Leray et al., 2013) and LepF1/MLepF1 (Brandon-Mong et al., 2015) were used, they generated highest detection rates for Diptera, Hymenoptera, Lepidoptera. These primers have been useful in the high throughput diet profiling analysis of insectivorous bats, i.e., such as free-tailed bat species *Chaerephon pumilus* and *Mops condylurus*, and Daubenton's bats *Myotis daubentoniid*, Natterer's bats *Myotis nattereri*, and birds, i.e., Western Bluebirds *Sialia Mexicana* and Rufous-winged Philentoma *Philentoma pyrhoptera* (Bohmann et al., 2011; Razgour et al., 2011; Jedlicka et al., 2013; Jedlicka et al., 2016; Vesterinen et al., 2013; Burgar et al., 2014; Hope et al., 2014; Piñol et al., 2015; Mansor et al., 2018a).

2.4.3 Argument between the taxonomic assignment methods

Some of the problems in the DNA metabarcoding analysis remain critical. Firstly, some primer sets showed bias towards certain arthropod orders. For example, lower detection rates towards Hymenoptera in the bulk PCR were observed by Yu et al. (2012) and Zhou et al. (2013) using LCO1490/HCO2198 primer pair (Folmer et al., 1994). Secondly, there were also arguments relating the completeness of DNA barcode reference libraries. Thirdly, is the accuracy of species identification methods. Therefore, the species identification methods are crucial to confirm the successful identification of the operational taxonomic units (OTU) obtaining from NGS.

The conventional and most widely used species assignment method has been based on the top similarity in BLAST search (Altschul et al., 1990). Some researchers suggested a threshold of 98% similarity for taxonomic assignment (Clare et al., 2009; Clare et al., 2011). Munch et al. (2008) justified the inappropriateness of performing taxonomic assignment based entirely on the top similar sequence in the GenBank database: (1) genetic variation across population and closely related species are ignored, and (2) the measure of confidence only reflects the local sequence similarity and not the significance of the species assignment. They have developed a phylogenetic-based software, Statistical Assignment Package, to provide a measure of statistical confidence (i.e., posterior probability) for the taxonomic assignment based on the GenBank database.

Wilson et al. (2011) proposed a tree-based assignment method in which the queries will be assigned "when they cluster with barcodes from their correct taxon". There were four tree-based taxonomic assignment criteria (liberal, strict, liberal and exclusive, strict and exclusive) based on the monophyly and the exclusivity of the query and several highly similar sequences provided in the search. Wilson et al. (2011) suggested that the conservative approach for a large-scale taxonomic assignment is the "strict" criterion. With this criterion, the query will be assigned to a taxon when it is nested within a clade formed by the members of the taxon, although some other members of this correct taxon can also be found elsewhere on the tree (Wilson et al., 2011). A more recent international bioinformatics workbench, the Barcode of Life Data (BOLD) Systems (Ratnasingham & Hebert, 2007), has adopted a tree-based

identification method advocated by Wilson et al. (2011) in assigning the “Best ID”. This tree-based method employs the neighbor-joining (NJ) algorithms only because the goal of DNA barcoding was species identification and not phylogenetic reconstruction. However, many query sequences still remain ambiguous (i.e., of uncertain high taxonomic rank, such as order) due to the conservative nature of the “strict” criterion (Wilson et al., 2011).

CHAPTER 3

MATERIALS AND METHODS

3.1 Sampling design and overall experimental design

Fresh faecal samples were collected from seven swiftlet house-farms located within five districts (Beruas, Gopeng, Ipoh, Pantai Remis and Sitiawan) across Perak State, during the month of October 2017. This month was chosen because this period of time is a monsoon transitional period between the Southwest Monsoon and Northeast Monsoon, where the rainfalls started to increase (Ramlan et al., 2017). Also, there was no occurrence of El Nino during this period of time. So, more insects can be expected to be found in the swiftlets' diet. The GPS coordinates of each swiftlet house were not provided due to the privacy concerns of the swiftlet house breeders.

Each of the swiftlet house-farm had a colony size estimated at 400 birds (personal observation). Ten pieces of cardboard of the size of 30 cm x 30 cm, each layered with plastic, were placed under the swiftlet nests in each swiftlet house-farm before darkness and were collected the next morning. The freshly collected faecal samples were kept dry in a sterile container and stored in a -20 ºC freezer prior to analysis.

The experimental design of this study employed three main approaches, namely, PCR-cloning, NGS and landscape characterisation. The steps are summarised in Figure 3.1. In PCR-cloning approach, faecal sample from Sitiawan was used to select the suitable primer-pair. After the primer selection process, the taxonomic assignment pipeline was developed using faecal samples from other six locations (except Sitiawan). The NGS approach was used to analyse the samples from six sampling sites (except Sitiawan). Landscape characterisation was used to link the diet profiles of house-farm swiftlets with the landscape features.

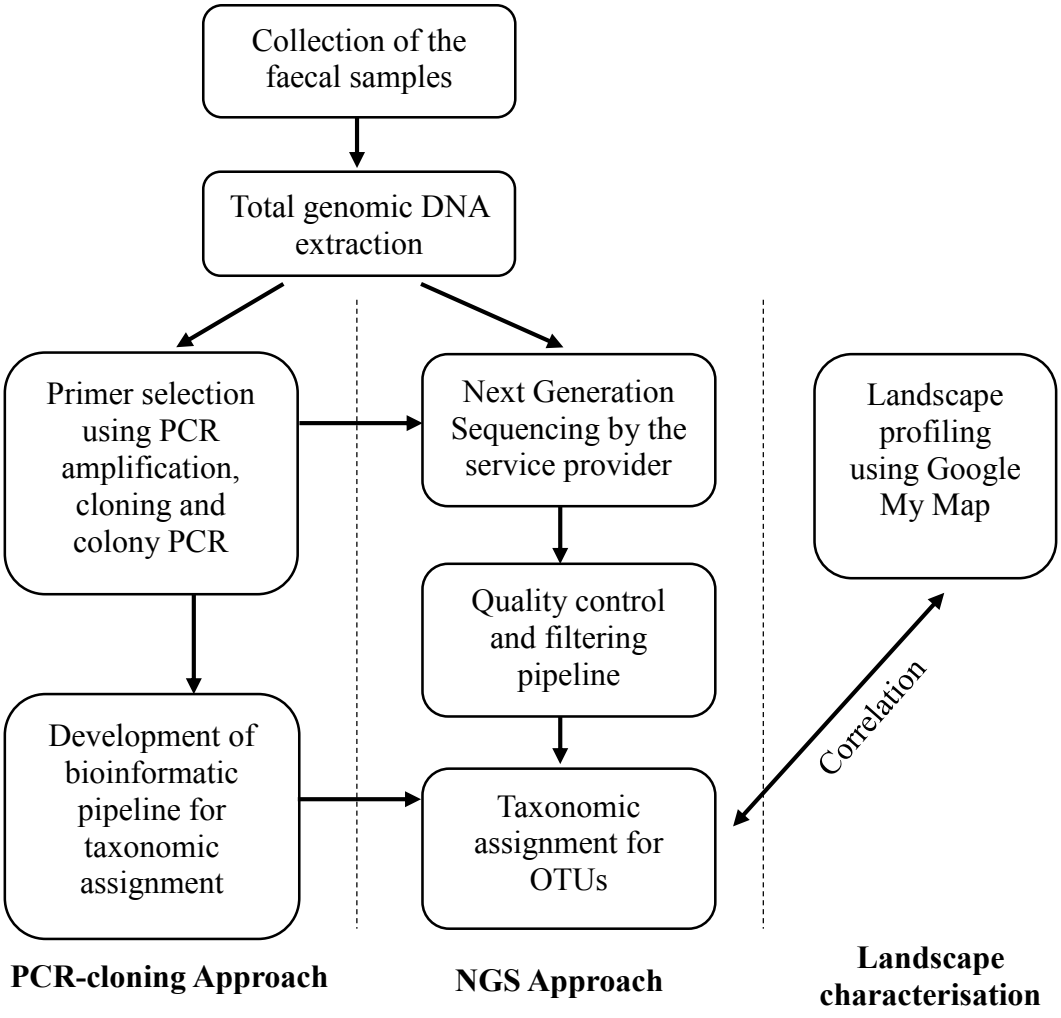


Figure 3.1 Summary of the experimental design used in this study.

3.2 Total DNA extraction

A total of 100 mg of faecal sample from each swiftlet house-farm was ground with liquid nitrogen using a prechilled pestle and mortar. Total DNA was extracted from the faecal samples using a PowerFecal DNA Kit (Qiagen, Germany) following the manufacturer's instructions. Next, total of five tubes of the extracted DNA from the same sampling site was pooled into a single tube. DNA purity and concentration were determined using a NanoDrop spectrophotometer (NanoDrop 2000c UV-Vis Spectrophotometer, Thermo Scientific).

3.3 DNA metabarcoding primer selection using PCR Amplification, cloning and colony PCR

A pilot study was conducted to select the DNA metabarcoding primer pairs. Two DNA metabarcoding primer pairs, LepF1/MLepF1_Rev and mICoIintF/HCO2198, were used to compare their sensitivity towards the insect DNA in the faecal samples. Both primers were developed for high-throughput metabarcoding analysis of insects by Brandon-Mong et al. (2015). The LepF1/MLepF1_Rev primer pair accounted for the region of ca. 218-bp (labelled as "Lep" region in the subsequent text). On the other hand, the mICoIintF/HCO2198 primer pair binds to a region of ca. 313-bp (indicated as "mlCO" region thereafter) (Leray et al., 2013; Brandon-Mong et al., 2015).

Each PCR tube contained around 60 ng of DNA samples, 1X of GoTaq® Green Master Mix (Promega, USA) and 0.5 μM of forward and reverse primers each.

The thermocycling conditions for Lep region were: an initial denaturation of 95 ºC for 2 min; 50 cycles of 95 ºC for 45 s, 53 ºC for 1 min and 72 ºC for 3 min; and a final extension of 72 ºC for 10 min. For the mlCO region, thermocycling conditions were: an initial denaturation of 95 ºC for 2 min; 39 cycles of 95 ºC for 30 s, 61□C for 45 s and 72 ºC for 1 min; and a final extension of 72 ºC for 2 min.

PCR products were electrophoresed and checked on a 1% agarose gel (First Base, Singapore) stained with ethidium bromide (Abcam, United Kingdom). The PCR products then purified using the Nucleospin Gel and PCR Clean-Up Kit (Macherey-Nagel, Germany) following the instructions by the manufacturer. The purified samples were subsequently quantified using the NanoDrop spectrophotometer before being used for ligation. Purified PCR products were ligated into pGEM®-T Easy Vector and transformed into JM109 competent cells following the instructions of the pGEM®-T Easy Vector Systems kit (Promega, USA).

Colony PCR were performed on 18 white colonies for each region. The successfully amplified colony PCR products were purified using the Nucleospin Gel and PCR Clean Up Kit following the manufacturer's instructions before they were sequenced by a local sequencing company, Apical Scientific Sdn. Bhd. Upon selection on the ideal DNA metabarcoding primer (LepF1/MLepF1_Rev), PCR cloning and colony PCR were performed on the 64 clones (10 from each sampling site) obtained from the other six swiftlet houses (3 from Beruas, 1 from Gopeng, 1 from Ipoh, and 1 from Pantai Remis).

3.4 Development of bioinformatic pipeline for taxonomic assignment

For the sensitivity comparison between the two primer-pairs towards arthropod DNA, 18 clones for each of the COI region were sequenced. The order of each haplotype was provisionally assigned based on the top hit in the searches against the GenBank database using the Basic Local Alignment Search Tool for nucleotides (BLASTn).

For the DNA sequences obtained from the 64 "Lep" clones, unique haplotypes were identified using DNaSP version 5.0 (Librado and Rozas, 2009). Each haplotype was sent for identification in the Barcode of Life Data (BOLD) database using two request types: (1) All Barcode Records on BOLD that may include private (non-accessible data), and the (2) Public Record Barcode Database which includes only the published data. Each haplotype was also searched in BLASTn and the top hits were recorded.

An NJ analysis was performed to show the genetic relatedness between each haplotype and the respective first top hits in BLASTn. The DNA sequences were then aligned using ClustalX version 2.1 (Larkin et al., 2007) and manually trimmed in Bioedit version 7.0.5 (Hall, 1999). NJ tree reconstruction with bootstrapping at 1,000 replicates was performed using Molecular Evolutionary Genetics Analysis (MEGA) version 7.0 (Kumar et al., 2016). Psocoptera was set as the outgroup as it is relatively more distant to all other arthropod orders identified here according to past phylogenomic studies (Kjer et al., 2016; Misof et al., 2014). All sequences obtained in this study were deposited in GenBank (The GenBank accession numbers to follow in the process of submission).

3.5 Next Generation Sequencing

High-throughput sequencing was commercially performed by MyTACG Bioscience Sdn. Bhd., using all Illumina Miseq Sequencer (2 x 250-bp paired-end read setting with 20,000 reads). All sequences related to this study were deposited in the National Center for Biotechnology Information (NCBI) (The GenBank accession numbers to follow in the process of submission).

3.6 Quality control and filtering pipeline for Operational Taxonomic Units (OTUs)

The pipelines for quality control and taxonomic assignment were as following. Firstly, the OTUs were filtered off if: (1) The identity of the OTU does not belong to the phylum Arthropoda and; (2) OTUs that are less than 200-bp in length.

The identity of each OTU was first searched using the Public Record Barcode Database (Ratnasingham and Hebert, 2007). If the identity of the OTU showed "no match" under the "BestID" in the Public Record Barcode Database, it would be searched again using the All Barcode Records on BOLD. The OTU would be assigned to order-level, if there is any association to "BestID" in the All Barcode Records on BOLD.

In this study, only when the "BestID" of the query sequence in both Public Record Barcode Database and All Barcode Records on BOLD is identical, the OTU is confirmed to species level. Discrepancies in identification were further

analysed using a Neighbor-joining (NJ) tree. NJ tree reconstruction was performed with bootstrapping at 1,000 replicates using top hit sequences from BLASTn, with the aid of Molecular Evolutionary Genetics Analysis (MEGA) version 7.0 (Kumar et al., 2016). OTUs that have bootstrap value more than 50% were assigned to order-level or they will be labelled as “ambiguous”.

3.7 Landscape characterisation using Google My Map

Considering the little scientific information available on the flight capabilities of any swiftlet species, the home range of the house-farm swiftlets in this study was defined as the area within a 6 km radius from a swiftlet house, following the suggestion that the average flying distance of the house-farm swiftlets ranged between 2 - 6 km (Burhanuddin and Noor, 2017).

An online mapping software, Google My Maps (https://www.google.com/mymaps) was used to characterise the landscape near to each swiftlet farm. Both the satellite and basic maps of Google My Maps were combined to provide more detailed information on the water bodies found close to each swiftlet house.

The GPS coordinates of each swiftlet house was imported and the “Measure distances and areas” functions were used to estimate the 6 km radial distance from the swiftlet house GPS coordinates. Each 6 km radius point was marked using the “Add marker” function, followed by the drawing of a circular polygon using the “Draw a line” function. Landscape features within the 6 km circular polygon from the swiftlet house were then calculated and categorised into five

types which are urban area, plantation site, forest and freshwater water bodies (such as lakes and rivers) and saltwater water bodies (sea).

Criteria and colour used for categorisation for each landscape features from Satellite Map setting of Google My Map were as follow:

1. When an area composed of settlements, industrial areas or any other form of developed land, yellow colour was used to label the landscape as urbanised area.
2. When a green area composed of well-arranged trees from aerial view, green colour was used to label the landscape as plantation area.
3. When a green area composed of cloud-shaped trees from aerial view, red colour was used to label the landscape as forest area.
4. When a river or lake were observed, blue colour was used to label these landscapes as freshwater water bodies.
5. When sea water near to the coastal were observed, purple colour was used to label the landscape as saltwater water bodies.

If any of the landscape features covered more than 80% within the 6 km radius, it would be assigned to the landscape type for the swiftlet house. If none of the category covered more than 80%, then the landscape type would be assigned as "mixed-use" landscape.

CHAPTER 4

RESULTS

4.1 Primer selection between LepF1/MLepF1_Rev and mICoIintF/HCO2198 primer-pair

From the preliminary assessment on the two primer pairs, LepF1/MLepF1_Rev (hereafter Lep primer) was shown to be a better DNA metabarcoding primer than mICoIintF/HCO2198 (hereafter mlCO primer). Of the 18 clones identified from Sitiawan swiftlet farm, using the Lep primer, five arthropod orders which included Coleoptera (11.11%), Diptera (66.67%), Hemiptera (5.56%), Hymenoptera (5.56%) and Psocoptera (5.56%) were identified for 17 clones. Another one clone was identified as *Wolbachia* (5.56%), an endosymbiotic bacterium which is commonly attached on the legs on arthropods (Jeyaprakash and Hoy, 2000; Zug and Hammerstein, 2012; Ali et al., 2018).

For the mlCO primer, 11 clones were identified as Sarcoptiformes (61.11%), 6 clones were Eurotiales (33.33%) and one clone belonged to the white-nest swiftlet (5.56%). The Sarcoptiformes (mostly mites and ticks), Eurotiales (sac fungi) and Apodiformes (white-nest swiftlets) were considered as contaminants as they were not arthropod orders. No arthropod order was identified from any of the clones using the mlCO primer. The identities of the clones assigned provisionally using BLASTn are shown in Table 4.1.

Table 4.1 Provisional assignments of the 18 clones to order-level, using the LepF1/MLepF1_Rev and mICoIintF/HCO2198 primer pairs, respectively.

Order identified	**LepF1/MLepF1_Rev**	**miCOIintF/HCO2198**
Rickettsiales	1 (5.56)	0 (0.00)
Diptera	12 (66.67)	0 (0.00)
Coleoptera	2 (11.11)	0 (0.00)
Psocoptera	1 (5.56)	0 (0.00)
Hymenoptera	1 (5.56)	0 (0.00)
Hemiptera	1 (5.56)	0 (0.00)
Sarcoptiformes	0 (0.00)	11 (61.11)
Eurotiales	0 (0.00)	6 (33.33)
Apodiformes	0 (0.00)	1 (5.56)
Total	**18 (100)**	**18 (100)**

4.2 Bioinformatic pipeline developed for taxonomic assignment of the clones

4.2.1 Identification of clones using BOLD system

A total of 27 unique haplotypes were identified from the 64 clones were bound to the Lep region. Searches based on the Public Record Barcode Database concluded 77.78% (21 out of 27 haplotypes) as “no match” in their Best IDs. Only 22.22% (6 out of 27 haplotypes; Haplotypes 4, 8, 11, 12, 14 and 20) could be identified to their respective species level (Table 4.2).

On the other hand, when the searches were carried out using the All Barcode Records on BOLD, the “no match” cases were reduced to 37.04% (10 out of 27 haplotypes; Haplotypes 1, 6, 9, 10, 13, 17, 23, 24, 25 and 27). A total of 51.86% (14 out of 27 haplotypes; Haplotypes 2, 3, 5, 7, 8, 11, 12, 15, 16, 18, 19, 21, 22 and 26) were identified to their order-level and 11.11% (3 out of 27 haplotypes; Haplotypes 4, 14 and 20) was identified to species-level (Table 4.2).

Between the Best IDs identified from the two databases, three cases of discrepancies were detected. Firstly, Haplotype 8 was identified as *Deronectes platynotus* within Coleoptera in the Public Record Barcode Database but assigned as Diptera using the All Barcode Records on BOLD. Next, Haplotypes 11 and 12 were identified as *Galathea* (order Decapoda) and *Deronectes platynotus* (order Coleoptera) in Public Record Barcode Database, respectively, yet, but were displayed as Diptera in the All Barcode Records on BOLD (Table 4.2).

Table 4.2 Summary of the identification of 27 haplotypes (out of 64 clones) in BOLD and the top hits in the BLASTn.

Haplotype	Best ID in BOLD (Public Record)		Best ID in BOLD (All Record)		Top hit in BLASTn				Decision
	Order	Genus / Species	Order	Genus / Species	Order	Genus / Species	Max Score	Accession number	
1	No match	No match	No match	No match	Lepidoptera	*Urania leilus*	283	KX781989.1	Lepidoptera
2	No match	No match	Diptera	No match	Diptera	*Nephrotoma alterna*	315	MF838043.1	Diptera
3	No match	No match	Hymenoptera	No match	Hymenoptera	*Pachycondyla* sp.	289	MF673717.1	Hymenoptera
4	Coleoptera	*Carpophilus marginellus*	Coleoptera	*Carpophilus marginellus*	Coleoptera	*Carpophilus marginellus*	363	KU914959.1	*Carpophilus marginellus*
5	No match	No match	Hymenoptera	*Hypoponera*	Hymenoptera	*Hypoponera* sp.	359	KY845694.1	*Hypoponera*
6	No match	No match	No match	No match	Diptera	Unclassified Limoniidae	270	KX053827.1	Uncertain taxa
7	No match	No match	Hemiptera	No match	Hemiptera	No match	230	KR578572.1	Hemiptera
8	Coleoptera	*Deronectes platynotus*	Diptera	No match	Diptera	*Nemorimyza* sp.	309	MF641766.1	Diptera
9	No match	No match	No match	No match	Coleoptera	*Blemus discus*	281	KU919098.1	Coleoptera
10	No match	No match	No match	No match	Hemiptera	*Eysarcoris* sp.	311	KY847240.1	Hemiptera
11	Decapoda	*Galathea*	Diptera	No match	Diptera	*Tricimba* sp.	324	KR639430.1	Diptera
12	Coleoptera	*Deronectes platynotus*	Diptera	No match	Diptera	*Nemorimyza* sp.	303	MF641766.1	Diptera
13	No match	No match	No match	No match	Psocoptera	*Liposcelis entomophila*	364	HQ658137.1	Psocoptera
14	Hymenoptera	*Odontomachus simillimus*	Hymenoptera	*Odontomachus simillimus*	Hymenoptera	*Odontomachus simillimus*	357	KU504909.1	*Odontomachus simillimus*
15	No match	No match	Coleoptera	No match	Coleoptera	*Stelidota geminata*	255	KM444965.1	Coleoptera
16*	No match	No match	Coleoptera	No match	Diptera	*Spilogona* sp.	250	KR438577.1	Uncertain taxa
17	No match	No match	No match	No match	Diptera	unclassified Cecidomyiidae	259	KM626905.1	Uncertain taxa
18*	No match	No match	Coleoptera	*Epuraea luteolus*	Coleoptera	*Epuraea signata*	298	KM442541.1	*Epuraea*
19*	No match	No match	Coleoptera	*Epuraea luteolus*	Coleoptera	*Epuraea signata*	279	KM442541.1	*Epuraea*

Table 4.2 (Cont'd)

Haplotype	Best ID in BOLD (Public Record)		Best ID in BOLD (All Record)		Top hit in BLASTn				Decision
	Order	Genus / Species	Order	Genus / Species	Order	Genus / Species	Max Score	Accession number	
20	Diptera	*Chironomus circumdatus*	Diptera	*Chironomus circumdatus*	Diptera	*Chironomus circumdatus*	370	KJ530965.1	*Chironomus circumdatus*
21	No match	No match	Hemiptera	No match	Hemiptera	*Trienopa* sp.	268	KX702955.1	Hemiptera
22	No match	No match	Hymenoptera	No match	Hymenoptera	*Crematogaster* sp.	235	KC501979.1	Hymenoptera
23	No match	No match	No match	No match	Coleoptera	*Euplatypus* sp.	244	MF804642.1	Uncertain taxa
24	No match	No match	No match	No match	Hymenoptera	*Dolichoris* sp.	193	JQ256562.1	Hymenoptera
25	No match	No match	No match	No match	Diptera	*Simulium chromatinum*	265	KM497573.1	Diptera
26	No match	No match	Hymenoptera	No match	Hymenoptera	*Philidris* sp.	355	MF804755.1	Hymenoptera
27	No match	No match	No match	No match	Diptera	*Drosophila*	292	DQ471549.1	Uncertain taxa

Note: Asterisk (*) showed discrepancies between BOLD and BLASTn at order-level (Haplotype 16) and species-level (Haplotypes 18 and 19).

Upon checking with the NJ trees generated in BOLD, the identifications for Haplotypes 8 (Figure 4.1a), 11 (Figure 4.1b) and 12 (Figure 4.1c) based on the Public Record Barcode Database were erroneous. The species assignment to these queries did not follow the "strict" criterion (Wilson et al., 2011). Furthermore, private data appears to have become mixed with the Public Record Barcode Database (as shown by the NJ trees for Haplotypes 8 and 12; Figure 4.2a and 4.2 b, respectively).

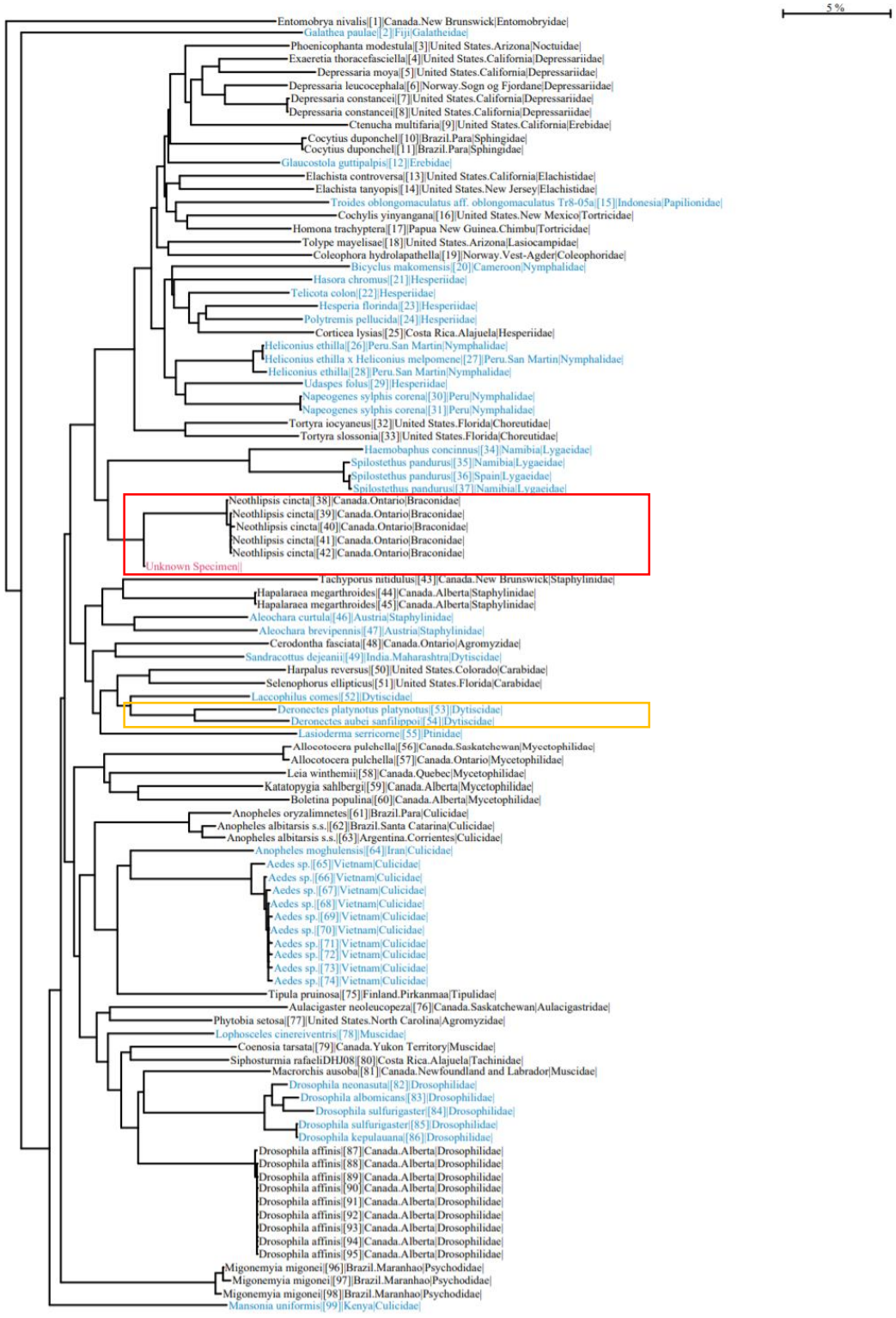


Figure 4.1a Haplotype 8 (unknown specimen) was identified as *Deronectes platynotus* (order Coleoptera) in the Public Record Barcode Database. However, haplotype 8 (in the red box) did not cluster within the same clade as *Deronectes platynotus* (in the orange box) as suggested in the Public Record Barcode Database.

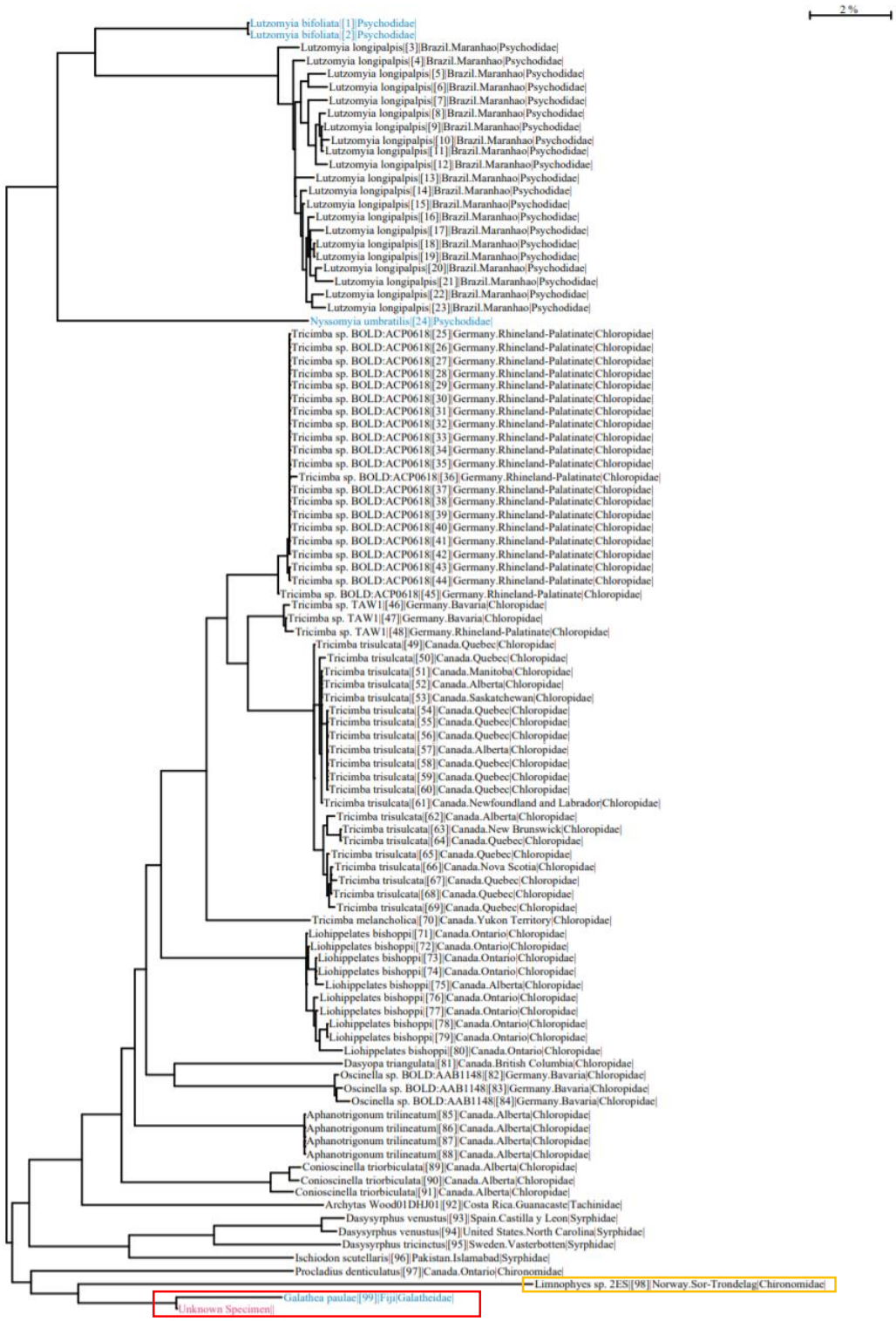


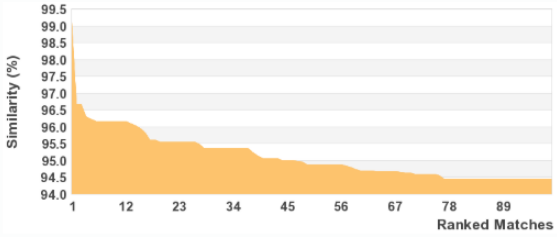
Figure 4.1b Haplotype 11 (unknown specimen) was identified as *Galathea* (order Decapoda) in the Public Record Barcode Database (in the red box). However, Haplotype 11 was clustered as *Galathea*, despite having another genus within the same clade (in the orange box), as suggested by the Public Record Barcode Database.

Figure 4.1c Haplotype 12 (unknown specimen) was identified as *Deronectes platynotus* (order Coleoptera) in the Public Record Barcode Database (in the orange box). However, Haplotype 12 did not cluster within the same clade as *Deronectes platynotus* (in the red box) as suggested by the Public Record Barcode Database.

Identification Summary:

Taxonomic Level	Taxon Assignment	Probability of Placement (%)
Phylum	Arthropoda	100
Class	Insecta	100
Order	Coleoptera	100
Family	Dytiscidae	100
Genus	Deronectes	100
Species	Deronectes platynotus	99.1

Similarity Scores of Top 99 Matches:



TOP 20 Matches :

Display option: Top 20

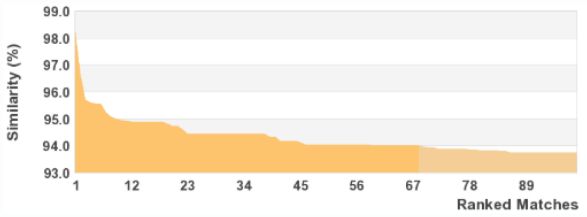
Phylum	Class	Order	Family	Genus	Species	Subspecies	Similarity (%)	Status
Arthropoda	Insecta	Coleoptera	Dytiscidae	*Deronectes*	*platynotus*	*platynotus*	99.12	Private
Arthropoda	Insecta	Lepidoptera	Lasiocampidae	*Tolype*	*mayelisae*		96.67	Published
Arthropoda	Insecta	Lepidoptera	Noctuidae	*Phoenicophanta*	*modestula*		96.67	Published
Arthropoda	Insecta	Lepidoptera	Hesperiidae	*Hesperia*	*florinda*		96.3	Published
Arthropoda	Insecta	Diptera	Mycetophilidae	*Leia*	*winthemii*		96.23	Published
Arthropoda	Insecta	Lepidoptera	Tortricidae	*Cochylis*	*yinyangana*		96.15	Published
Arthropoda	Insecta	Diptera	Agromyzidae	*Cerodontha*	*fasciata*		96.15	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		96.15	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		96.15	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		96.15	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		96.15	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		96.15	Published
Arthropoda	Insecta	Diptera	Culicidae	*Anopheles*	*moghulensis*		96.1	Published
Arthropoda	Insecta	Coleoptera	Carabidae	*Harpalus*	*reversus*		96.03	Published
Arthropoda	Insecta	Diptera	Culicidae	*Anopheles*	*albitarsis s.s.*		95.96	Published
Arthropoda	Insecta	Diptera	Culicidae	*Anopheles*	*oryzalimnetes*		95.83	Published
Arthropoda	Insecta	Coleoptera	Dytiscidae	*Laccophilus*	*comes*		95.61	Published
Arthropoda	Insecta	Diptera	Muscidae	*Coenosia*	*tarsata*		95.61	Published
Arthropoda	Insecta	Hemiptera	Lygaeidae	*Spilostethus*	*pandurus*		95.56	Published
Arthropoda	Insecta	Hemiptera	Lygaeidae	*Spilostethus*	*pandurus*		95.56	Published

Figure 4.2a Haplotype 8 was shown to be a private data although the data contained in the Public Record Barcode Database should be published.

Identification Summary:

Taxonomic Level	Taxon Assignment	Probability of Placement (%)
Phylum	Arthropoda	100
Class	Insecta	100
Order	Coleoptera	100
Family	Dytiscidae	100
Genus	Deronectes	98.2

Similarity Scores of Top 99 Matches:



TOP 20 Matches :

Display option: Top 20 ▼

Phylum	Class	Order	Family	Genus	Species	Subspecies	Similarity (%)	Status
Arthropoda	Insecta	Coleoptera	Dytiscidae	*Deronectes*	*platynotus*	*platynotus*	98.25	Private
Arthropoda	Insecta	Diptera	Culicidae	*Anopheles*	*albitarsis s.s.*		96.67	Published
Arthropoda	Insecta	Diptera	Tachinidae	*Siphosturmia*	*rafaeliDHJ08*		95.7	Published
Arthropoda	Insecta	Diptera	Mycetophilidae	*Leia*	*winthemii*		95.6	Published
Arthropoda	Insecta	Lepidoptera	Lasiocampidae	*Tolype*	*mayelisae*		95.56	Published
Arthropoda	Insecta	Lepidoptera	Noctuidae	*Phoenicophanta*	*modestula*		95.56	Published
Arthropoda	Insecta	Coleoptera	Carabidae	*Harpalus*	*reversus*		95.24	Published
Arthropoda	Insecta	Lepidoptera	Hesperiidae	*Hesperia*	*florinda*		95.06	Published
Arthropoda	Insecta	Diptera	Culicidae	*Anopheles*	*oryzalimnetes*		95	Published
Arthropoda	Insecta	Diptera	Culicidae	*Anopheles*	*albitarsis s.s.*		94.95	Published
Arthropoda	Insecta	Diptera	Aulacigastridae	*Aulacigaster*	*neoleucopeza*		94.9	Published
Arthropoda	Insecta	Lepidoptera	Tortricidae	*Cochylis*	*yinyangana*		94.87	Published
Arthropoda	Insecta	Diptera	Agromyzidae	*Cerodontha*	*fasciata*		94.87	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		94.87	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		94.87	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		94.87	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		94.87	Published
Arthropoda	Insecta	Hymenoptera	Braconidae	*Neothlipsis*	*cincta*		94.87	Published
Arthropoda	Insecta	Diptera	Culicidae	*Anopheles*	*moghulensis*		94.81	Published
Arthropoda	Insecta	Coleoptera	Dytiscidae	*Laccophilus*	*comes*		94.74	Published

Figure 4.2b Haplotype 12 was shown to be a private data although the data contained in the Public Record Barcode Database should be published.

4.2.2 BLASTn search and Neighbor-Joining (NJ) Analysis

The first top hits for each haplotype are listed in Table 4.2. As shown in the NJ tree constructed using top hits BLASTn sequences, 18 haplotypes (1, 4, 5, 7, 9, 10, 11, 13, 14, 15, 18, 19, 20, 21, 22, 24, 25 and 26) formed clusters with their respective BLASTn top hits (Figure 4.3).

Nine haplotypes (2, 3, 6, 8, 12, 16, 17, 23 and 27) were not resolved in the NJ tree even though their top hits were included in the NJ analysis. It is noted in Figure 4.3 that Haplotypes 16 and 17 formed a cluster but their top hits are sequences, from the families of Muscidae and Cecidomyiidae, respectively, which are not within the cluster. Haplotype 23 was clustered with members of Hemiptera (87%) but its first top hit (max score = 244) in the BLASTn search was *Euplatypus* of Coleoptera (Figure 4.3).

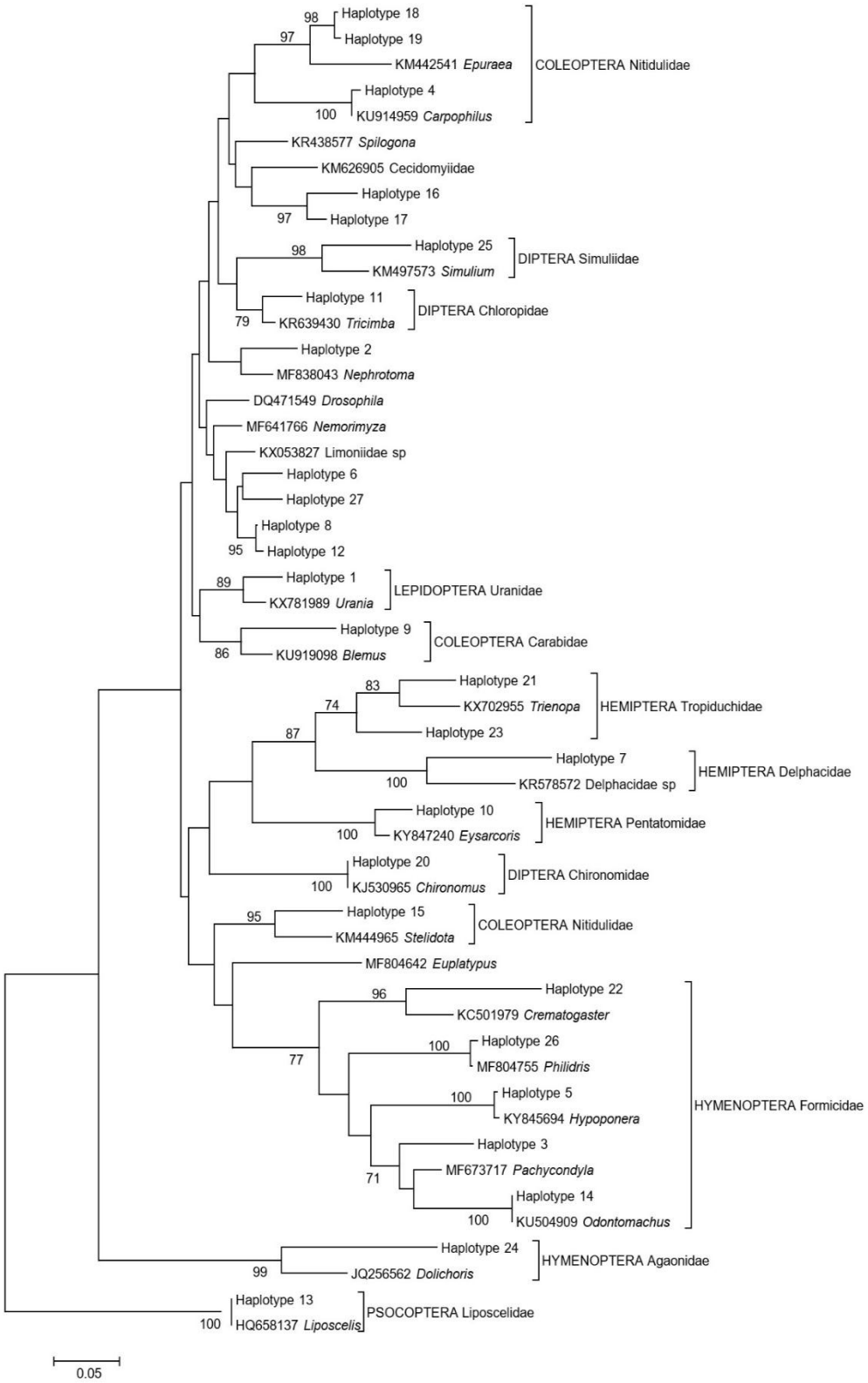


Figure 4.3 NJ tree reconstruction based on the "Lep" region. Bootstrap value more than 70% are shown above the nodes.

4.2.3 Decision for taxonomic assignment

The order/species identities suggested by the BLASTn top hits were tallied with the BOLD identification for 14 haplotypes (out of 17 haplotypes identified in the All Barcode Records on BOLD). Taxonomic assignment for these haplotypes therefore follows the BOLD identification based on the All Barcode Records on BOLD.

Three cases where the identifications using BOLD and BLASTn did not tally, i.e., Haplotype 16 (order-level) and Haplotypes 18 and 19 (species-level), were carefully investigated. Considering that the genetic affinity between Haplotype 16 and its BLASTn top hit was not supported in NJ tree, and that the private reference data used in BOLD identification was not available to the users, Haplotype 16 was therefore concluded as uncertain order. It was suggested that the genus assignment which was agreed upon by both identification systems was used for Haplotypes 18 and 19.

Four “no match” haplotypes (6, 17, 23 and 27) in the All Barcode Records on BOLD were not resolved in the NJ tree as well. Therefore, their identities remained uncertain. For the haplotypes which were “no match” in the All Barcode Records on BOLD but appeared to be related to their respective top hits in the NJ analysis, their orders were assigned based on the BLASTn top hits. These cases are Haplotypes 1, 9, 10,13, 24 and 25.

In summary, among the 27 haplotypes extracted from the house-farm swiftlet faecal samples, a total of 3.70% (1 haplotype) of Lepidoptera, 22.22% (6

haplotypes) of Hymenoptera, 11.11% (3 haplotypes) of Hemiptera, 22.22% (6 haplotypes) of Diptera, 18.52% (5 haplotypes) of Coleoptera and 3.70% (1 haplotype) of Psocoptera were detected. A total of six haplotypes were confirmed down to species level. Among Diptera, one haplotype was *Chironomus circumdatus*. Two haplotypes among Coleoptera were *Epuraea* and *Carpophilus marginellus*. Among Hymenoptera, one haplotype was identified as *Hypoponera* while another was *Odontomachus simillimus*.

4.2.4 Pipeline developed for the 64 identified clones from six sampling sites

The bioinformatic pipeline developed based on the cloning results in this study could be summarised into four steps for taxonomic assignment of arthropod COI region involve four steps:

(1) The Best ID, if provided in BOLD searches using the Public Record Barcode Database, would be considered as the taxon identity. If "no match" (i.e., no Best ID) in this setting,

(2) The Best ID, if provided in BOLD searches using the All Barcode Records on BOLD, would be considered as the taxon identity. If "no match" (i.e., no Best ID) in this setting,

(3) The order of the first top hit in BLASTn, if it formed a cluster with the query sequence, would be considered as the taxon identity. If there is no clustering between the query sequence and the first top hit,

(4) The query would be considered as an uncertain taxon.

4.3 Operational Taxonomic Units (OTUs) obtained from NGS

Out of the 4,852 OTUs generated by the NGS service provider, 93.22% of the non-arthropod sequences (4,523 OTUs) were removed, based on the searches using the All Barcode Records on BOLD. Another 1.30% (63 OTUs) of the sequences were also discarded as they were shorter than 200-bp sequence length (Figure 4.4). Taxonomic assignment was performed on the remaining 266 OTUs using the pipeline described in section 4.2.4. Of the 266 OTUs, 79.70% (212 OTUs) were resolved either to order or species level, whereas 20.30% (54 OTUs) were considered as unresolved taxa.

Of the 212 OTUs, 21 OTUs were identified down to the species level based on matching identities between the All Barcode Records on BOLD and Public Record Barcode Database. Another 188 OTUs that showed no matched "BestID" using the Public Record Barcode Database, were reassigned to order-level using the All Barcode Records on BOLD (Figure 4.4, Figure 4.5 and Appendix A).

Of the 57 OTUs that remained as uncertain taxa, three OTUs were clustered with their respective BLASTn top hit sequence. OTU 138 showed clustering

with family Muscidae of Diptera at a bootstrap value of 56%. Besides, OTU 1977 and OTU 2558 showed genetic affinity to family Dolichopidae and Ephydridae of Diptera at bootstrap values of 77% and 99%, respectively. Thus, these three OTUs were reassigned as Diptera and made up a total of 191 OTUs that were identified to order-level (Figure 4.4).

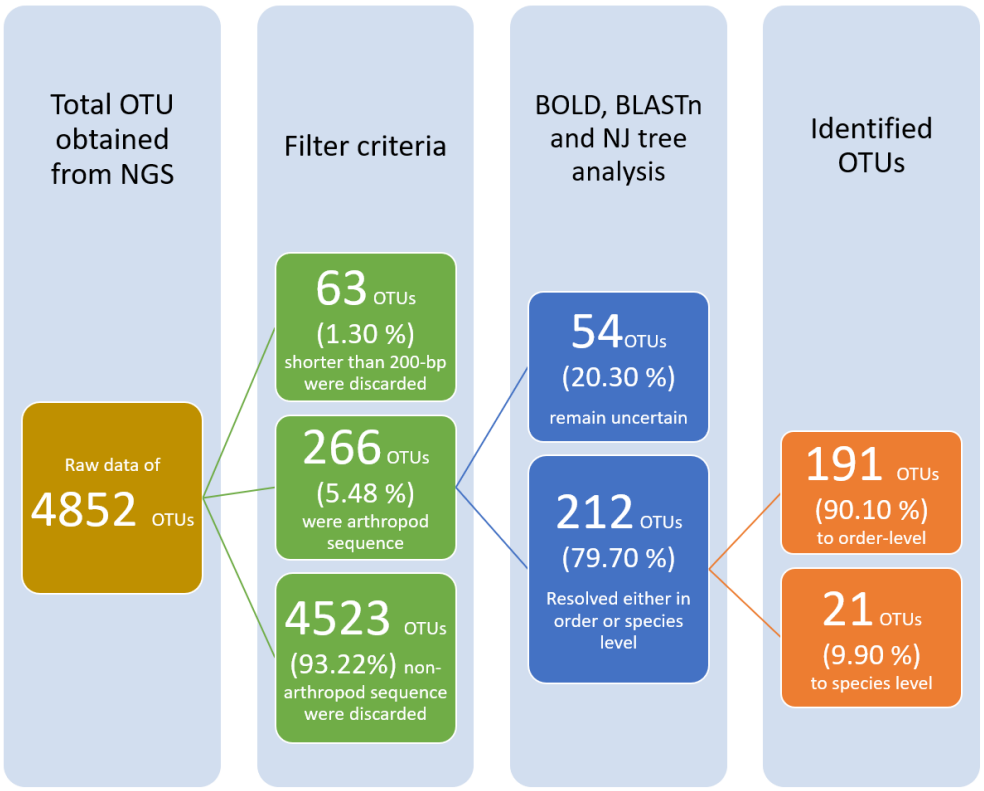


Figure 4.4 Conceptual illustration of the flow of taxonomic assignment for Operational Taxonomic Units (OTUs) generated using NGS. The percentage was calculated based on the connecting line to each box (e.g. the total percentage from the green boxes is equal to the sum of the brown box).

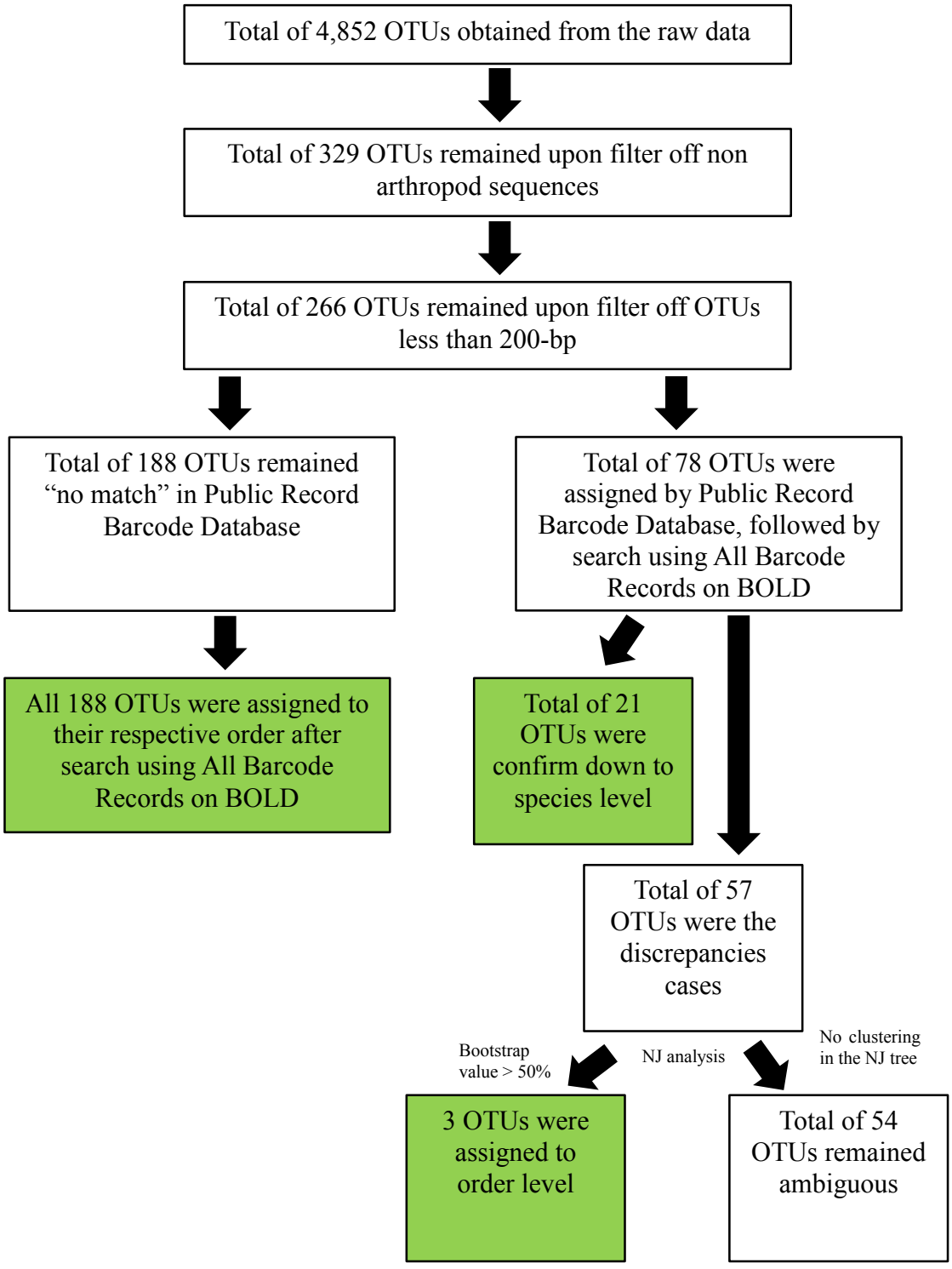


Figure 4.5 Detailed summary of the taxonomic assignment pipeline used for the 4852 OTUs obtained from NGS.

The remaining 54 OTUs recorded order-level discrepancies between the Public Record Barcode Database and All Barcode Records on BOLD searches (Table 4.3). Furthermore, these 54 OTUs showed no genetic affinity to any of the BLASTn top hit sequences as shown in the NJ tree (Figure 4.6a and Figure 4.6b), and hence remained as unidentified taxa. All the BLASTn top hits were as listed in Table 4.3.

In summary, seven arthropod orders were identified for 212 OTUs out of 266 OTUs (Figure 4.7). The arthropod orders included Diptera (62.74%), Hemiptera (18.87%), Coleoptera (12.26%), Lepidoptera (2.36%), Hymenoptera (1.89%), Blattodea (0.94%) and Odonata (0.94%).

Nine species were identified from the 21 species confirmed OTUs. Taxonomic identities were as shown in Figure 4.7 and Table 4.4. Although nine species were identified using the pipeline developed, only three species *Tetramorium bicarinatum*, *Odontotermes hainanensis* and *Ceriagrion auranticum* could be confirmed due to their geographical distribution in the Peninsular Malaysia (Table 4.4). Another five species were reassigned to a higher taxonomic rank at family-level due to the absence of species record in Peninsular Malaysia. Among the families identified were Braconidae (Hymenoptera), Noctuidae (Lepidoptera), Choreutidae (Lepidoptera), Dytiscidae (Coleoptera) and Chironomidae (Diptera). The remaining species was assigned to generic level as the sequence remained uncertain to be found in East Malaysia or Peninsular Malaysia (Table 4.4).

Table 4.3 Comparison of 54 OTUs that showed discrepancies at the order-level among searches using the All Barcode Records on BOLD and Public Record Barcode Database, together with the BLASTn top hit sequences.

No.	OTU	BOLD Public Record	BOLD All record	BLASTn			
		Order	Order	Order	Family	Genus/Species	Accession number
1	OTU 14	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
2	OTU 34	Coleoptera	Diptera	Diptera	Muscidae	*Drymeia* sp.	MF890509.1
3	OTU 46	Coleoptera	Diptera	Coleoptera	Carabidae	*Calleida viridis*	JX259813.1
4	OTU 140	Lepidoptera	Diptera	Diptera	Muscidae	*Drymeia* sp.	MF890509.1
5	OTU 154	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
6	OTU 228	Coleoptera	Diptera	Diptera	Mycetophilidae	*Leia winthemii*	MF889420.1
7	OTU 248	Coleoptera	Diptera	Coleoptera	Carabidae	*Calleida viridis*	JX259813.1
8	OTU 336	Coleoptera	Diptera	Diptera	Muscidae	*Drymeia* sp.	MF890509.1
9	OTU 427	Coleoptera	Diptera	Coleoptera	Carabidae	*Calleida viridis*	JX259813.1
10	OTU 493	Coleoptera	Diptera	Diptera	Limoniidae	*Limoniidae* sp.	KX053827.1
11	OTU 613	Coleoptera	Diptera	Coleoptera	Carabidae	*Calleida viridis*	JX259813.1
12	OTU 618	Coleoptera	Diptera	Diptera	Lauxaniidae	*Lauxaniidae* sp.	KR396249.1
13	OTU 700	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
14	OTU 779	Coleoptera	Diptera	Diptera	Muscidae	*Helina* sp.	MF884850.1
15	OTU 780	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
16	OTU 786	Coleoptera	Hemiptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
17	OTU 815	Coleoptera	Lepidoptera	Diptera	Tachinidae	*Leucostoma simplex*	KX843880.1
18	OTU 951	Coleoptera	Diptera	Diptera	Drosophilidae	*Leucophenga zhenfangae*	JX235940.1
19	OTU 972	Coleoptera	Diptera	Coleoptera	Carabidae	*Calleida viridis*	JX259813.1
20	OTU 1192	Coleoptera	Diptera	Diptera	Phoridae	*Megaselia* sp.	MF871071.1

Table 4.3 (Cont'd)

No.	OTU	BOLD Public Record	BOLD All record	BLASTn			
		Order	Order	Order	Family	Genus/Species	Accession number
21	OTU 1357	Coleoptera	Diptera	Diptera	Phoridae	*Megaselia* sp.	MF871071.1
22	OTU 1389	Plecoptera	Lepidoptera	Coleoptera			JQ344783.1
23	OTU 1558	Coleoptera	Diptera	Diptera	Psychodidae	*Migonemyia migonei*	KP112713.1
24	OTU 1569	Coleoptera	Lepidoptera	Diptera	Muscidae	*Drymeia* sp.	MF890509.1
25	OTU 1643	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
26	OTU 1742	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
27	OTU 2021	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
28	OTU 2030	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
29	OTU 2126	Coleoptera	Diptera	Diptera	Phoridae	*Megaselia* sp.	MF871071.1
30	OTU 2179	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
31	OTU 2384	Coleoptera	Diptera	Diptera	Tabanidae	*Plinthina binotata*	KC592610.1
32	OTU 2714	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
33	OTU 2748	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
34	OTU 2837	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
35	OTU 2877	Coleoptera	Diptera	Diptera	Lygistorrhinidae	*Asiorrhina parasiatica*	KT316832.1
36	OTU 2925	Coleoptera	Diptera	Diptera	Muscidae	*Drymeia* sp.	MF890509.1
37	OTU 3050	Coleoptera	Diptera	Lepidoptera	Gracillariidae	*Phyllonorycter sorbi*	KX045792.1
38	OTU 3070	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
39	OTU 3238	Coleoptera	Lepidoptera	Coleoptera	Carabidae	*Calleida viridis*	JX259813.1
40	OTU 3281	Coleoptera	Diptera	Coleoptera	Carabidae	*Calleida viridis*	JX259813.1
41	OTU 3384	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1

42	OTU 3406	Coleoptera	Diptera	Diptera	Muscidae	*Polietina prima*	AJ879601.1

Table 4.3 (Cont'd)

No.	OTU	BOLD Public Record	BOLD All record	BLASTn			
		Order	**Order**	**Order**	**Family**	**Genus/Species**	**Accession number**
43	OTU 3600	Coleoptera	Diptera	Diptera	Mycetophilidae	*Leia winthemii*	MF884936.1
44	OTU 3693	Coleoptera	Diptera	Diptera	Tabanidae	*Plinthina binotata*	KC592610.1
45	OTU 3782	no match	no match	Coleoptera	Staphylinidae	*Philonthus discoideus*	KU909625.1
46	OTU 3879	Coleoptera	Diptera	Diptera	Muscidae	*Drymeia* sp.	MF890509.1
47	OTU 3907	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
48	OTU 4074	Lepidoptera	Lepidoptera	Diptera	Tabanidae	*Myioscaptia calliphora*	KC592589.1
49	OTU 4137	Coleoptera	Diptera	Diptera	Mycetophilidae	*Leia winthemii*	MF889420.1
50	OTU 4466	Coleoptera	Diptera	Diptera	Muscidae	*Drymeia* sp.	MF890509.1
51	OTU 4518	Diptera	Lepidoptera	Diptera	Muscidae	*Drymeia* sp.	MF890509.1
52	OTU 4592	Coleoptera	Diptera	Diptera	Drosophilidae	*Drosophila fraburu*	EU493602.1
53	OTU 4744	Coleoptera	Diptera	Diptera	Phoridae	*Megaselia* sp.	MF871071.1
54	OTU 4840	Coleoptera	Diptera	Diptera	Mycetophilidae	*Leia winthemii*	KY833113.1

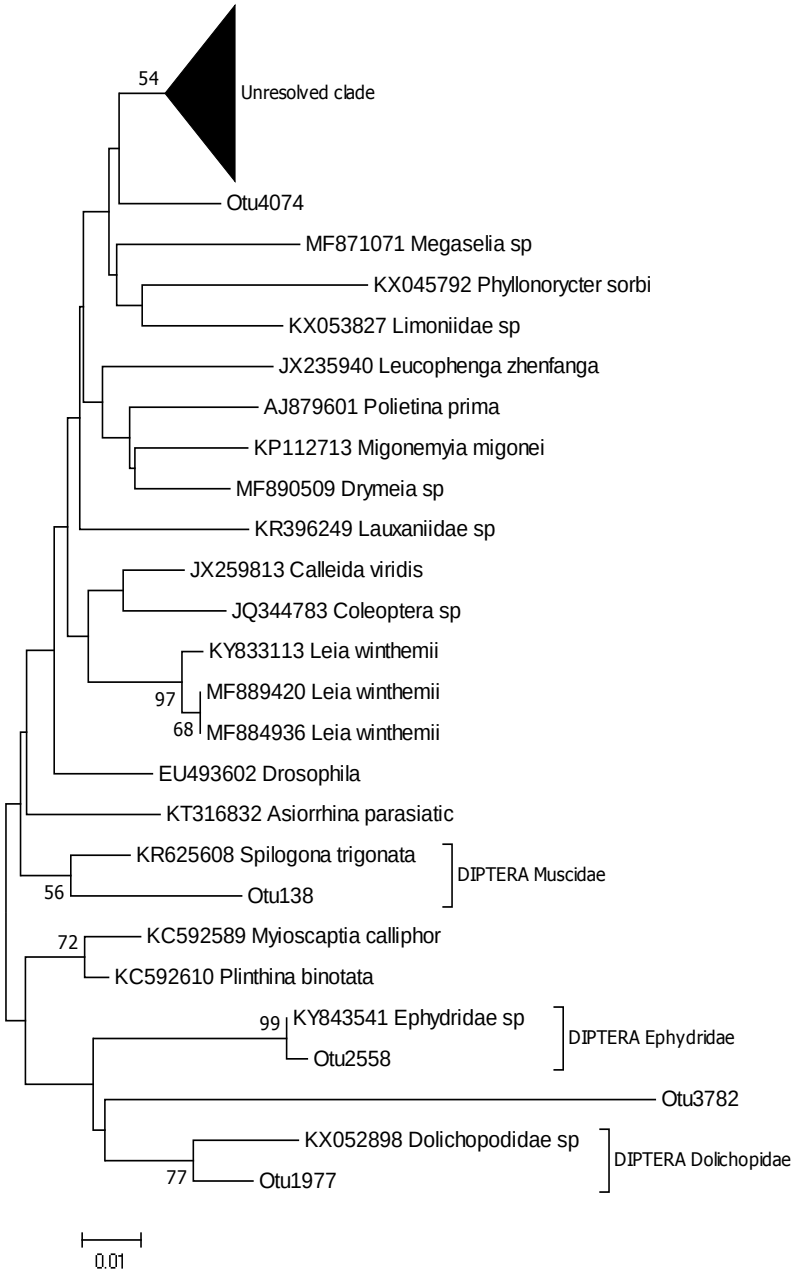


Figure 4.6a Overall NJ tree reconstructed based on sequence from BLASTn (bootstrap more than 50%).

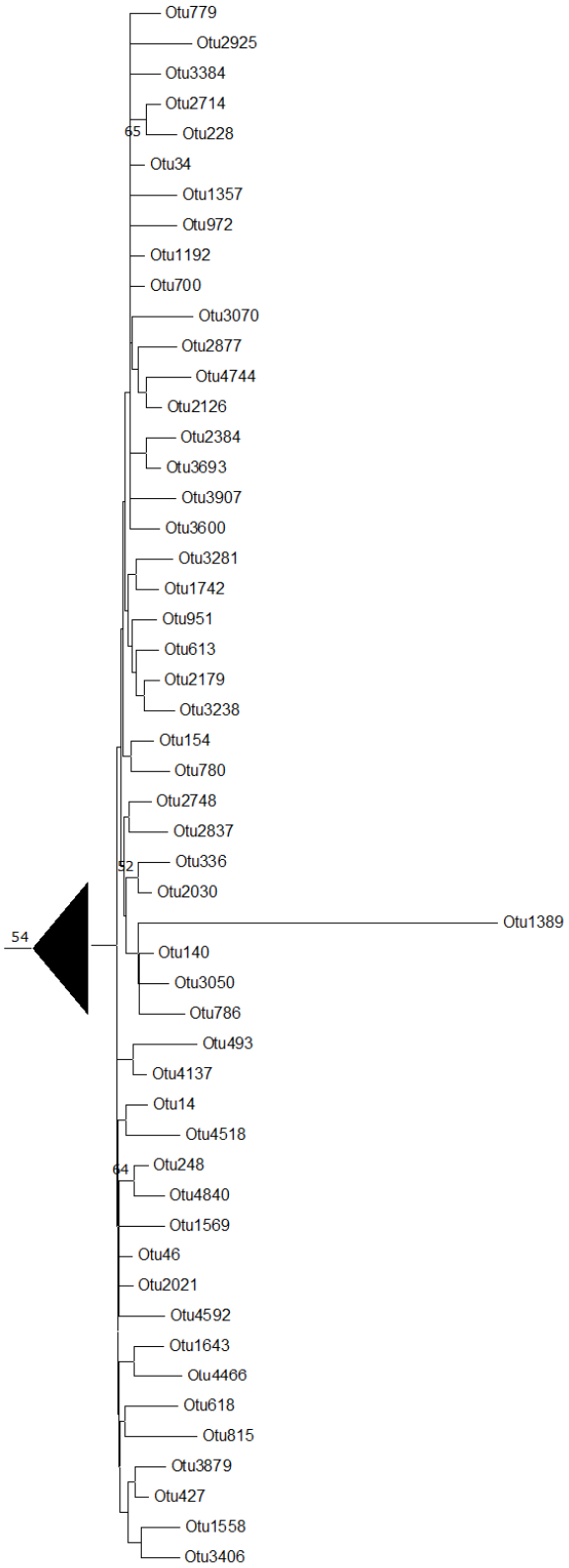


Figure 4.6b Inlet of the NJ tree reconstructed based on sequence from BLASTn (bootstrap more than 50%).

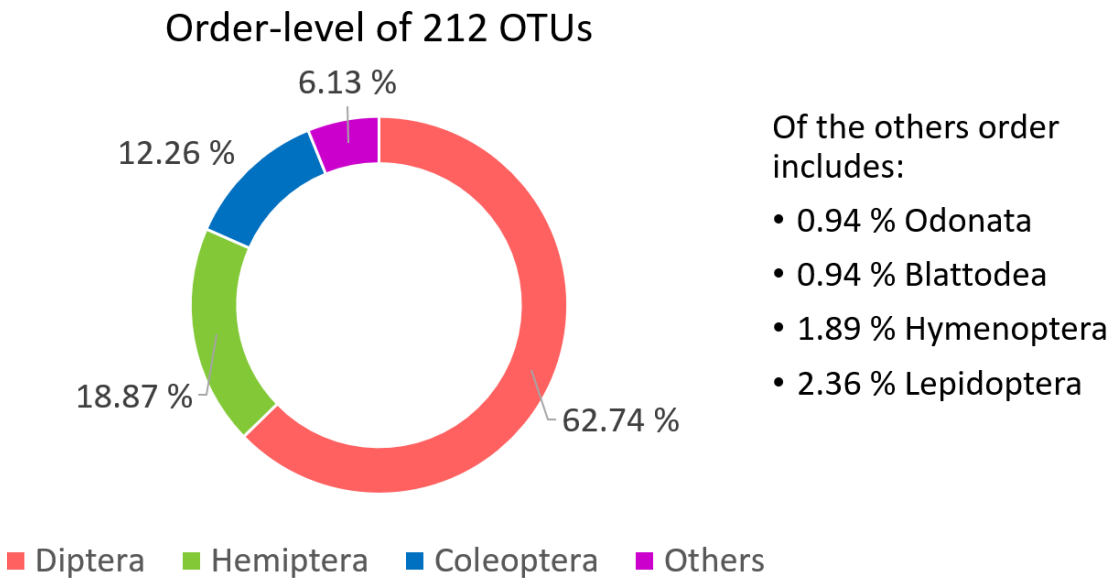


Figure 4.7 Summary of arthropod orders assigned using BOLD (Public Record Barcode Database and All Barcode Records on BOLD) and BLASTn/Neighbor-joining analysis. A total of 212 out of 266 OTUs were assigned to their respective order level. Within the 212 OTUs identified orders, 21 OTUs were confirmed to species-level. *Kiefferulus tainanus* (2) *Spelobia bifrons* (1) *Deronectes platynotus* (10) *Phoenicophanta modestula* (1) *Tortyra iocyaneus* (1) *Tetramorium bicarinatum* (1) *Neothlipsis cincta* (1) *Odontotermes hainanensis* (2) *Ceriagrion auranticum* (2).

Table 4.4 Nine species that were identified from the 21 OTUs assigned to species-level using the developed pipeline. However, they were reassigned to either family-level, generic-level or species-level based on geographical distribution. If the species was not found in Peninsular Malaysia, it was assigned to familial-level; If the species is uncertain to be found in Peninsular or East Malaysia, it would be assigned to generic-level. If the species could be found in the Peninsular Malaysia, the species identity could be confirmed.

Arthropod Order identified	**Species Included (frequency)**	**Geographical distribution record**	**Reference(s)**	**Decision**
Diptera	*Kiefferulus tainanus* (2)	China, Japan and Thailand. (No record found in Peninsular Malaysia)	http://v3.boldsystems.org/index.php/Taxbrowser_Taxonpage?taxon=Kiefferulus+tainanus+&searchTax=	Assigned to family-level Chironomidae
	Spelobia bifrons (1)	Bangladesh, Egypt, United States, Bulgaria, Malaysia, Norway, Canada, Australia, Saudi Arabia, New Zealand, Germany and South Africa. For Malaysia (uncertain in East or Peninsular Malaysia).	http://v3.boldsystems.org/index.php/Taxbrowser_Taxonpage?taxon=Spelobia+bifrons+&searchTax=	Assigned to generic-level *Spelobia*
Coleoptera	*Deronectes platynotus* (10)	Germany. (No record found in Peninsular Malaysia)	http://v3.boldsystems.org/index.php/Taxbrowser_Taxonpage?taxon=Deronectes+platynotus+&searchTax=	Assigned to family-level Dytiscidae

Table 4.4 (Cont'd)

Arthropod Order identified	Species Included (frequency)	Geographical distribution record	Reference(s)	Decision
Lepidoptera	*Phoenicophanta modestula* (1)	Unites States and Mexico. (No record found in Peninsular Malaysia)	http://v3.boldsystems.org/index.php/Taxbrowser_Taxonpage?taxon=Phoenicophanta+modestula+&searchTax=	Assigned to family-level Noctuidae
	Tortyra iocyaneus (1)	Unites States. (No record found in Peninsular Malaysia)	http://v3.boldsystems.org/index.php/Taxbrowser_Taxonpage?taxon=Tortyra+iocyaneus&searchTax=	Assigned to family-level Choreutidae
Hymenoptera	*Tetramorium bicarinatum* (1)	Malaysia (both East and Peninsular), Indonesia, Myanmar, Cambodia, Philippines, Vietnam, Taiwan etc.	Yahya and Hamdan (2014) http://antmaps.org/?mode=species&species=Tetramorium.bicarinatum	Assigned to species-level
	Neothlipsis cincta (1)	Unites states and Canada. (No record found in Peninsular Malaysia)	http://v3.boldsystems.org/index.php/Taxbrowser_Taxonpage?taxon=Neothlipsis+cincta+&searchTax=	Assigned to family-level Braconidae

Table 4.4 (Cont'd)

Arthropod Order identified	**Species Included (frequency)**	**Geographical distribution record**	**Reference(s)**	**Decision**
Blattodea	*Odontotermes hainanensis* (2)	Peninsular Malaysia.	Cheng et al. (2011)	Assigned to species-level
Odonata	*Ceriagrion auranticum* (2)	Malaysia (both East and Peninsular), Indonesia, Laos, Myanmar, Thailand, Vietnam.	https://www.iucnredlist.org/species/164790/5927151	Assigned to species-level
Total	(21)			

4.4 Landscape characterization using Google My Map

All six sampling sites were plotted as maps using their respective GPS coordinates (Figure 4.8). The sampling sites subjected to landscape characterisation were Beruas, Beruas OP1, Beruas OP2, Gopeng, Ipoh and Pantai Remis (Figure 4.9).

The area of each landscape was calculated using formula πr^2 and each landscape has an estimated area of 113.1 km^2. In this study, the landscape type was classified as urban area, plantation, forest, river and lake if the landscape profile had more than 80% of land coverage. Based on the landscape profiles found at each sampling site, three landscape types (mixed-use, monocrop and urban) could be categorised (Table 4.5).

Firstly, a mixed-use of landscape was identified for sampling sites such as Beruas, Gopeng and Pantai Remis. For Beruas, the landscape profiles were predominated by 77.8% of plantations, and consisted of 2.0% urban area and 20.1% of forest area. For Gopeng, the landscape profiles of the swiftlet house showed by 46.1% of plantations (mostly *Aquilaria* sp. trees, as the sampling site was located at the Gaharu Tea Valley), 40.5% of forest and 12.4% of urban area. At Pantai Remis, plantations and forest covered more than 60% of the total landscape (54.0% and 9.4%, respectively), followed by 17.0% of urban area.

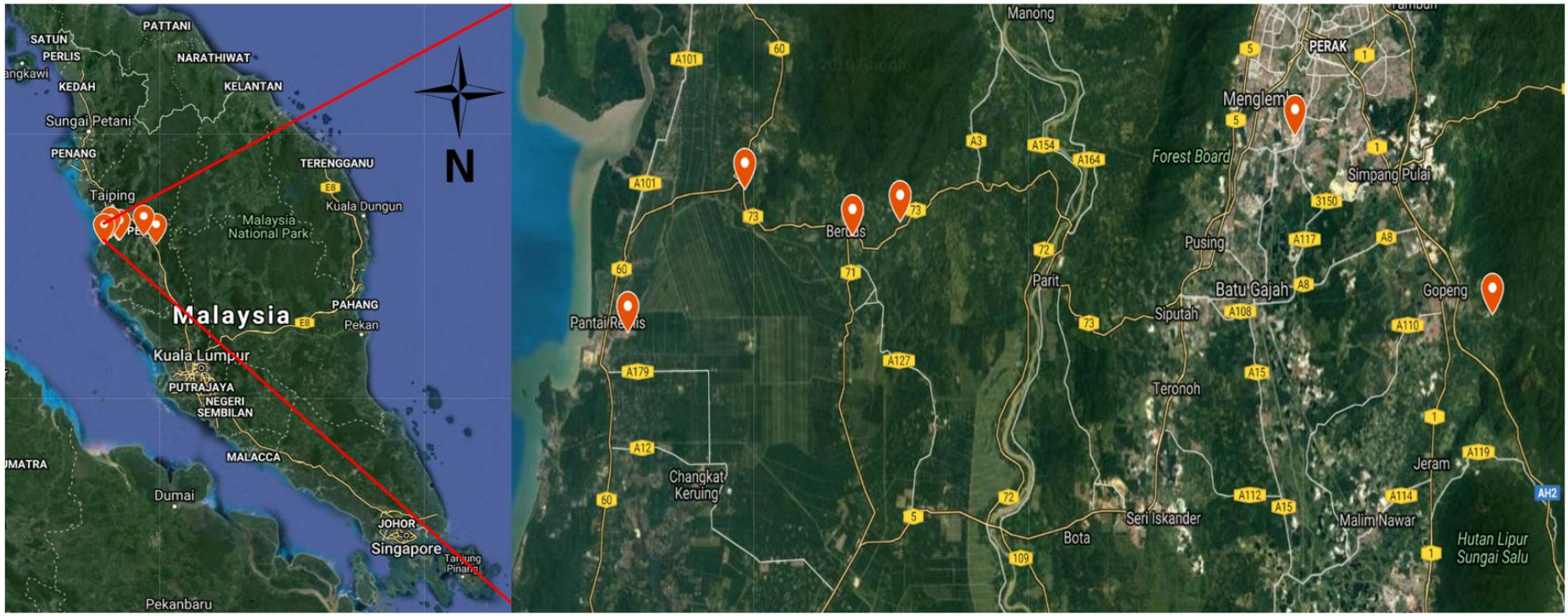


Figure 4.8 All six sampling sites were sited in the Perak state of Malaysia. Three swiftlet houses were located in Beruas while another three located in Gopeng, Ipoh and Pantai Remis, respectively.

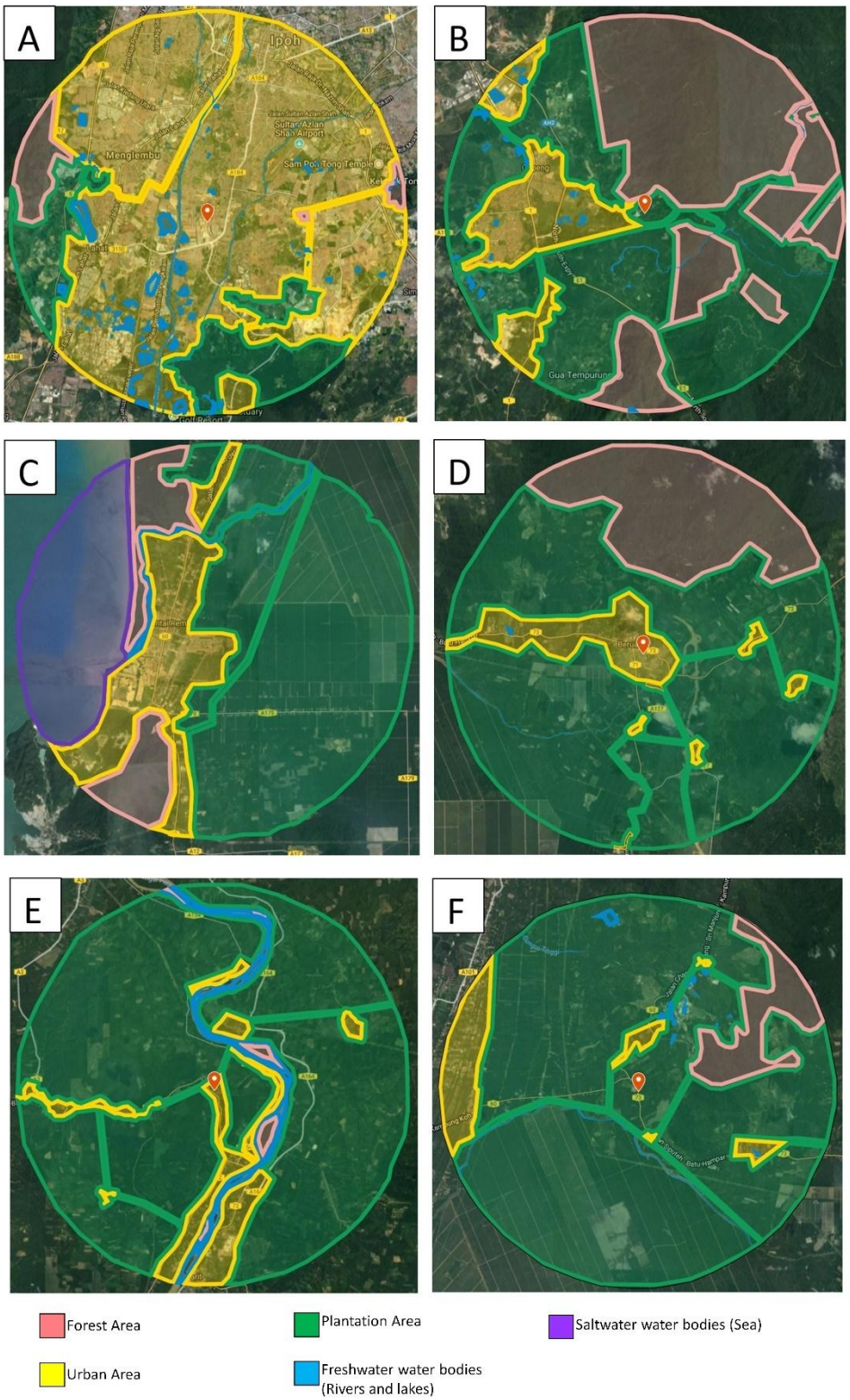


Figure 4.9 Landscape profiles within a radius of 6 km for each sampling site were estimated using Google My Map: (A) Ipoh, (B) Gopeng, (C) Pantai Remis, (D) Beruas, (E) Beruas OP1, and (F) Beruas OP2.

Table 4.5 Landscape profiles of the sampling sites based on an area within a 6 km radius of the swiftlet house, estimated using Google My Map.

	Urban Area (%)	Plantation (%)	Forest (%)	River and Lake (%)	Sea (%)	Total Area Size (%)	Landscape type
Beruas	2.3 (2.0)	88.0 (77.8)	22.7 (20.1)	0.1 (0.1)	0.0 (0.0)	113.1 (100.0)	Mixed-use
Beruas OP1	9.1 (8.0)	98.0 (86.7)	1.5 (1.4)	4.4 (3.9)	0.0 (0.0)	113.1 (100.0)	Monocrop
Beruas OP2	6.9 (6.1)	97.6 (86.3)	7.6 (6.7)	1.0 (0.9)	0.0 (0.0)	113.1 (100.0)	Monocrop
Gopeng	14.0 (12.4)	52.1 (46.1)	45.8 (40.5)	1.2 (1.1)	0.0 (0.0)	113.1 (100.0)	Mixed-use
Ipoh	91.2 (80.6)	12.7 (11.2)	4.4 (3.9)	4.8 (4.3)	0.0 (0.0)	113.1 (100.0)	Urban
Pantai Remis	19.2 (17.0)	61.1 (54.0)	10.6 (9.4)	1.0 (0.9)	21.2 (18.7)	113.1 (100.0)	Mixed-use

Monocrop landscape was seen at sampling sites Beruas OP1 and Beruas OP2. The landscape profiles of Beruas OP1 and Beruas OP2 were predominantly plantations, at 86.7% and 86.3%, respectively. Also, forest areas were sparse in the surroundings and were only observed at 1.4% for Beruas OP1 and 6.7% for Beruas OP2. Urban area for Beruas OP1 and Beruas OP2 were accounted at 8.0% and 6.1%, respectively.

The urban landscape was observed only at the Ipoh sampling site. The landscape profiles were mainly urban area (80.6%), followed by 11.2% of plantations and 3.9% of forest area.

Additionally, freshwater water bodies such as rivers and lakes and saltwater water bodies, sea were also categorised. On average, only around 1.87% of freshwater water bodies, including both rivers and lakes, could be found at each sampling site. Of the six sampling sites, Ipoh has the largest freshwater water bodies (4.3% of the total landscape profiles). Two rivers (Sungai Kinta and Sungai Pinji) and 116 lakes were observed using Google My Map (Figure 4.9-A). On the other hand, Beruas has the smallest freshwater water bodies (0.1%). Saltwater water body (sea) is only found at Pantai Remis as this sampling site is a coastal town. The sea covered 18.7% of the total landscape profiles.

4.5 Arthropod distribution in different landscape types

The distribution of OTUs across the six sampling sites, namely Beruas, Beruas OP1, Beruas OP2, Ipoh, Gopeng and Pantai Remis is summarized in Table 4.6. A total of 245 OTUs (instead of 212) was observed as some OTUs were recorded in more than one location. The arthropod abundance refers to the species richness found in the faecal sample of the swiftlet.

Ipoh recorded the largest numbers of OTUs (n=111), followed by Beruas OP2 (n=60), Gopeng (n=32), Beruas (n=15), Beruas OP1 (n=14) and Remis (n=13). Among the six sampling sites, seven arthropod orders were identified, including Diptera (64.49%) which was the most common order found in all house-farms. Both Odonata and Blattodea (0.82%) were the least common orders.

Five arthropod orders were recorded at the Pantai Remis and Ipoh sampling sites. In Ipoh, a total of 111 OTUs was assigned to Diptera (81.98%), Coleoptera (10.81%), Hymenoptera (2.70%), Lepidoptera (2.70%) and Odonata (1.80%). For Pantai Remis, Diptera was found to be most abundant (46.15%), followed by Hemiptera (23.08%), Blattodea (15.38%), Hymenoptera (7.69%) and Coleoptera (7.69%).

Table 4.6 Number of OTUs for each arthropod order and their abundance (in percentage) at six sampling sites in Perak, Malaysia. OTUs that were identified to species level were also included in the footnote below.

	Beruas (%)[1]	**Beruas OP1 (%)**	**Beruas OP2 (%)**	**Gopeng (%)[2]**	**Ipoh (%)[3]**	**Pantai Remis (%)[4]**	**Total (%)**
Blattodea	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	2 (15.38)	2 (0.82)
Coleoptera	1 (6.67)	13 (92.86)	0 (0.00)	6 (18.75)	12 (10.81)	1 (7.69)	33 (13.47)
Diptera	12 (80.00)	1 (7.14)	23 (38.33)	25 (78.13)	91 (81.98)	6 (46.15)	158 (64.49)
Hemiptera	0 (0.00)	0 (0.00)	37 (61.67)	1 (3.13)	0 (0.00)	3 (23.08)	41 (16.73)
Hymenoptera	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	3 (2.70)	1 (7.69)	4 (1.63)
Lepidoptera	2 (13.33)	0 (0.00)	0 (0.00)	0 (0.00)	3 (2.70)	0 (0.00)	5 (2.04)
Odonata	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	2 (1.80)	0 (0.00)	2 (0.82)
Total	**15 (100.00)**	**14 (100.00)**	**60 (100.00)**	**32 (100.00)**	**111 (100.00)**	**13 (100.00)**	**245 (100.00)**

Note:

[1] Species identified for Coleoptera : *Deronectes platynotus*

[2] Species identified for Coleoptera : *Deronectes platynotus*

[3] Species identified for Coleoptera : *Deronectes platynotus*
Species identified for Odonata : *Ceriagrion auranticum*
Species identified for Diptera : *Kiefferulus tainanus*, *Spelobia bifrons*
Species identified for Hymenoptera : *Neothlipsis cincta*, *Tetramorium bicarinatum*,
Species identified for Lepidoptera : *Phoenicophanta modestula*, *Tortyra iocyaneus*

[4] Species identified for Coleoptera : *Deronectes platynotus*
Species identified for Hymenoptera : *Odontotermes hainanensis,*

Three arthropod orders were reported at the Beruas and Gopeng sampling sites. Of the 15 OTUs observed in Beruas, Diptera was found to display the highest composition (80.00%), followed by Lepidoptera (13.33%) and Coleoptera (6.67%). In Gopeng, out of the 32 OTUs, Diptera predominated in the house-farm swiftlet diet (78.13%), followed by Coleoptera (18.75%) and Hemiptera (3.13%).

Two arthropod orders were observed at the Beruas OP1 and Beruas OP2 sampling sites. In Beruas OP1, Coleoptera (92.86%) and Diptera (7.14%) were identified from the 14 OTUs. Meanwhile, in Beruas OP2, Hemiptera (61.67%) was more common, as compared to Diptera (38.33%).

CHAPTER 5

DISCUSSION

5.1 Taxonomic identification pipeline for the dietary arthropods

As compared to species assignment based directly on similarity index or threshold of BOLD database (Clare et al., 2009; Clare et al., 2011), the present study suggests that BOLD is a more robust approach as it is based on a tree-based criterion (Wilson et al., 2011). The bioinformatic pipeline developed based on the cloning results in this study can be summarised into four steps for taxonomic assignment of arthropod COI region.

(1) The Best ID, if given in BOLD searches using the public database, will be considered as the taxon identity. If "no match" (i.e., no Best ID) in this setting,

(2) The Best ID, if given in BOLD searches using the full database, will be considered as the taxon identity. If "no match" (i.e., no Best ID) in this setting,

(3) The order of the first top hit in BLASTn, if forms a cluster with the query sequence, will be considered as the taxon identity. If there is no clustering between the query sequence and the first top hit,

(4) The query will be considered uncertain taxa.

However, caution should be exercised when using the BOLD Systems. For the cases where the Best ID is provided based on the Public Record Barcode Database in BOLD (in Step 1), the queries should also be searched again based on the Public Record Barcode Database. This additional step is necessary as 56 cases of identification discrepancies were observed in this study (Tables 4.2 and 4.4). When any discrepancy occurs, it should be referred to the BLASTn results and phylogenetic analysis.

Therefore, it can be suggested that the Best ID (by BOLD Systems) should be verified by using the BLASTn search and NJ analysis to improve the accuracy of the taxonomic identification. The BLASTn search, coupled with NJ analysis, are also useful in determining the identities, at least at a higher taxonomic level (i.e., order), for a large number of "no match" cases in BOLD.

5.2 Overall diet composition of house-farm swiftlets

This study provides information on the diet profiles of the house farmed swiftlets (Table 4.6). Diptera appears to be one of the favourite prey items (64.49%) of the house-farm swiftlets, followed by Hemiptera (16.73%) and Coleoptera (13.47%).

In the present study, Diptera is the only arthropod order that is present in the faecal samples of all six sampling sites. Diptera consists of diverse members that occupy a wide range of habitats. It is not surprising to see the presence of

Dipterans in all kinds of landscapes such as mixed-type, urban and monocrop. One of the Dipterans identified belongs to the genus *Spelobia*, which is commonly associated with decaying materials, human or animal faecal samples (Roháček et al., 2001).

Hemiptera is ranked as the second highest (16.73%) consumed prey item in Table 4.6. This frequency is mainly attributed to the prey items in Beruas OP2, Gopeng and Pantai Remis. Being the third most found prey items (13.47%) in the diet of house-farm swiftlets, Coleoptera was found in all sampling sites except for Beruas OP2 (Table 4.6). This order was abundant in Beruas OP1 (39.39%) and Ipoh (36.36%), as compared to other locations. Lepidoptera accounted for 2.04% in the overall diet composition. Two moth families thawere confirmed to their identity are Noctuidae (owlet moths) and Choreutidae (metalmark moths) (Table 4.4).

The low frequency of Hymenoptera in this study contributes to one of the most notable differences in the diet profiles of the house-farm swiftlets and those of the white-nest swiftlets in the previous studies. In this study, Hymenoptera only made up 1.63% of the house-farm swiftlets' diet. However, in other dietary studies of white-nest swiftlets, Hymenoptera was the top consumed prey item (Langham, 1980; Lourie and Tompkins, 2000; Figure 5.1). Some of the OTUs were identified as tramp ants (*Tetramorium bicarinatum*), which is common in the oriental region, i.e. India, Nicobar Island, Thailand, Vietnam etc. (Guénard et al., 2017).

Odonata comprised 0.82% in the diet of house-farm swiftlets in this study. The damselflies species assigned was *Ceriagrion auranticum* that belongs to the family Coenagrionidae or recognised as Orange-tailed Sprite. Their habitat ranges from forest (tropical swamp) to inland wetlands (ponds, lakes, rivers, streams, marshes, pools etc.) and artificial ponds and water storage areas. Lastly, the overall composition of Blattodea in the house-farm swiftlet diet was 0.82%. Langham (1980) and Lourie and Tompkins (2000) also reported low percentage of termites in the food boluses of white-nest swiftlets, ranging from 0.1% to 2.1%. In this study, the Blattodea species identified was *Odontotermes hainanensis*.

Figure 5.1 shows the overall dietary arthropod profile obtained in this study and those from the previous studies for white-nest swiftlets (Langham, 1980; Lourie and Tompkins, 2000). Lourie and Tompkins suggested that the presence of large numbers of Ephemeroptera (mayflies) in Lourie and Tompkins's study was due to the co-incidence of swarming episodes. The occurrence of the swarming episodes could due to many reasons such as weather, rainfall, temperature, humidity etc (Nutting, 1969; Neoh and Lee, 2009). For the same reasoning, absence of Ephemeropterans in this study could be attributed to the monsoon season of sample collection.

Langham's observation was based on the house-farm swiftlets in Penang in 1970s when swiftlet farming in Malaysia has yet to become popular (Cranbrook et al. 2013), while Lourie and Tompkins' results were based on the white-nest swiftlet colony of the Gomantong Cave. The house-farm swiftlets were

suggested to be a new domesticated white-nest swiftlet group (Goh et al., 2018). The differences in the feeding profiles are possible indications of the distinct feeding behaviours between the cave swiftlets and the house-farm swiftlets, in line with Lourie and Tompkins' suggestion of diet resource partitioning among the swiftlet species, i.e., glossy swiftlets, mossy-nest swiftlet, black-nest swiftlets and white-nest swiftlets (Lourie and Tompkins, 2000). However, the diet profiles are inconsistent among the house-farm swiftlet colonies, as shown by the comparison between Langham's and the present studies (Figure 5.1).

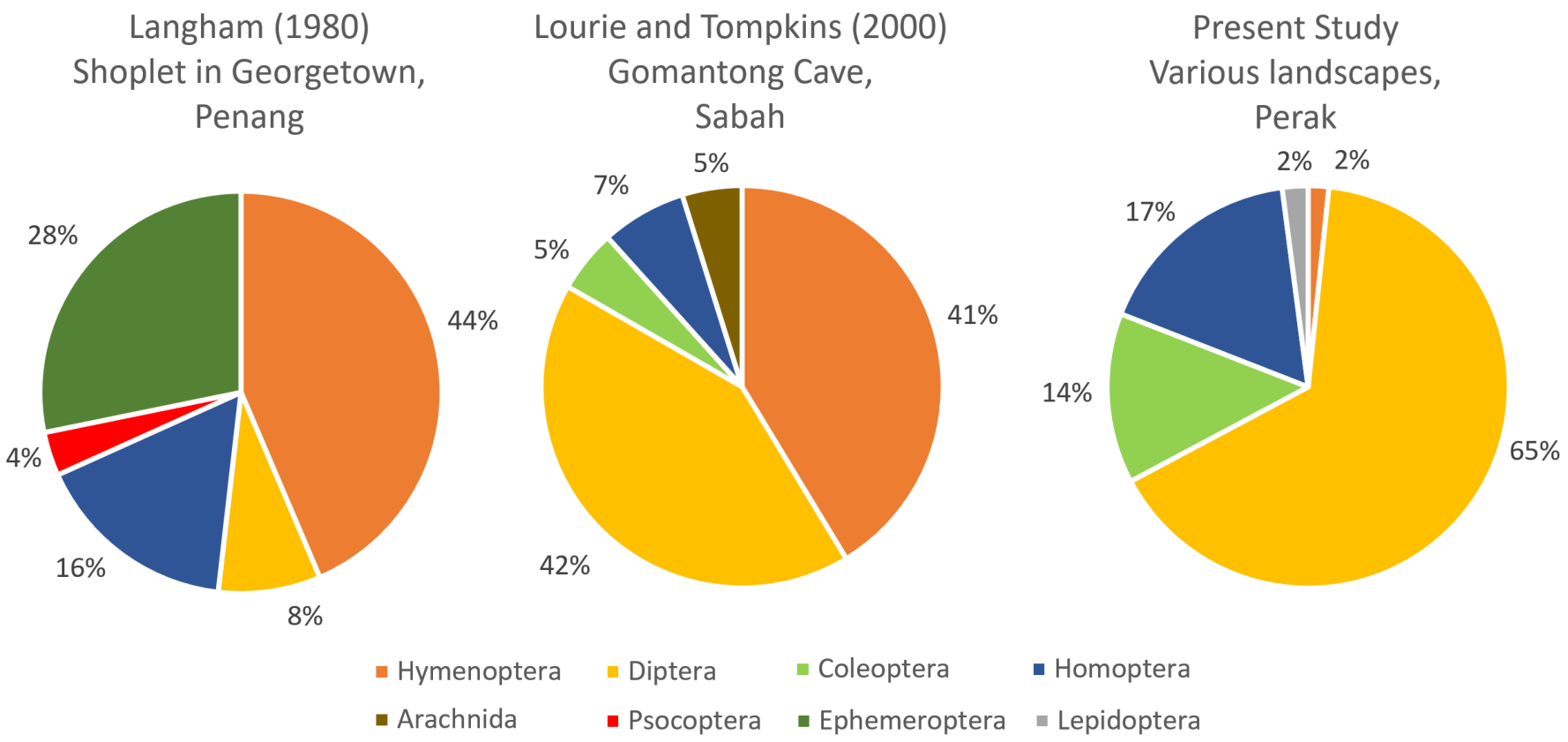


Figure 5.1 Comparison of the diet profiles of house-farm swiftlets (present study) with two other studies (Langham, 1980; Lourie and Tompkins, 2000. Only the top five arthropod orders consumed by the white-nest swiftlets are shown in each pie chart.

5.3 Urban landscape

The landscape characteristics of the Ipoh sampling site (hereafter referred to Ipoh) was (Figure 4.9 and Table 4.5) predominantly urban area (80.6%). Interestingly, the dietary arthropods were found to be highly diverse (consisting of five orders) and the highest in abundance (111 OTUs) as compared to other landscape types.

Diptera was the dominant group (81.98%) in the diet of the house-farm swiftlets. It is likely attributed to the presence of water bodies in the surrounding environment, as the high capture rate of insect prey was observed over the water bodies (Petkliang et al., 2017). A total of two rivers (Kinta River and Pinji River) and 116 lakes/ponds was observed within the 6 km radius from the swiftlet farm in Ipoh. These water bodies support the growth of aquatic insect larvae, many of which belong to the order Diptera (Fukui et al., 2006). Aquatic insects such as adult damselflies generally display a reduced flight ability due to wingbeat frequencies and the wing stroke pattern as compared to dragonflies (Wakeling and Ellington, 1997; Klym and Quinn, 2003), hence making them relatively easier preys. Ipoh could be an ideal for the Orange-tailed Sprite (Coenagrionidae) to breed due to the abundant freshwater water bodies such as lakes and rivers that are present (Wilson, 2009). However, the Orange-tailed Sprite was less abundant in the diet of house-farm swiftlets (Tables 4.4 and 4.6). Other damselflies families could be too large in body size (Larvae: 14 - 15mm; Adults: 19 - 31 mm; Wingspan: up to 41 mm) to be fed upon by the house-farm swiftlets. This may account for the low frequency of Odonata consumed (Zwick, 2001).

Another possible explanation for the high frequency of dietary dipterans at the sampling site is the introduction of fruit flies to the swiftlet farms. According to the swiftlet house breeder at Ipoh sampling site, some rotten fruits infested with insect larvae are routinely placed inside the swiftlet house, a practice which is believed to be able to continuously supply fruit flies for the house-farm swiftlets (personal communication). However, the genera of the fruit flies were not able to be confirmed as the swiftlet farmers thought that it was their trade secret.

An urban landscape always perceived as a concrete jungle that are dominated by non-native and homogeneous species, i.e. introduced garden flower and trees along the roadside (Aronson et al., 2017; Unterweger et al., 2017; Southon et al., 2018). So, many of the swiftlet house breeders have usually assumed that an urban landscape is not ideal for setting up swiftlet houses. This idea was further supported by Petkliang et al. (2017), who suggested that urban areas are not ideal habitats for the *Aerodramus inexpectatus germani* as less prey capture was observed. They justified that urban habitats have lower insect biomass than other habitats such as forests and agricultural land.

However, as shown in Table 4.6, the diet profile of house-farm swiftlets in the urban landscape of Ipoh shows has shown a diverse arthropod order, suggesting that some urban areas might be able to provide more food choices for the house-farm swiftlets. Two possible reasons are: (1) the diverse habitats surrounding the swiftlet house; (2) and fewer less insectivorous bird competitors.

Firstly, the diverse habitats surrounding the swiftlet house in an urban landscape is important for arthropods to persist. Ipoh is rich in natural habitats (such as mountains, fragmented forests, lakes, rivers and caves) and man-made habitats (such as remnants-patches of native vegetation, parks, home gardens, green roof-manmade artificial ecological system, urban wasteland, riparian corridor and stagnant drains). Some of these habitats are considered "urban green space" which can support urban arthropod diversity and species composition (Jaganmohan et al., 2018; Leonard et al., 2018; Melliger et al., 2018). For instance, some of the native plant species at Ipoh (personal observation) such as Banyan (*Ficus benghalensis*), mango (*Mangifera indica*) and introduced plant species Yellow Tulip Tree (*Spathodea campanulata Aurea*) in the urban green space can provide a habitat for urban animals and insects to persist. Besides, anthropogenic standing waters (such as rainwater barrels) in the urban landscape also serve as a platform to support the growth of mosquitoes (Diptera: Culicidae) as well as protect mosquito larvae from the attack of birds, reptiles and amphibians (Zittra et al., 2017).

Secondly, fewer insectivorous bird competitors in the city could have resulted in greater food availability for the house-farm swiftlets. The high noise levels such as the noise along roadways and in urban habitats, were found to have imposed negative effects on the bird species richness (Francis et al., 2009; Rodrigues et al., 2018). These noises are usually in high amplitudes and low spectral frequencies at about 0-10,000 Hz in the human altered urban landscape. Forest birds (i.e. *Turdus philomelos*, *Carduelis spinus*) migratory birds (i.e. *Columba palumbus*, *Cuculus canorus*) and ground nesting species (i.e.

Phylloscopus sibilatrix, *Erithacus rubecula*) were shown to have been affected the most in the urban area (Dale, 2018). To name some examples, the insectivorous birds such as Common Iora (*Aegithina tiphia*), Common Tailor Bird (*Orthotomus sutorius*), and Sunda Pygmy Woodpecker (*Dendrocopos moluccensis)* have declined in their population sizes in highly urbanized Singapore (Lim and Sodhi, 2004). However, the high feeding rate of the Ipoh house-farm swiftlets possibly reflects their adaptation they are adapted to the urban life and that the level of noise disturbance in Ipoh was tolerable to them.

5.4 Monocrop landscape

Beruas OP1 and Beruas OP2 are covered by more than 80% of the monocrop plantation, namely the oil palms, *Elaeis guineensis* (Table 4.5). These two locations have the lowest arthropod diversity (only two orders) as compared to the other locations (Table 4.6).

The formation of monocrop landscape is often a result of deforestation of the primary rainforests in Southeast Asia, followed by the expansion of oil palm plantations (Sodhi et al., 2010; Wilcove et al., 2013; Barnes et al., 2014). The change of land use has caused an extensive loss of biodiversity, massive reduction in species richness and functional diversity (Foster et al., 2011; Barnes et al., 2014; Edwards et al., 2014). The native plant species in the forest plays important roles in providing refuges to the diverse arthropod taxa. Studies have suggested that the primary forest houses a greater diversity of arthropods than the oil palm plantation (Fayle et al., 2010; Edwards et al., 2014). For instance,

Turner and Foster (2009) reported that the arthropod diversity in the primary forest in Sabah became reduced drastically by to 77%, following the massive clearance of the area. Fayle et al. (2010) revealed the loss of the ant species was as high as 81% (250 of the total 309 ant species surveyed) in the oil palm plantation when compared to the primary forest. Some of the arthropod species even dominated in monocrop landscapes as compared to forest (Asfiya et al., 2015). For instance, tramp ant species *Anoplolepis gracilipes*, *Monomorium floricola* and *Nylanderia vaga* dominated the cocoa plantations that were previously rainforests. Besides, Zheng et al. (2015) also reported a 42.6%-50.0% drop in spider diversity in monocrop landscape which was converted from primary forests.

Despite the observation that the Beruas OP1 and Beruas OP2 sites share a similar landscape characteristic, i.e., dominated by oil palm plantations (around 86%), the diet profiles were greatly different. Beruas OP1 has recorded a low OTU abundance (14 OTUs) of which 94.68% were coleopterans, while Beruas OP2 had 61.67% hemipterans and 38.33% dipterans (Table 4.7). The presence of Hemiptera in Beruas OP2 (but absent in Beruas OP1) could be due to the small paddy plantation (area not estimated) in Beruas OP2, which could support a phytophagous hemipteran population. Hemipterans are known to fly up to an elevation of 1,000 m assisted by the air current and could possibly be caught by house-farm swiftlets when they swarm in the dawn and dusks (Brodsky, 1994).

Oil palm cultivation is dependent on oil palm pollinating weevils (*Elaeidobius kamerunicus*). The oil palm pollinating weevil were introduced to Malaysia from Cameroon in 1980 due to poor natural oil palm pollination problem in Malaysia (Syed, 1979). This species first released in Johor and Sabah Pamol plantations and soon spread quickly across the whole Malaysia within a few years (Syed et al., 1982). After pollination by this pollinator, the oil palm fruit yield increased by 20% in Peninsular Malaysia and 53% in Sabah (Ponnamma, 1999). Then, hand pollination was soon obsolete in most parts of the country. This species was found to adapt both wet and dry seasons in Malaysia and able to carry the most pollen cartage compare other weevil pollinators species (Syed, 1979; Syed et al., 1982).

Syed (1979) found that the *Elaeidobius kamerunicus* is very host-specific to the oil palm. This species dependent on the male inflorescence of the oil palm. Adult weevils usually feed on the oil palm pollens and lay eggs in anthesizing male inflorescences. There is no evidence that show whether *Elaeidobius kamerunicus* weevil swarms like other insects like Hymenoptera and termites, as there was no report on the swarming behaviour on this species.

Weevils commonly reported to fly as high as 2-3m high as compared to the flying altitude of white-nest swiftlet at 100-200m. However, it was told by the swiftlet farmers that the house-farm swiftlets commonly fly in between the oil palm trees (personal communication). It was therefore of interest to know if house-farm swiftlets were consuming these beneficial insects since Beruas OP1 and Beruas OP2 swiftlet farms were located in the monocrop landscape.

Although none of the OTUs obtained was identified as this species, a verification analysis was performed to check the similarity of the 13 Coleopteran OTUs obtained for Beruas OP1 against the COI data sequenced for four weevil individuals. The NJ tree (Figure 5) showed all 12 Coleoptera OTUs were distinct from the four weevil individuals. Therefore, it can be concluded that none of these OTUs was the oil palm weevil.

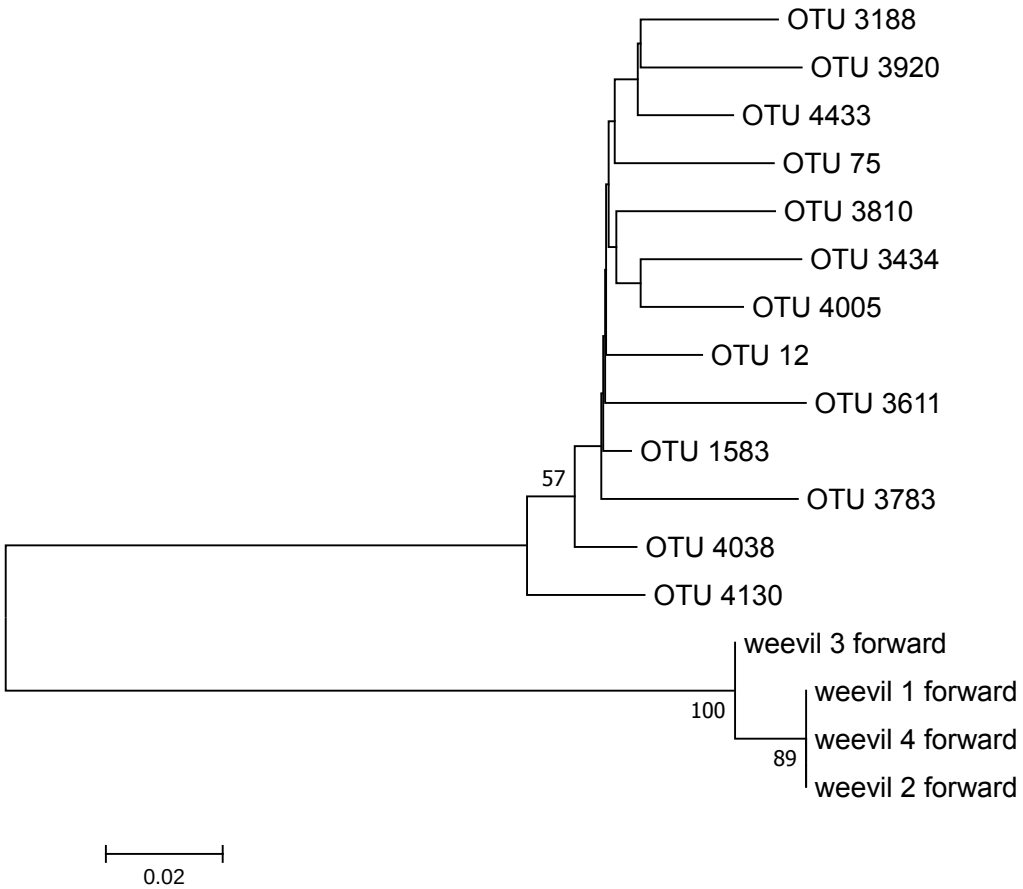


Figure 5.2 NJ tree reconstructed based on the 218-bp COI region of 13 Coleoptera OTU sequences from Beruas OP1 and four weevils obtained from the oil palm plantations (bootstrap more than 50%).

5.5 Mixed-used landscape

Three to five arthropod orders occurred in the diet of the swiftlets in the mixed-use landscapes (Beruas, Gopeng and Pantai Remis). Diptera was the most consumed arthropod order for all three sites (Table 4.6). Termites (Blattodea: *Odontotermes hainanensis*) were in the diet at Pantai Remis only. The high biodiversity of even small forest patches in the mixed-use landscape was likely to account for the greater variety of arthropods in the swiftlet diet (Basset et al., 2012).

The arthropod diversity increases with the trophic rank in the forest communities of urban areas (Melliger et al., 2018). In Melliger et al (2018) study, Hymenoptera and Arachnida were observed with the increase of forest size, despite the expansion of urbanisation. This was further supported by Lewis and Basset (2007), that in the tropical forests, some of the insect order such as Coleoptera and Hymenoptera decreased at the disturbed forests or the cleared forest (deforestation). Hymenoptera such as ants (Formicidae) and fig wasps (Agaonidae) are the major food components in the diet of white-nest swiftlets (Langham, 1980; Lourie and Tompkins, 2000). Therefore, the present study suggests that the heterogeneous landscape which consists of plantations, urban areas, water bodies and fragmented forest can be important for the continuous food supply (arthropods) of the house-farm swiftlets.

5.6 Implications for swiftlet farming practices

The varying diet profiles of the house-farm swiftlets in Perak reflect the availability of insect preys as swiftlets are opportunistic feeders. Our results implied that the house-farm swiftlet colonies can persist in the urban and mixed-type landscapes because these landscapes provide the house-farm swiftlets a diverse and continuous food (arthropod) supply. In addition, the swiftlets in the urban areas are also less exposed to predation, negating the usual practice of setting up electric nets and traps, or poisonous food, to get rid of predators like barn owls and snakes etc. On the other hand, house-farm swiftlets in the monocrop plantations are potentially negatively affected, as the diverse arthropod diversity can be diminished upon land clearance and by the insecticide applications in large-scale for insect pest control.

5.7 Limitations of the present study and recommendations for the future studies

The observations made in the present study were mostly limited by the small number of swiftlet faecal samples, sampling locations and single sampling episodes. The results would be more convincing if the faecal samples of house-farm swiftlets were compared daily throughout a year, which took account into monsoon and dry seasons, and also El Nino effect, whenever possible. Differences of the diet profiles between the cave swiftlets and the house-farm swiftlets can be further attested using a DNA meta-profiling approach in the future studies. Since Diptera was the major food component in the swiftlet diet, the potential roles of the house-farm swiftlets as a biological control agent for urban dipterans pests (i.e. mosquitoes) can also be investigated.

Ground truthing such as insect survey through various trapping method, should be carried out to provide a more reliable abundance of the insect found near to the swiftlet farm. With the obtained data, the feeding ecology of the house-farm swiftlet could be postulated and see if house-farm swiftlets would consume the high abundance of insect in a specific area or they would have preference in their diet. However, the ground truthing could not be performed as most of the covered land was private owned properties which is not accessible and allowing insect survey to be conducted.

Limited DNA record in both BOLD and GenBank database resulted in the difficulty for species-identification using the partial COI region. Kvist (2013) revealed that, out of 1,242,040 recognised arthropod COI sequences used in his study, only 149,997 sequences (12.08%) could be represented in BOLD whereas 69,123 (5.56%) were represented in GenBank or NCBI (Kvist, 2013). Both databases only made up a total of 189,319 (15.24%), in which seven times lesser than the recognised arthropod COI sequences (Kvist, 2013; Zhang, 2011). Thus, more arthropod specimen should be collected and deposited the COI sequences into both databases to make the databases more complete.

Due to the incomplete DNA database for the insects found in Malaysia, the insect taxa identified using molecular methods should also be cross-checked with the species distribution record (as mentioned in section 4.3 and 5.1; Table 4.4). For instance, Crisol-Martinez et al. (2016) compared the insect records in Australia before concluding the insect taxa in the diet of birds. Since a comprehensive insect checklist for Malaysia is unavailable, expertise from

regional entomologists would helpful in confirming the presence of arthropod species within the geographical range. Online databases of the neighbouring countries, such as the Digital Reference Collection for Singapore's Biodiversity (https://singapore.biodiversity.online/) by Ng et al. (2011) would provide a good reference.

CHAPTER 6

CONCLUSIONS

The present study suggests that the Barcodes of Life Data (BOLD) identifications based on both records (All Barcode Records on BOLD and Public Record Barcode Database), BLASTn and Neighbour-joining (NJ) analysis are needed for a relatively more precise taxonomic assignment of arthropod in swiftlet diet. A bioinformatic pipeline was developed using the partial mitochondrial COI sequences cloned from the swiftlet faeces collected in Sitiawan. Due to the incomplete of the DNA database for the insects found in Malaysia, the insect taxa identified using molecular methods should also be cross-checked with the species distribution record.

High-throughput sequencing in this study generated a total of 4,852 operational taxonomic units (OTUs). Of the 266 arthropod DNA sequences, 20.30% (total of 54) of OTUs showed no genetic affinity with their respective top BLASTn hits and were therefore regarded as uncertain identity, while 79.70% of the OTUs (total of 212) were resolved either at order-level or species-level. Following the taxonomic assignment pipeline established in the preliminary assessment, the diet composition of the house-farm swiftlets in Perak was found to comprise Diptera (62.74%), followed by Hemiptera (18.87%), Coleoptera (12.26%), Lepidoptera (2.36%), Hymenoptera (1.89%), Blattodea (0.94%) and Odonata (0.94%).

This study also showed that the urban area and mixed-used landscape could likely to be suitable habitats to establish swiftlet houses, as indicated by the diverse orders of insects found in these two landscape types. Although the urban landscape is commonly perceived as having low arthropod diversity, the presence or absence of the "urban green space" can greatly influence the persistence and diversity of the arthropods. Mixed-used landscape that is composed of heterogeneous landscape features can provide a continuous food supply for the house-farm swiftlets.

On the contrary, monocrop landscapes such as large-scale oil palm plantations may not be ideal for the house-farm swiftlets because the continuous food supply in the monocrop landscape could be adversely affected by certain farming practices such as pesticide application. In the present study, the possibility of the presence of the domestic pests, such as houseflies (Diptera), and agricultural pests, such as the rhinoceros beetle (Coleoptera) in the house-farm swiftlet diet could not be ruled out. The role of house-farm swiftlets to provide ecological service as insect pest predators in all three landscape types should be further investigated.

REFERENCES

Ali, H., Muhammad, A., Sanda Bala, N. and Hou, Y., 2018. The Endosymbiotic Wolbachia and Host *COI* Gene Enables to Distinguish Between Two Invasive Palm Pests; Coconut Leaf Beetle, *Brontispa longissima* and Hispid Leaf Beetle, *Octodonta nipae*. *Journal of Economic Entomology*, 111 (6), pp. 2894 - 2902.

Altschul, S.F. et al., 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215 (3), pp. 403 - 410.

Aronson, M.F. et al., 2017. Biodiversity in the city: key challenges for urban green space management. *Frontiers in Ecology and the Environment*, 15 (4), pp. 189 - 196.

Asfiya, W.A.R.A. et al., 2015. Intensive agroforestry practices negatively affect ant (Hymenoptera: Formicidae) diversity and composition in southeast Sulawesi, Indonesia. *Asian Myrmecology*, 7 (1), pp. 87 - 104.

Barnes, A.D. et al., 2014. Consequences of tropical land use for multitrophic biodiversity and ecosystem functioning. *Nature Communications*, 5, p. 5351.

Basiron, Y., 2007. Palm oil production through sustainable plantations. *European Journal of Lipid Science and Technology*, 109 (4), pp. 289 - 295.

Basset, Y. et al., 2012. Arthropod diversity in a tropical forest. *Science*, 338 (6113), pp. 1481 - 1484.

Bohmann, K. et al., 2011. Molecular diet analysis of two African free-tailed bats (Molossidae) using high throughput sequencing. *PloS ONE*, 6 (6), p. e21441.

Brandon-Mong, G.J. et al., 2015. DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. *Bulletin of Entomological Research*, 105 (6), pp. 717 - 727.

Brooke, R.K., 1970. Taxonomic and evolutionary notes on the subfamilies, tribes, genera and subgenera of the swifts (Aves: Apodidae). *Durban Museum Novitates*, 9 (2), pp. 13 - 24.

Brodsky. A.K., 1994. *The evolution of insect flight*. Oxford University Press, London.

Burgar, J.M. et al., 2014. Who's for dinner? High□throughput sequencing reveals bat dietary differentiation in a biodiversity hotspot where prey taxonomy is largely undescribed. *Molecular Ecology*, 23 (15), pp. 3605 - 3617.

Burhanuddin, M. and Noor, H.M., 2017. Ranging Behaviour of Edible Nest Swiftlet (*Aerodramus* sp.) in Kuala Langat District, Selangor, Malaysia. *Malaysian Applied Biology*, 46 (2), pp. 59 - 66.

Cheng, S. et al., 2011. Evidence for a higher number of species of *Odontotermes* (Isoptera) than currently known from Peninsular Malaysia from mitochondrial DNA phylogenies. *PLoS ONE*, 6 (6), p. e20992.

Clare, E.L. et al., 2009. Species on the menu of a generalist predator, the eastern red bat (*Lasiurus borealis):* using a molecular approach to detect arthropod prey. *Molecular Ecology*, 18 (11), pp. 2532 - 2542.

Clare, E.L. et al., 2011. Eating local: influences of habitat on the diet of little brown bats (*Myotis lucifugus*). *Molecular Ecology*, 20 (8), pp. 1772 - 1780.

Connolly, C., 2017. "Bird cages and boiling pots for potential diseases": contested ecologies of urban 'Swiftlet farming' in George Town, Malaysia. *Journal of Political Ecology*, 24 (1), pp. 24-43.

Cranbrook, E.O. and Medway, L., 1965. Lack of ultrasonic frequencies in the calls of swiftlets (*Collocalia* spp.). *Ibis*, 107 (2), pp. 258 - 258.

Cranbrook, E.O., Goh, W.L., Lim, C.K. and Mustafa, A.R., 2013. The species of white nest swiftlets (Apodidae, Collocaliini) of Malaysia and the origins of house-farm birds: Morphometric and genetic evidence. *Forktail*, 29, pp. 107 - 119.

Crisol-Martínez, E. et al., 2016. Using next-generation sequencing to contrast the diet and explore pest-reduction services of sympatric bird species in macadamia orchards in Australia. *PLoS ONE*, 11 (3), p. e0150159.

Dale, S., 2018. Urban bird community composition influenced by size of urban green spaces, presence of native forest, and urbanization. *Urban Ecosystems*, 21 (1), pp. 1 - 14.

Deagle, B.E. et al., 2007. Studying seabird diet through genetic analysis of faeces: a case study on macaroni penguins (*Eudyptes chrysolophus*). *PLoS ONE*, 2 (9), p. e831.

Deagle, B.E., Chiaradia, A., McInnes, J. and Jarman, S.N., 2010. Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out?. *Conservation Genetics*, 11 (5), pp. 2039 - 2048.

Dillery, D.G., 1965. Post-mortem digestion of stomach contents in the Savannah Sparrow. *The Auk*, 82 (2), pp. 281 - 281.

Duckett, J.E. and Karuppiah, S., 1990. A guide to the planter in utilizing barn owls (*Tyto alba*) as an effective biological control in mature oil palm plantations. *Proceedings of the 1989 International Palm Oil Development Conference-Agriculture.* 5-9 September 1989 Kuala Lumpur, Malaysia. Kuala Lumpur: Palm Oil Research Institute of Malaysia, pp. 357 - 372.

Edwards, F.A. et al., 2014. Does logging and forest conversion to oil palm agriculture alter functional diversity in a biodiversity hotspot?. *Animal Conservation*, 17 (2), pp. 163 - 173.

Fayle, T.M. et al., 2010. Oil palm expansion into rain forest greatly reduces ant biodiversity in canopy, epiphytes and leaf-litter. *Basic and Applied Ecology*, 11 (4), pp. 337 - 345.

Folmer, O. et al., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3 (5), pp. 294 - 299.

Foster, W.A. et al., 2011. Establishing the evidence base for maintaining biodiversity and ecosystem function in the oil palm landscapes of South East Asia. *Philosophical Transactions of the Royal Society B*, 366 (1582), pp. 3277 - 3291.

Francis, C.D., Ortega, C.P. and Cruz, A., 2009. Noise pollution changes avian communities and species interactions. *Current Biology*, 19 (16), pp. 1415 - 1419.

Fukui, D.A.I., Murakami, M., Nakano, S. and Aoi, T., 2006. Effect of emergent aquatic insects on bat foraging in a riparian forest. *Journal of Animal Ecology*, 75 (6), pp. 1252 - 1258.

Goh, W.L. et al., 2018. Genetic diversity among white-nest swiftlets of the genus *Aerodramus* (Aves: Apodidae: Collocaliini) of house-farms in Malaysia. *Raffles Bulletin of Zoology*, 66, pp. 350 - 360.

Guenard, B.S. et al., 2017. The Global Ant Biodiversity Informatics (GABI) database: synthesizing data on the geographic distribution of ant species (Hymenoptera: Formicidae). *Myrmecological News*, 24, pp. 83 - 89.

Guo, C.T. et al., 2006. Edible bird's nest extract inhibits influenza virus infection. *Antiviral Research*, 70 (3), pp. 140 - 146.

Hails, C.J. and Amiruddin, A., 1981. Food samples and selectivity of white-bellied Swiftlets, *Collocalia esculenta*. *Ibis*, 123, pp. 329 - 333.

Hajibabaei, M. et al., 2006. A minimalist barcode can identify a specimen whose DNA is degraded. *Molecular Ecology Notes*, 6 (4), pp. 959 - 964.

Hajibabaei, M. et al., 2011. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS ONE*, 6 (4), p. e17497.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series*, January 1999. London, United Kingdom. London: Information Retrieval Ltd., pp. 95-98.

Harrisson, T., 1974. The food of *Collocalia* swiftlets (Aves, Apodidae) at Niah Great Cave in Borneo. *Journal Bombay National History Society*, 71 (3), pp. 376 - 393.

Hebert, P.D., Cywinska, A. and Ball, S.L., 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, 270 (1512), pp. 313 - 321.

Hope, P.R. et al., 2014. Second generation sequencing and morphological faecal analysis reveal unexpected foraging behaviour by *Myotis nattereri* (Chiroptera, Vespertilionidae) in winter. *Frontiers in Zoology*, 11 (1), p. 39.

Hu, Q. et al., 2016. Edible bird's nest enhances antioxidant capacity and increases lifespan in *Drosophila melanogaster*. *Cellular and Molecular Biology*, 62 (4), pp. 116 - 122.

Ibrahim, S.H., Teo, W.C. and Baharun, A., 2009. A study on suitable habitat for swiftlet farming. *Journal of Civil Engineering, Science and Technology*, 1 (1), pp. 1 - 7.

Ibrahim, W.K.W., Yaacob, M.R. and Abdullah, A., 2015. The importance of technical knowledge in sustainability of Malay bird's nest industry in Malaysia. *Journal of Applied Environmental and Biological Sciences*, 5 (4), pp. 190 - 196.

Jaganmohan, M., Vailshery, L.S., Mundoli, S. and Nagendra, H., 2018. Biodiversity in sacred urban spaces of Bengaluru, India. *Urban Forestry and Urban Greening*, 32, pp. 64 - 70.

Jedlicka, J.A., Sharma, A.M. and Almeida, R.P., 2013. Molecular tools reveal diets of insectivorous birds from predator fecal matter. *Conservation Genetics Resources*, 5 (3), pp. 879 - 885.

Jedlicka, J.A., Vo, A.T.E. and Almeida, R.P., 2016. Molecular scatology and high-throughput sequencing reveal predominately herbivorous insects in the diets of adult and nestling Western Bluebirds (*Sialia mexicana*) in California vineyards. *The Auk*, 134 (1), pp. 116 - 127.

Jeyaprakash, A. and Hoy, M.A., 2000. Long PCR improves Wolbachia DNA amplification: *wsp* sequences found in 76% of sixty-three arthropod species. *Insect Molecular Biology*, 9 (4), pp. 393 - 405.

Kamarudin, M.I. and Khoo, C.K., 2011. *Megaselia scalaris* as feed of domesticated *Aerodramus fuciphagus* swiftlet. *Proceedings of International Conference and Training on Swiftlet Ranching*, 17 - 19 July 2011 Kuala Terengganu, Malaysia. Kuala Terengganu: Universiti Sultan Zainal Abidin, pp. 1 – 5.

King, R.A., Symondson, W.O.C. and Thomas, R.J., 2015. Molecular analysis of faecal samples from birds to identify potential crop pests and useful biocontrol agents in natural areas. *Bulletin of Entomological Research*, 105 (3), pp. 261 - 272.

Kjer, K.M., Simon, C., Yavorskaya, M. and Beutel, R.G., 2016. Progress, pitfalls and parallel universes: a history of insect phylogenetics. *Journal of the Royal Society Interface*, 13 (121), p. 20160363.

Klym, M. and Quinn, M.A., 2003. *Introduction to Dragonfly and Damselfly Watching*. Texas: Texas Parks and Wildlife.

Kohn, M.H. and Wayne, R.K., 1997. Facts from feces revisited. *Trends in Ecology and Evolution*, 12 (6), pp. 223 - 227.

Kong, Y.C. et al., 1987. Evidence that epidermal growth factor is present in swiftlet's (Collocalia) nest. *Comparative Biochemistry and Physiology*, 87 (2), pp. 221 - 226.

Kumar, S., Stecher, G. and Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33 (7), pp. 1870 - 1874.

Kvist, S., 2013. Barcoding in the dark?: a critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge. *Molecular Phylogenetics and Evolution*, 69 (1), pp. 39 - 45.

Langham, N., 1980. Breeding biology of the edible□nest swiftlet *Aerodramus fuciphagus*. *Ibis*, 122 (4), pp. 447 - 461.

Larkin, M.A. et al., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23 (21), pp. 2947 - 2948.

Lee, P.L., Clayton, D.H., Griffiths, R. and Page, R.D., 1996. Does behavior reflect phylogeny in swiftlets (Aves: Apodidae)? A test using cytochrome b mitochondrial DNA sequences. *Proceedings of the National Academy of Sciences*, 93 (14), pp. 7091 - 7096.

Leonard, R.J., McArthur, C. and Hochuli, D.F., 2018. Habitat complexity does not affect arthropod community composition in roadside greenspaces. *Urban Forestry and Urban Greening*, 30, pp. 108 - 114.

Leray, M. et al., 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, 10 (1), p. 34.

Lewis, O.T. and Basset, Y.V.E.S., 2007. Insect conservation in tropical forests. *Insect Conservation Biology*, pp. 34-56.

Librado, P. and Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25 (11), pp. 1451 - 1452.

Lim, C.K. and Cranbrook, E.O., 2014. *Swiftlets of Borneo: Builders of Edible Nests*, 2nd ed. Kota Kinabalu: Natural History Publications.

Lim, H.C. and Sodhi, N.S., 2004. Responses of avian guilds to urbanisation in a tropical city. *Landscape and Urban Planning*, 66 (4), pp. 199 - 215.

Looi, Q.H., Ideris, A., Zakaria, M.Z.B.A.B. and Omar, A.R.B., 2015. Morphology comparison of swiftlet species from natural and man-made habitats in Malaysia. *Sains Malaysiana*, 44 (4), pp. 497 - 502.

Lourie, S.A. and Tompkins, D.M., 2000. The diets of Malaysian swiftlets. *Ibis*, 142 (4), pp. 596 - 602.

Manchi, S. and Sankaran, R., 2009. Predators of swiftlets and their nests in the Andaman and Nicobar Islands. *Indian Birds*, 5 (4), pp. 118 - 120.

Mansor, M.S., Nor, S.M. and Ramli, R., 2018a. Assessing Diet of the Rufous-Winged Philentoma (*Philentoma pyrhoptera*) in Lowland Tropical Forest using Next-Generation Sequencing. *Sains Malaysiana*, 47 (5), pp. 1045 - 1050.

Mansor, M.S., Abdullah, N.A., Abdullah Halim, M.R., Md. Nor, S. and Ramli, R., 2018b. Diet of tropical insectivorous birds in lowland Malaysian rainforest. *Journal of Natural History*, 52 (35-36), pp. 2301 - 2316.

Mansor, M.S., Nor, S.M., Ramli, R. and Sah, S.A.M., 2018c. Niche shift in three foraging insectivorous birds in lowland Malaysian forest patches. *Behavioural Processes*, 157, pp. 73 - 79.

Marzuki, H.A.F., 1994. *Prinsip-prinsip budidaya pemeliharaan burung wallet*. Surabaya: Biro Pusat Rehabilitasi Sarang Burung.

Melliger, R.L., Braschler, B., Rusterholz, H.P. and Baur, B., 2018. Diverse effects of degree of urbanisation and forest size on species richness and functional diversity of plants, and ground surface-active ants and spiders. *PLoS ONE*, 13 (6), p. e0199245.

Meusnier, I. et al., 2008. A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*, 9 (1), p. 214.

Misof, B. et al., 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346 (6210), pp. 763 – 767.

Munch, K. et al., 2008. Statistical assignment of DNA sequences using Bayesian phylogenetics. *Systematic Biology*, 57 (5), pp. 750 - 757.

Ng, M.H., Chan, K.H. and Kong, Y.C., 1986. Potentiation of mitogenic response by extracts of the swiftlet's (Collocalia) nest. *Biochemistry International*, 13 (3), pp. 521 - 531.

Ng, P.K., Corlett, R. and Tan, H.T., 2011. Singapore Biodiversity: An Encyclopedia of The Natural Environment and Sustainable Development, 1st ed. Singapore: Editions Didier Millet.

Nurshuhada, S. et al., 2015. Study on the performance of the ekasih swiftlet house–a low cost alternative to promote the swiftlet industry. *Malaysian Journal of Veterinary Research*, 6 (2), pp. 9 - 22.

Othman, A., Yatim, S.H., Razak, N.A. and Azimat, N.A., 2008. Distribution of edible-nest swiftlet (*Aerodramus fuciphagus*) bird house in Peninsular Malaysia. *Journal of Wildlife and Parks*, 25, pp. 8-11.

Petkliang, N., Gale, G.A., Brunton, D.H. and Bumrungsri, S., 2017. Wetland, forest, and open paddy land are the key foraging habitats for Germain's swiftlet (*Aerodramus inexpectatus germani*) in Southern Thailand. *Tropical Conservation Science*, 10, p. 1940082917698467.

Piñol, J., Mir, G., Gomez□Polo, P. and Agustí, N., 2015. Universal and blocking primer mismatches limit the use of high□throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, 15 (4), pp. 819 - 830.

Ponnamma, K.N., 1999. Diurnal variation in the population of *Elaeidobius kamerunicus* on the anthesising male inflorescences of oil palm. *Planter*, 75 (881), pp. 405 - 410.

Price, J.J., Johnson, K.P. and Clayton, D.H., 2004. The evolution of echolocation in swiftlets. *Journal of Avian Biology*, 35 (2), pp. 135 - 143.

Rahman, M.A., Ghazali, P.L. And Lian, C.J., 2018. Environmental parameters in successful edible bird nest swiftlet houses in Terengganu. *Journal of*

Sustainability Science and Management, 13 (1), pp. 127 - 131.

Rahman, S.F.S., Nor, H.M. and Burhanuddin, M., 2016. Diversity of the insects in the diet of edible nest swiftlets in oil palm plantations. *Journal of Biodiversity and Environmental Sciences*, 8 (4), pp. 39 - 48.

Ramji, M.F.S., Lim, C.K. and Rahman, M.A., 2013. Roosting and nest-building behaviour of the white-nest swiftlet *Aerodramus fuciphagus* (Thunberg) (Aves: Apodidae) in farmed colonies. *Raffles Bulletin of Zoology*, 29, pp. 225 - 235.

Ramlan, A.R. et al., 2017, Jabatan Meteorologi Malaysia Annual Report 2017 [Online]. Available at: http://www.met.gov.my/content/pdf/penerbitan/laporantahunan/laporantahunan2017.pdf [Accessed: 27 June 2019].

Ratnasingham, S. and Hebert, P.D., 2007. BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). *Molecular Ecology Notes*, 7 (3), pp. 355 - 364.

Razgour, O. et al., 2011. High□throughput sequencing offers insight into mechanisms of resource partitioning in cryptic bat species. *Ecology and Evolution*, 1 (4), pp. 556 - 570.

Rodrigues, A.G., Borges-Martins, M. and Zilio, F., 2018. Bird diversity in an urban ecosystem: the role of local habitats in understanding the effects of urbanization. *Iheringia. Série Zoologia*, 108, p. e2018017.

Roháček, J. et al., 2001. World Catalog of Sphaeroceridae (Diptera) Opava: Slezské Zemské Museum, p. 414.

Rosenberg, K.V. and Cooper, R.J., 1990. Approaches to avian diet analysis. *Studies in Avian Biology*, 13, pp. 80 - 90.

Sien, L.S., Chuan, C.H., Lihan, S. and Yee, L.T., 2013. Isolation and identification of airborne bacteria inside swiftlet houses in Sarawak, Malaysia. *Makara Journal of Science*. 17 (3), pp. 104 - 108.

Sodhi, N.S. et al., 2010. Conserving Southeast Asian forest biodiversity in human-modified landscapes. *Biological Conservation*, 143 (10), pp. 2375 - 2384.

Southon, G.E. et al., 2018. Perceived species-richness in urban green spaces: Cues, accuracy and well-being impacts. *Landscape and Urban Planning*, 172, pp. 1 - 10.

Suzanti, F., Kuswardani, R.A., Rahayu, S., Susanto, A., 2016. Contribution of epiphytes on the canopy insect population in oil palm plantations in North Sumatera. *ARPN Journal of Engineering and Applied Sciences*, 11 (11), pp. 6982 - 6998.

Syed, R.A., 1979. Studies on oil palm pollination by insects. *Bulletin of Entomological Research*, 69 (2), pp. 213 - 224.

Syed, R.A., Law, I.H. and Corley, R.H.V., 1982. Insect pollination of oil palm: introduction, establishment and pollinating efficiency of *Elaeidobius kamerunicus* in Malaysia. *Planter*, 58 (681), pp. 547 - 561.

Tan, H.K., Chia, F.C. and Ong, A.H.K., 2018. Approach to improve edible bird nest quality and establishing better bird nest cleaning process facility through best value approach. *Journal for the Advancement of Performance Information and Value*, 10 (1), pp. 38 - 50.

Thomassen, H.A. et al., 2003. A new phylogeny of swiftlets (Aves: Apodidae) based on *cytochrome-b* DNA. *Molecular Phylogenetics and Evolution*, 29 (1), pp. 86 - 93.

Thomassen, H.A., den Tex, R.J., de Bakker, M.A. and Povel, G.D.E., 2005. Phylogenetic relationships amongst swifts and swiftlets: a multi locus approach. *Molecular Phylogenetics and Evolution*, 37 (1), pp. 264 - 277.

Thorburn, C.C., 2015. The edible nest swiftlet industry in Southeast Asia: Capitalism meets commensalism. *Human Ecology*, 43 (1), pp. 179 - 184.

Turner, E.C. and Foster, W.A., 2009. The impact of forest conversion to oil palm on arthropod abundance and biomass in Sabah, Malaysia. *Journal of tropical ecology*, *25*(1), pp.23-30.

Unterweger, P.A., Schrode, N. and Betz, O., 2017. Urban nature: perception and acceptance of alternative green space management and the change of awareness after provision of environmental information. a chance for biodiversity protection. *Urban Science*, 1 (3), p. 24.

Vesterinen, E.J., Lilley, T., Laine, V.N. and Wahlberg, N., 2013. Next generation sequencing of fecal DNA reveals the dietary diversity of the widespread insectivorous predator Daubenton's bat (*Myotis daubentonii*) in Southwestern Finland. *PLoS ONE*, 8 (11), p. e82168.

Vimala, B., Hussain, H. and Nazaimoon, W.W., 2012. Effects of edible bird's nest on tumour necrosis factor-alpha secretion, nitric oxide production and cell viability of lipopolysaccharide-stimulated RAW 264.7 macrophages. *Food and Agricultural Immunology*, 23 (4), pp. 303 - 314.

Wakeling, J.M. and Ellington, C.P., 1997. Dragonfly flight. II. Velocities, accelerations and kinematics of flapping flight. *Journal of Experimental Biology*, 200 (3), pp. 557 - 582.

Waugh, D.R. and Hails, C.J., 1983. Foraging ecology of a tropical aerial feeding bird guild. *Ibis*, 125 (2), pp. 200 - 217.

Wilcove, D.S. et al., 2013. Navjot's nightmare revisited: logging, agriculture, and biodiversity in Southeast Asia. *Trends in Ecology and Evolution*, 28 (9), pp.531 - 540.

Wilson, K.D.P., 2009. *Ceriagrion auranticum* [Online]. Available at: http://dx.doi.org/10.2305/IUCN.UK.2009-2.RLTS.T164790A5927151.en. [Accessed: 4 July 2019].

Wilson, J.J. et al., 2011. When species matches are unavailable are DNA barcodes correctly assigned to higher taxa? An assessment using sphingid moths. *BMC Ecology*, 11 (1), p. 18.

Wood, B.J. and Fee, C.G., 2003. A critical review of the development of rat control in Malaysian agriculture since the 1960s. *Crop Protection*, 22 (3), pp.445 - 461.

Yahya, B.E.B. and Hamdan, M.A.A., 2014. Inventori fauna semut di kawasan rezab hidupan liar Sungkai, Perak, Malaysia. *Journal of Wildlife and Parks*, 27, pp. 133 - 136.

Yew, M.Y. et al., 2018. Edible bird's nest improves motor behavior and protects dopaminergic neuron against oxidative and nitrosative stress in Parkinson's disease mouse model. *Journal of Functional Foods*, 48, pp. 576 - 585.

Yu, D.W. et al., 2012. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, 3 (4), pp. 613 - 623.

Zeale, M.R. et al., 2011. Taxon□specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, 11 (2), pp. 236 - 244.

Zhang, Z.Q., 2011. *Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness*. Auckland: Magnolia press.

Zheng, G., Li, S. and Yang, X., 2015. Spider diversity in canopies of Xishuangbanna rainforest (China) indicates an alarming juggernaut effect of rubber plantations. *Forest Ecology and Management*, *338*, pp.200-207.

Zhou, X. et al., 2013. Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. *Gigascience*, 2 (1), p. 4.

Zittra, C. et al., 2017. Landscape structure affects distribution of potential disease vectors (Diptera: Culicidae). *Parasites and Vectors*, 10 (1), p. 205.

Zug, R. and Hammerstein, P., 2012. Still a host of hosts for Wolbachia: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. *PLoS ONE*, 7 (6), p. e38544.

Zwick, P., 2001. Corbet, PS (1999): Dragonflies: behaviour and ecology of Odonata. *Aquatic Insects*, 23 (1), p. 83.

APPENDICES

APPENDIX A

Table A: Decision on the taxonomic assignment down to order-level for 188 OTUs.

No.	OTUs	BOLD Public Record			BOLD All Record			Decision
		Order	Family	Genus/Species	Order	Family	Genus/Species	
1	OTU 1	No match	-	-	Diptera	-	-	Diptera
2	OTU 4	No match	-	-	Hemiptera	-	-	Hemiptera
3	OTU 12	No match	-	-	Coleoptera	Nitidulidae	*Epuraea luteolus*	Coleoptera
4	OTU 18	No match	-	-	Coleoptera	-	-	Coleoptera
5	OTU 20	No match	-	-	Hemiptera	-	-	Hemiptera
6	OTU 30	No match	-	-	Coleoptera	-	-	Coleoptera
7	OTU 36	No match	-	-	Diptera	-	-	Diptera
8	OTU 37	No match	-	-	Diptera	-	-	Diptera
9	OTU 39	No match	-	-	Coleoptera	-	-	Coleoptera
10	OTU 44	No match	-	-	Diptera	-	-	Diptera
11	OTU 52	No match	-	-	Diptera	-	-	Diptera
12	OTU 75	No match	-	-	Coleoptera	-	-	Coleoptera
13	OTU 90	No match	-	-	Hemiptera	-	-	Hemiptera
14	OTU 148	No match	-	-	Diptera	-	-	Diptera
15	OTU 165	No match	-	-	Diptera	-	-	Diptera
16	OTU 188	No match	-	-	Diptera	-	-	Diptera
17	OTU 206	No match	-	-	Hemiptera	-	-	Hemiptera
18	OTU 238	No match	-	-	Diptera	-	-	Diptera
19	OTU 243	No match	-	-	Diptera	-	-	Diptera
20	OTU 270	No match	-	-	Diptera	-	-	Diptera

Table A (Cont'd)

No.	OTUs	BOLD Public Record			BOLD All Record			Decision
		Order	Family	Genus/Species	Order	Family	Genus/Species	
21	OTU 278	No match	-	-	Diptera	-	-	Diptera
22	OTU 300	No match	-	-	Diptera	-	-	Diptera
23	OTU 319	No match	-	-	Diptera	-	-	Diptera
24	OTU 326	No match	-	-	Diptera	-	-	Diptera
25	OTU 332	No match	-	-	Diptera	-	-	Diptera
26	OTU 335	No match	-	-	Diptera	-	-	Diptera
27	OTU 343	No match	-	-	Diptera	-	-	Diptera
28	OTU 348	No match	-	-	Diptera	-	-	Diptera
29	OTU 388	No match	-	-	Hemiptera	-	-	Hemiptera
30	OTU 421	No match	-	-	Diptera	-	-	Diptera
31	OTU 423	No match	-	-	Hemiptera	-	-	Hemiptera
32	OTU 428	No match	-	-	Hemiptera	-	-	Hemiptera
33	OTU 492	No match	-	-	Diptera	-	-	Diptera
34	OTU 505	No match	-	-	Diptera	-	-	Diptera
35	OTU 509	No match	-	-	Diptera	-	-	Diptera
36	OTU 522	No match	-	-	Hemiptera	-	-	Hemiptera
37	OTU 564	No match	-	-	Hemiptera	-	-	Hemiptera
38	OTU 596	No match	-	-	Diptera	-	-	Diptera
39	OTU 597	No match	-	-	Diptera	-	-	Diptera
40	OTU 601	No match	-	-	Diptera	-	-	Diptera
41	OTU 605	No match	-	-	Diptera	-	-	Diptera
42	OTU 617	No match	-	-	Diptera	-	-	Diptera

Table A (Cont'd)

No.	OTUs	BOLD Public Record			BOLD All Record			Decision
		Order	Family	Genus/Species	Order	Family	Genus/Species	
43	OTU 748	No match	-	-	Diptera	-	-	Diptera
44	OTU 777	No match	-	-	Diptera	-	-	Diptera
45	OTU 825	No match	-	-	Diptera	-	-	Diptera
46	OTU 874	No match	-	-	Diptera	-	-	Diptera
47	OTU 876	No match	-	-	Diptera	-	-	Diptera
48	OTU 898	No match	-	-	Hemiptera	-	-	Hemiptera
49	OTU 940	No match	-	-	Diptera	-	-	Diptera
50	OTU 950	No match	-	-	Diptera	-	-	Diptera
51	OTU 962	No match	-	-	Diptera	-	-	Diptera
52	OTU 965	No match	-	-	Diptera	-	-	Diptera
53	OTU 977	No match	-	-	Diptera	-	-	Diptera
54	OTU 978	No match	-	-	Diptera	-	-	Diptera
55	OTU 986	No match	-	-	Hemiptera	-	-	Hemiptera
56	OTU 995	No match	-	-	Hymenoptera	-	-	Hymenoptera
57	OTU 1000	No match	-	-	Diptera	-	-	Diptera
58	OTU 1030	No match	-	-	Diptera	-	-	Diptera
59	OTU 1135	No match	-	-	Diptera	-	-	Diptera
60	OTU 1136	No match	-	-	Diptera	-	-	Diptera
61	OTU 1145	No match	-	-	Diptera	-	-	Diptera
62	OTU 1208	No match	-	-	Diptera	-	-	Diptera
63	OTU 1328	No match	-	-	Diptera	-	-	Diptera
64	OTU 1329	No match	-	-	Diptera	-	-	Diptera

Table A (Cont'd)

No.	OTUs	BOLD Public Record			BOLD All Record			Decision
		Order	Family	Genus/Species	Order	Family	Genus/Species	
65	OTU 1402	No match	-	-	Diptera	-	-	Diptera
66	OTU 1405	No match	-	-	Diptera	-	-	Diptera
67	OTU 1414	No match	-	-	Diptera	-	-	Diptera
68	OTU 1417	No match	-	-	Hemiptera	-	-	Hemiptera
69	OTU 1418	No match	-	-	Hymenoptera	-	-	Hymenoptera
70	OTU 1448	No match	-	-	Hemiptera	-	-	Hemiptera
71	OTU 1461	No match	-	-	Diptera	-	-	Diptera
72	OTU 1471	No match	-	-	Diptera	-	-	Diptera
73	OTU 1483	No match	-	-	Diptera	-	-	Diptera
74	OTU 1548	No match	-	-	Diptera	-	-	Diptera
75	OTU 1582	No match	-	-	Diptera	-	-	Diptera
76	OTU 1583	No match	-	-	Coleoptera	-	-	Coleoptera
77	OTU 1590	No match	-	-	Diptera	-	-	Diptera
78	OTU 1636	No match	-	-	Hemiptera	-	-	Hemiptera
79	OTU 1661	No match	-	-	Diptera	-	-	Diptera
80	OTU 1674	No match	-	-	Hemiptera	-	-	Hemiptera
81	OTU 1695	No match	-	-	Diptera	-	-	Diptera
82	OTU 1735	No match	-	-	Diptera	-	-	Diptera
83	OTU 1736	No match	-	-	Lepidoptera	Noctuidae	*Noctua fimbriata*	Lepidoptera
84	OTU 1754	No match	-	-	Diptera	-	-	Diptera
85	OTU 1755	No match	-	-	Hemiptera	-	-	Hemiptera
86	OTU 1757	No match	-	-	Diptera	-	-	Diptera
87	OTU 1774	No match	-	-	Diptera	-	-	Diptera

Table A (Cont'd)

No.	OTUs	BOLD Public Record			BOLD All Record			Decision
		Order	Family	Genus/Species	Order	Family	Genus/Species	
88	OTU 1819	No match	-	-	Hemiptera	-	-	Hemiptera
89	OTU 1855	No match	-	-	Hemiptera	-	-	Hemiptera
90	OTU 1873	No match	-	-	Diptera	-	-	Diptera
91	OTU 1908	No match	-	-	Hemiptera	-	-	Hemiptera
92	OTU 1916	No match	-	-	Diptera	-	-	Diptera
93	OTU 1984	No match	-	-	Diptera	-	-	Diptera
94	OTU 2044	No match	-	-	Diptera	-	-	Diptera
95	OTU 2045	No match	-	-	Diptera	-	-	Diptera
96	OTU 2061	No match	-	-	Diptera	-	-	Diptera
97	OTU 2082	No match	-	-	Diptera	-	-	Diptera
98	OTU 2099	No match	-	-	Hemiptera	-	-	Hemiptera
99	OTU 2148	No match	-	-	Diptera	-	-	Diptera
100	OTU 2150	No match	-	-	Diptera	-	-	Diptera
101	OTU 2172	No match	-	-	Diptera	-	-	Diptera
102	OTU 2210	No match	-	-	Diptera	-	-	Diptera
103	OTU 2276	No match	-	-	Diptera	-	-	Diptera
104	OTU 2297	No match	-	-	Lepidoptera	-	-	Lepidoptera
105	OTU 2397	No match	-	-	Diptera	-	-	Diptera
106	OTU 2574	No match	-	-	Diptera	-	-	Diptera
107	OTU 2595	No match	-	-	Hemiptera	-	-	Hemiptera
108	OTU 2828	No match	-	-	Hemiptera	-	-	Hemiptera
109	OTU 2831	No match	-	-	Diptera	-	-	Diptera
110	OTU 2836	No match	-	-	Diptera	-	-	Diptera

Table A (Cont'd)

No.	OTUs	BOLD Public Record			BOLD All Record			Decision
		Order	Family	Genus/Species	Order	Family	Genus/Species	
111	OTU 2897	No match	-	-	Diptera	-	-	Diptera
112	OTU 2927	No match	-	-	Diptera	-	-	Diptera
113	OTU 2946	No match	-	-	Diptera	-	-	Diptera
114	OTU 3046	No match	-	-	Diptera	-	-	Diptera
115	OTU 3136	No match	-	-	Hemiptera	-	-	Hemiptera
116	OTU 3174	No match	-	-	Diptera	-	-	Diptera
117	OTU 3179	No match	-	-	Diptera	-	-	Diptera
118	OTU 3188	No match	-	-	Coleoptera	Nitidulidae	*Epuraea luteolus*	Coleoptera
119	OTU 3206	No match	-	-	Diptera	-	-	Diptera
120	OTU 3225	No match	-	-	Diptera	-	-	Diptera
121	OTU 3325	No match	-	-	Diptera	-	-	Diptera
122	OTU 3346	No match	-	-	Diptera	-	-	Diptera
123	OTU 3385	No match	-	-	Diptera	-	-	Diptera
124	OTU 3407	No match	-	-	Diptera	-	-	Diptera
125	OTU 3434	No match	-	-	Coleoptera	Nitidulidae	*Epuraea luteolus*	Coleoptera
126	OTU 3526	No match	-	-	Diptera	-	-	Diptera
127	OTU 3559	No match	-	-	Diptera	-	-	Diptera
128	OTU 3596	No match	-	-	Hemiptera	-	-	Hemiptera
129	OTU 3597	No match	-	-	Hemiptera	-	-	Hemiptera
130	OTU 3611	No match	-	-	Coleoptera	Nitidulidae	*Epuraea luteolus*	Coleoptera
131	OTU 3613	No match	-	-	Diptera	-	-	Diptera
132	OTU 3622	No match	-	-	Diptera	-	-	Diptera
133	OTU 3653	No match	-	-	Diptera	-	-	Diptera

Table A (Cont'd)

No.	OTUs	BOLD Public Record			BOLD All Record			Decision
		Order	Family	Genus/Species	Order	Family	Genus/Species	
134	OTU 3678	No match	-	-	Diptera	-	-	Diptera
135	OTU 3743	No match	-	-	Hemiptera	-	-	Hemiptera
136	OTU 3746	No match	-	-	Diptera	-	-	Diptera
137	OTU 3753	No match	-	-	Diptera	-	-	Diptera
138	OTU 3775	No match	-	-	Diptera	-	-	Diptera
139	OTU 3779	No match	-	-	Diptera	-	-	Diptera
140	OTU 3783	No match	-	-	Coleoptera	Nitidulidae	*Epuraea luteolus*	Coleoptera
141	OTU 3801	No match	-	-	Diptera	-	-	Diptera
142	OTU 3810	No match	-	-	Coleoptera	-	-	Coleoptera
143	OTU 3829	No match	-	-	Diptera	-	-	Diptera
144	OTU 3830	No match	-	-	Diptera	-	-	Diptera
145	OTU 3845	No match	-	-	Diptera	-	-	Diptera
146	OTU 3867	No match	-	-	Lepidoptera	Nymphalidae	*Pierella luna*	Lepidoptera
147	OTU 3886	No match	-	-	Diptera	-	-	Diptera
148	OTU 3920	No match	-	-	Coleoptera	Nitidulidae	*Epuraea luteolus*	Coleoptera
149	OTU 3951	No match	-	-	Diptera	-	-	Diptera
150	OTU 3961	No match	-	-	Diptera	-	-	Diptera
151	OTU 3974	No match	-	-	Hemiptera	-	-	Hemiptera
152	OTU 4005	No match	-	-	Coleoptera	-	-	Coleoptera
153	OTU 4038	No match	-	-	Coleoptera	Nitidulidae	*Epuraea luteolus*	Coleoptera
154	OTU 4043	No match	-	-	Hemiptera	-	-	Hemiptera
155	OTU 4075	No match	-	-	Diptera	-	-	Diptera
156	OTU 4126	No match	-	-	Hemiptera	-	-	Hemiptera

Table A (Cont'd)

No.	OTUs	BOLD Public Record			BOLD All Record			Decision
		Order	Family	Genus/Species	Order	Family	Genus/Species	
157	OTU 4130	No match	-	-	Coleoptera	Nitidulidae	*Epuraea luteolus*	Coleoptera
158	OTU 4149	No match	-	-	Diptera	-	-	Diptera
159	OTU 4174	No match	-	-	Diptera	-	-	Diptera
160	OTU 4181	No match	-	-	Diptera	-	-	Diptera
161	OTU 4200	No match	-	-	Diptera	-	-	Diptera
162	OTU 4231	No match	-	-	Diptera	-	-	Diptera
163	OTU 4243	No match	-	-	Diptera	-	-	Diptera
164	OTU 4264	No match	-	-	Diptera	-	-	Diptera
165	OTU 4269	No match	-	-	Diptera	-	-	Diptera
166	OTU 4301	No match	-	-	Diptera	-	-	Diptera
167	OTU 4309	No match	-	-	Hemiptera	-	-	Hemiptera
168	OTU 4376	No match	-	-	Diptera	-	-	Diptera
169	OTU 4391	No match	-	-	Hemiptera	-	-	Hemiptera
170	OTU 4433	No match	-	-	Coleoptera	Nitidulidae	*Epuraea luteolus*	Coleoptera
171	OTU 4440	No match	-	-	Diptera	-	-	Diptera
172	OTU 4453	No match	-	-	Diptera	-	-	Diptera
173	OTU 4470	No match	-	-	Diptera	-	-	Diptera
174	OTU 4484	No match	-	-	Diptera	-	-	Diptera
175	OTU 4488	No match	-	-	Hemiptera	-	-	Hemiptera
176	OTU 4497	No match	-	-	Diptera	-	-	Diptera
177	OTU 4501	No match	-	-	Hemiptera	-	-	Hemiptera
178	OTU 4536	No match	-	-	Hemiptera	-	-	Hemiptera
179	OTU 4539	No match	-	-	Diptera	-	-	Diptera

Table A (Cont'd)

No.	OTUs	BOLD Public Record			BOLD All Record			Decision
		Order	Family	Genus/Species	Order	Family	Genus/Species	
180	OTU 4571	No match	-	-	Diptera	-	-	Diptera
181	OTU 4623	No match	-	-	Hemiptera	-	-	Hemiptera
182	OTU 4637	No match	-	-	Hemiptera	-	-	Hemiptera
183	OTU 4675	No match	-	-	Hemiptera	-	-	Hemiptera
184	OTU 4739	No match	-	-	Hemiptera	-	-	Hemiptera
185	OTU 4773	No match	-	-	Hemiptera	-	-	Hemiptera
186	OTU 4777	No match	-	-	Diptera	-	-	Diptera
187	OTU 4833	No match	-	-	Diptera	-	-	Diptera
188	OTU 4844	No match	-	-	Hemiptera	-	-	Hemiptera