# RESOLVING COMPLEXITY WITHIN *Rattus rattus* IN RELATION TO LEPTOSPIROSIS

NURUL HUDA BINTI MOHAMAD IKBAL

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

# RESOLVING COMPLEXITY WITHIN *Rattus rattus* IN RELATION TO LEPTOSPIROSIS

**NURUL HUDA BINTI MOHAMAD IKBAL**

**THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**

**INSTITUTE OF BIOLOGICAL SCIENCES**
**FACULTY OF SCIENCE**
**UNIVERSITY OF MALAYA**
**KUALA LUMPUR**

**2019**

**UNIVERSITY OF MALAYA**

**ORIGINAL LITERARY WORK DECLARATION**

Name of Candidate: **NURUL HUDA BINTI MOHAMAD IKBAL**

Matric No: **SHC150053**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of Thesis ("this Work"):

**RESOLVING COMPLEXITY WITHIN *Rattus rattus* IN RELATION TO LEPTOSPIROSIS**

Field of Study:

**GENETICS AND MOLECULAR BIOLOGY**

I do solemnly and sincerely declare that:

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

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:
Designation:

**RESOLVING COMPLEXITY WITHIN *Rattus rattus* IN RELATION TO LEPTOSPIROSIS**

## ABSTRACT

*Rattus rattus* is an economically important pest widely distributed throughout Peninsular Malaysia. The taxonomic problems of this species are still debated to date. Although asymptomatic, they are currently recognised as main carriers of various zoonotic diseases, inclusive of leptospirosis which represents a serious threat for human health. This species complex has huge range of morphological variations which creates difficulties in species identification. It is vital to elucidate the complexities within this species in order to discover the population structure, associated diseases' transmission patterns, and host-pathogen relationships. This study aims to clarify the species complexities and population structure of *R. rattus* in Peninsular Malaysia using DNA-based molecular method. Conventional morphometrics was incorporated to examine the skull variation within *R. rattus*. Moreover, as a recognised carrier of leptospirosis in Malaysia, the prevalence of *Leptospira* spp. circulating among urban rats and the disease's transmission patterns within the populations are utmost critical. Transcriptomics was used to determine the functional genes; to understand the adaptation of these genes in context to disease susceptibility and resistance in rats. The result from molecular data revealed a distant relationship between Southeast Asian *R. rattus* towards *R. rattus* sensu stricto and Himalayan *R. rattus* (up to 5.6% K2P). The presence of *R. rattus cf*. lineage IV sensu Aplin et al., 2011 and *R. tanezumi* as a valid sympatric species was also disclosed. Through skull variation analysis, only a single morphotype of *R. rattus* was identified throughout Peninsular Malaysia. Sexual dimorphism and ontogenetic variation were the evident of non-geographic variations. Population structure of *R. rattus* was investigated based on three mitochondrial markers (COI, Cyt b, D-loop) and one nuclear gene (IRBP). High level of nucleotide and haplotype diversities were detected in *R. rattus* from

Peninsular Malaysia, indicating a high level of genetic diversity which demonstrated low population differentiation with significant *p*-value ($p < 0.001$). Besides, phylogenetic analysis of leptospiral DNA found in either kidney or liver of *R. rattus* revealed circulation of pathogenic *L. interrogans* and *L. borgpetersenii* in Malaysia's house rat. The prevalence patterns (e.g. sex, age-groups, niches) were also discovered. This study reports an update of molecular evidence on rodent leptospirosis in Peninsular Malaysia and will be useful for further research in this area. It also confirms the efficiency of 16S rRNA and LipL32 for the diagnosis of leptospirosis. Due to inflammation which can cause sepsis during leptospirosis, gene expressions of four cytokines; tumor necrosis factor-α (TNF-α), interleukin (IL)-1β, IL-6, IL-10, and chemokines of macrophage inflammatory protein-1α (MIP-1α/C-C-type chemokine ligand 3, CCL3) and gamma interferon (IFN-γ)-inducible protein 10 kDa (IP-10/C-X-C-type chemokine ligand 10, CXCL10) were studied in both kidney and liver of *R. rattus*. Interleukin-6 was found to be expressed during the early stage in the liver of *R. rattus*, while all the six genes were upregulated during the late stage of leptospirosis in the kidney of *R. rattus*. Overall, these studies formulate a conceptual framework based on host and pathogen relationships which can lead to the formation and documentation of leptospirosis transmission patterns within *R. rattus*.

**Keywords**: conventional morphometrics, molecular approaches, population structure, leptospiral prevalence, transcriptomics.

**MELERAIKAN KEKOMPLEKSAN DALAM *Rattus rattus* BERHUBUNG LEPTOSPIROSIS**

## ABSTRAK

*Rattus rattus* merupakan salah satu spesies tikus perosak ekonomi yang penting dan tersebar di seluruh Semenanjung Malaysia. Masalah taksonomi melibatkan spesies ini masih diperdebatkan sehingga kini. Walaupun asimptomatik, tikus ini terkenal sebagai pembawa utama pelbagai penyakit zoonotik, termasuk leptospirosis yang mengakibatkan ancaman serius terhadap kesihatan manusia. Spesies kompleks ini mempunyai variasi morfologi yang sangat ketara menyebabkan kesukaran dalam proses pengecaman spesies. Kerumitan yang melibatkan spesies ini adalah sangat penting untuk dirungkai bagi membolehkan struktur populasi, pola transmisi penyakit, dan hubungan perumah-patogen dapat dikaji. Kajian ini dijalankan untuk menjelaskan kekompleksan spesies dan struktur populasi *R. rattus* di Semenanjung Malaysia menggunakan kaedah molekul asid deoksiribonukleik (DNA). Kaedah morfometri konvensional digunakan untuk memeriksa variasi tengkorak pada *R. rattus*. Disebabkan *R. rattus* terkenal sebagai agen pembawa penyakit leptospirosis di Malaysia, kelaziman spesies bakteria *Leptospira* yang dibawa oleh tikus bandar ini dan pola transmisi penyakit leptospirosis adalah sangat kritikal untuk dikaji. Pendekatan transkriptomik digunakan dalam menentukan gen berfungsi, untuk memahami adaptasi gen berfungsi tersebut dalam konteks suseptibiliti dan ketahanan *R. rattus*. Hasil kajian melalui data molekul mendedahkan hubungan genetik yang jauh antara *R. rattus* dari Asia Tenggara dengan *R. rattus* sensu stricto dan *R. rattus* dari Himalaya (jarak genetik sehingga 5.6%). Kehadiran *R. rattus cf*. lineage IV sensu Aplin et al., 2011 dan *R. tanezumi* sebagai spesies simpatrik turut didedahkan. Melalui analisis variasi tengkorak, hanya satu morfotip *R. rattus* yang dikenalpasti di Semenanjung Malaysia. Dimorfisme seks dan variasi ontogenetik merupakan bukti kewujudan variasi bukan geografi. Struktur populasi *R. rattus* pula diselidik menggunakan tiga penanda

mitokondria (COI, Cyt b, D-loop) dan satu gen nuklear (IRBP). Aras diversiti nukleotida dan haplotip yang tinggi dikenalpasti pada populasi *R. rattus* dari Semenanjung Malaysia, menunjukkan aras keragaman genetik yang tinggi dan pembezaan populasi yang rendah dengan perbezaan keertiaan ($p < 0.001$). Di samping itu, analisis filogenetik DNA *Leptospira* yang ditemui sama ada di ginjal atau hati *R. rattus* mendedahkan peredaran dua spesies patogen *L. interrogans* dan *L. borgpetersenii* pada tikus rumah. Pola kelaziman (jantina, kumpulan umur, nic) turut dikaji. Kajian ini melaporkan pengemaskinian kaedah molekul dalam pengesanan leptospirosis melalui tikus di Semenanjung Malaysia dan akan bermanfaat untuk kajian akan datang dalam bidang ini. Kajian ini juga mengesahkan kecekapan gen 16S rRNA dan LipL32 dalam mendiagnosis leptospirosis. Disebabkan keradangan yang boleh mengakibatkan sepsis semasa jangkitan leptospirosis, ungkapan gen empat sitokin iaitu *tumor necrosis factor-α* (*TNF-α*), *interleukin* (*IL*)*-1β*, *IL-6*, *IL-10*, dan simokin *macrophage inflammatory protein-1α* (*MIP-1α/C-C-type chemokine ligand 3, CCL3*) serta *gamma interferon* (*IFN-γ*)*-inducible protein 10 kDa* (*IP-10/C-X-C-type chemokine ligand 10, CXCL10*) telah diuji pada ginjal dan hati *R. rattus*. *Interleukin-6* terekspres semasa jangkitan leptospirosis tahap awal pada hati *R. rattus*, manakala keseluruhan enam gen yang dikaji menunjukkan pengawalaturan tinggi semasa tahap lewat jangkitan pada ginjal *R. rattus*. Secara keseluruhan, penyelidikan ini merumuskan rangka kerja konseptual berdasarkan hubungan antara perumah dan patogen yang boleh memacu kepada pembentukan dan dokumentasi pola transmisi leptospirosis oleh *R. rattus*.

**Kata kunci**: morfometri konvensional, pendekatan molekul, struktur populasi, kelaziman leptospirosis, transkriptomik.

## ACKNOWLEDGEMENTS

"In the Name of Allah, the Most Gracious, the Most Merciful". Alhamdulillah, I am thankful to Almighty Allah, who in His infinite blessings has strengthen and guided me to complete this research. May the peace and blessings of Allah be upon His Prophet Muhammad (peace be upon him), his family, and his companions. First and foremost, I would like to express my sincere appreciation to my dedicated supervisors; Dr. Hasmahzaiti Omar, Assoc. Prof. Dr. Subha Bhassu, and Dr. Khanom Simarani who gave me the opportunity to work under their supervisions. Their invaluable guidance and constructive suggestions were the key for me to complete this study. Warmest thanks to the Ministry of Higher Education for giving me scholarship to further my study. The financial support provided by the University of Malaya was greatly appreciated. It is also my pleasure to say thank you to Faculty of Science, mainly Institute of Biological Sciences. Deepest thanks to my beloved family who deserve the utmost gratitude especially my wonderful parents and my adorable siblings for their unconditional love, support, understanding, and warm encouragements that inspired me. They truly have been my biggest supporters throughout graduate school and really, my entire life. I would also like to express my heartfelt dedication to all of my friends, especially AGAGEL labmates, for their guidance, contributions, generous assistance, and moral supports. Chan Chew Chin really deserved this dedication for helping me in many ways throughout my research life. I owe you so much, Chen. I wish to acknowledge this debt and profoundly thank each and every one of those individuals who generously gave their time and expertise to review this research work. They may remain anonymous but their efforts are publicly acknowledged. To those whom I failed to mention but have helped in one way or another, may Allah bless them. Lastly, I am gratified to everyone who lent their hand in this venture, directly or indirectly, that eventually driven me to produce this final masterpiece. To all of you, thank you, without you, it would be a near impossible task.

## TABLE OF CONTENTS

**ORIGINAL LITERARY WORK DECLARATION..... ii**

**ABSTRACT..... iii**

**ABSTRAK..... v**

**ACKNOWLEDGEMENTS..... vii**

**TABLE OF CONTENTS..... viii**

**LIST OF FIGURES..... xiii**

**LIST OF TABLES..... xvi**

**LIST OF SYMBOLS AND ABBREVIATIONS..... xix**

**LIST OF APPENDICES..... xxii**

**CHAPTER 1: INTRODUCTION..... 1**

1.1 General Introduction..... 1

1.2 Problem Statements..... 6

1.3 Significant of the Research..... 7

1.4 Research Questions..... 8

1.5 Research Objectives..... 9

**CHAPTER 2: LITERATURE REVIEW..... 10**

2.1 Order Rodentia and Its Diversity..... 10

2.2 The Black Rat (*Rattus rattus*)..... 11

2.3 Cytogenetic Analyses on *Rattus rattus* Complex..... 12

2.4 The Taxonomy of the *Rattus* Complex and Problems Encountered During Species Classification..... 14

2.5 Address on Classification Using Morphometrics Tool..... 19

2.5.1 Conventional morphometrics in identification of rodents..... 20

2.6 Molecular Analysis as a Powerful Complementary Tool..... 22

2.7 Animal Mitochondrial DNA (mtDNA) and Nuclear DNA (nDNA)..... 27

2.8 Rodents as Carrier of Zoonotic Diseases..... 30

2.9 Rodents and Leptospirosis..... 31

2.10 Diagnosis of Rodent Leptospirosis..... 34

2.11 Rodents Immune Response Against Leptospiral Infection..... 37

2.12 Host-Pathogen Transcriptomics by RNA Sequencing..... 38

2.13 Novel Findings on Cytokines and Chemokines Expressions During Leptospirosis Infection..... 40

**CHAPTER 3: METHODOLOGY..... 43**

3.1 Ethics and Permits Approval..... 43

3.2 Sampling and Study Area..... 43

3.3 Species Diversity of *Rattus* (Rodentia: Muridae) on the Malay Peninsula and Their Phylogenetic Relationships Inferred from Complete Mitochondrial Cytochrome b Gene Sequences..... 46

3.3.1 Extraction of genomic DNA……... 47

3.3.2 Determination of DNA quality and concentration……..... 47

3.3.3 DNA amplification by the polymerase chain reaction (PCR)……..... 48

3.3.4 Gel electrophoresis…… 49

3.3.5 Purification of PCR products…….. 49

3.3.6 DNA sequencing……..... 50

3.3.7 Sequence editing and alignment……..... 50

3.3.8 Sequence and phylogenetic analysis……..... 51

3.3.9 Molecular clock analyses……..…….. 52

3.4 Population Structure of *Rattus rattus* in Peninsular Malaysia Based on Mitochondrial and Nuclear DNA Data……..…… 53

3.4.1 Sequence analysis……..……. 55

3.4.2 Phylogenetic analysis……. 56

3.4.3 Population genetics analysis……... 56

3.5 Morphometric Analysis of Skull Characteristics of *Rattus rattus*……..…….. 57

3.5.1 Skinning..... 57

3.5.2 Skull extraction..... 58

3.5.3 Conventional morphometrics..... 59

3.5.4 Statistical analysis..... 60

3.6 Prevalence and Genetic Divergence of *Leptospira* spp. in *Rattus rattus*..…..... 62

3.6.1 Sequence and phylogenetic analysis..... 64

3.7 Transcriptome Study of *Rattus rattu*s Infected with *Leptospira* spp……. 65

3.7.1 Isolation of total RNA…… 65

3.7.2 Determination of RNA quality and concentration……. 67

3.7.3 RNA sequencing (RNA-Seq) library construction……..... 68

3.7.4 Bioinformatics workflow……... 68

3.7.4.1 Sequencing reads filtering…….. 69

3.7.4.2 Genome mapping…… 69

3.7.4.3 Novel transcripts prediction…… 69

3.7.4.4 SNP and INDEL detection……. 70

3.7.4.5 Differentially splicing gene detection……. 70

3.7.4.6 Gene expression analysis…… 71

3.7.4.7 Differentially expressed gene detection…….. 71

3.7.4.8 Hierarchical clustering analysis of DEGs……..... 71

3.7.4.9 Gene ontology analysis of DEGs…… 71

3.7.4.10 Pathway analysis of DEGs……..... 72

3.7.5 Reverse transcription and first strand complementary DNA (cDNA) synthesis……..…….... 72

3.7.6 Primer design and optimisation for quantitative RT-PCR

(qRT-PCR) amplification……..…… 73

3.7.7 Quantification of mRNA expression on healthy and leptospiral-infected *Rattus rattus* by qRT-PCR……..…… 76

3.7.8 qRT-PCR statistical analysis……..……... 78

**CHAPTER 4: RESULTS..... 79**

4.1 Species Diversity of *Rattus* (Rodentia: Muridae) on the Malay Peninsula and Their Phylogenetic Relationships Inferred from Complete Mitochondrial Cytochrome b Gene Sequences..... 79

4.2 Population Structure of *Rattus rattus* in Peninsular Malaysia Based on Mitochondrial and Nuclear DNA Data..... 91

4.2.1 Sequence analysis..... 91

4.2.2 Phylogenetic analysis..... 102

4.2.3 Population genetics analysis..... 105

4.3 Morphometric Analysis of Skull Characteristics of *Rattus rattus*……. 120

4.4 Prevalence and Genetic Divergence of *Leptospira* spp. in *Rattus rattus*……. 129

4.5 Transcriptome Study of *Rattus rattu*s Infected with *Leptospira* spp……. 143

4.5.1 Sequencing reads filtering…….. 143

4.5.2 Genome mapping…… 145

4.5.3 Novel transcripts prediction…… 146

4.5.4 SNP and INDEL detection……. 146

4.5.5 Differentially splicing gene detection……..... 149

4.5.6 Gene expression analysis…… 151

4.5.6.1 Gene mapping and expression……..... 151

4.5.6.2 Correlation between samples…….. 152

4.5.6.3 The distribution of gene expression…… 153

4.5.7 Differentially expressed gene detection…….. 156

4.5.8 Hierarchical clustering analysis of DEGs…… 158

4.5.9 Gene ontology analysis of DEGs……. 158

4.5.10 Pathway analysis of DEGs……..... 162

4.5.11 Quantitative real-time PCR (qRT-PCR) analysis……. 168

**CHAPTER 5: DISCUSSION..... 173**

5.1 Species Diversity of *Rattus* (Rodentia: Muridae) on the Malay Peninsula and Their Phylogenetic Relationships Inferred from Complete Mitochondrial Cytochrome b Gene Sequences..... 173

5.1.1 The number and identity of *Rattus* species in Peninsular Malaysia and the *Rattus rattus* complex……..……... 173

5.1.2 The disclosure of the *Rattus* spp..... 179

5.2 Population Structure of *Rattus rattus* in Peninsular Malaysia Based on Mitochondrial and Nuclear DNA Data……..…… 184

5.2.1 Mitochondrial and nuclear DNA as potential markers in elucidating the population structure of *Rattus rattus*..... 184

5.2.2 Sequence and phylogenetic analysis..... 186

5.2.3 Population genetics analysis..... 189

5.3 Morphometric Analysis of Skull Characteristics of *Rattus rattus*……. 194

5.4 Prevalence and Genetic Divergence of *Leptospira* spp. in *Rattus rattus*……... 198

5.5 Transcriptome Study of *Rattus rattu*s Infected with *Leptospira* spp…….. 202

**CHAPTER 6: CONCLUSION..... 211**

**REFERENCES..... 214**

**LIST OF PUBLICATIONS AND PAPERS PRESENTED..... 254**

**APPENDICES..... 255**

## LIST OF FIGURES


| | | | |
|---|---|---|---|
| Figure 3.1 | : | Sampling location of *Rattus* spp. throughout Peninsular Malaysia…….. | 44 |
| Figure 3.2 | : | Different age classes of *R. rattus* described on the basis of molar wear stages (M3)……..... | 59 |
| Figure 3.3 | : | Craniodental measurements of *R. rattus* based on (a) dorsal, (b) ventral, and (c) lateral views……..... | 61 |
| Figure 4.1 | : | Extraction results of genomic DNA from several samples of *Rattus rattus*, on 1% agarose gel……. | 79 |
| Figure 4.2 | : | Minimum-spanning network of genus *Rattus* haplogroup…….. | 83 |
| Figure 4.3 | : | Minimum-spanning network of *Rattus* spp. haplogroup with specific geographic origin of the haplotypes……... | 84 |
| Figure 4.4 | : | The neighbor-joining (NJ) phylogenetic tree of *Rattus* species estimated using the Kimura 2-parameter algorithm and 1000 bootstrap replications…….. | 88 |
| Figure 4.5 | : | Bayesian inference (BI) of the 50% majority rule consensus tree of the *Rattus* species…… | 89 |
| Figure 4.6 | : | Result of BEAST analysis for the *Rattus* species based on Cyt b gene…….. | 91 |
| Figure 4.7 | : | The minimum-spanning network (MSN) generated illustrating the relationships of *R. rattus* from different locality, based on mtDNA sequences…….. | 97 |
| Figure 4.8 | : | The minimum-spanning network (MSN) generated illustrating the relationships of *R. rattus* from different locality, based on nDNA sequences……. | 98 |
| Figure 4.9 | : | The neighbor-joining (NJ) phylogenetic tree of *R. rattus* species based on mtDNA sequences…… | 103 |
| Figure 4.10 | : | The neighbor-joining (NJ) phylogenetic tree of *R. rattus* species based on nDNA sequences…….. | 104 |
| Figure 4.11 | : | Mismatch distribution of pairwise nucleotide differences of mtDNA sequences of *R. rattus* with unimodal distribution……. | 116 |
| Figure 4.12 | : | Mismatch distribution of pairwise nucleotide differences of nDNA sequences of *R. rattus* with unimodal distribution…….... | 120 |
| Figure 4.13 | : | PC1 and PC2 of the PCA for 14 craniodental measurements of 67 male *R. rattus*……. | 124 |

Figure 4.14 : PC1 and PC2 of the PCA for 14 craniodental measurements of 63 female *R. rattus*…….. 124

Figure 4.15 : LDA plot of the 14 craniodental measurements of male *R. rattus*……... 125

Figure 4.16 : LDA plot of the 14 craniodental measurements of female *R. rattus*……... 127

Figure 4.17 : Detection of leptospiral DNA in kidney and liver of *Rattus rattus* by PCR…….. 129

Figure 4.18 : Prevalence of leptospiral infection in *R. rattus* collected in Peninsular Malaysia from December 2016 to February 2018….. 130

Figure 4.19 : Prevalence of *Leptospira* spp. circulating among *R. rattus*……. 131

Figure 4.20 : Prevalence of *Leptospira* spp. based on age group and sex of *R. rattus*……... 132

Figure 4.21 : Prevalence of *Leptospira* spp. according to cities in Peninsular Malaysia…….. 132

Figure 4.22 : Prevalence of *Leptospira* spp. according to niches…… 133

Figure 4.23 : Prevalence of *Leptospira* spp. in different organs…….. 134

Figure 4.24 : The neighbor-joining (NJ) phylogenetic tree of *Leptospira* spp. estimated using the Kimura 2-parameter algorithm and 1000 bootstrap replications, based on 16S rRNA gene…… 136

Figure 4.25 : The neighbor-joining (NJ) phylogenetic tree of *L. interrogans* and *L. borgpetersenii* estimated using the Kimura 2-parameter algorithm and 1000 bootstrap replications, based on LipL32 gene……. 137

Figure 4.26 : Minimum-spanning network of *L. interrogans* and *L. borgpetersenii* haplogroup defining 18 haplotypes, based on 16S rRNA gene……... 140

Figure 4.27 : Minimum-spanning network of *L. interrogans* and *L. borgpetersenii* haplogroup defining five haplotypes, based on LipL32 gene…… 140

Figure 4.28 : Result of RNA quality of *R. rattus* samples inspected using Agilent 2100 Bioanalyzer……... 143

Figure 4.29 : Raw data filter composition chart…… 144

Figure 4.30 : SNP variant type distribution…….. 147

Figure 4.31 : Distribution of SNP location……... 148

Figure 4.32 : Distribution of INDEL location…….. 149

Figure 4.33 : Statistic of splicing…….. 150

Figure 4.34 : Heatmap of Pearson correlation between samples…….. 152

Figure 4.35 : Hierarchical clustering between samples…… 153

Figure 4.36 : Gene expression distribution……... 154

Figure 4.37 : Statistics of novel genes and known genes…….. 155

Figure 4.38 : Expression distribution of novel and known genes……. 156

Figure 4.39 : Summary of DEGs…….. 157

Figure 4.40 : Scatter plot of DEGs……... 157

Figure 4.41 : Heatmap of hierarchical clustering of DEGs……... 158

Figure 4.42 : GO classification of DEGs…….. 159

Figure 4.43 : GO classification of upregulated and downregulated genes…… 161

Figure 4.44 : Pathway classification of DEGs…….. 163

Figure 4.45 : Pathway functional enrichment of DEGs…… 165

Figure 4.46 : Pathway functional enrichment result for up/down regulation genes……... 167

Figure 4.47 : Optimisation of qRT-PCR primers using PCR…… 169

Figure 4.48 : Comparison of the expression profiles in the kidney of *R. rattus* as determined by BGISEQ-500 sequencing and qRT-PCR……. 170

Figure 4.49 : Comparison of the expression profiles in the liver of *R. rattus* as determined by BGISEQ-500 sequencing and qRT-PCR……. 170

Figure 4.50 : Gene ontology (GO) classification of the six annotated genes… 171

Figure 4.51 : KEGG pathway classification for annotated genes……. 172

## LIST OF TABLES


| | | | |
|---|---|---|---|
| Table 3.1 | : | The localities and samples sizes (*N*) from which *R. rattus* populations were collected in Peninsular Malaysia…… | 46 |
| Table 3.2 | : | Details of primers used to deduce the population structure of *R. rattus* from Peninsular Malaysia…… | 54 |
| Table 3.3 | : | PCR profile used for amplification of mitochondrial and nuclear genes……... | 54 |
| Table 3.4 | : | PCR chemical components used for amplification of mitochondrial and nuclear genes…… | 54 |
| Table 3.5 | : | Details of primers used to presume the prevalence and genetic divergence of *Leptospira* spp. in *R. rattus*……... | 63 |
| Table 3.6 | : | PCR profile used for leptospiral DNA screening……. | 63 |
| Table 3.7 | : | PCR chemical components used for leptospiral DNA screening. | 63 |
| Table 3.8 | : | First strand cDNA synthesis components…… | 73 |
| Table 3.9 | : | Reverse transcription profile……... | 73 |
| Table 3.10 | : | Details of primers designed for qRT-PCR……... | 74 |
| Table 3.11 | : | PCR chemical components used for optimisation of designed primers…… | 75 |
| Table 3.12 | : | PCR profile used for optimisation of designed primers……... | 75 |
| Table 3.13 | : | The chemical components used in qRT-PCR amplification…… | 77 |
| Table 3.14 | : | Thermal profile of the SYBR Green real-time PCR…… | 77 |
| Table 4.1 | : | Correlation between the mtDNA lineage terminology and taxonomic usage in this study with that implemented in several previous studies…….. | 81 |
| Table 4.2 | : | Average percentages of pairwise genetic distances among *Rattus* species and *Leopoldamys sabanus* (Outgroup), based on the Kimura 2-Parameter model…….. | 86 |
| Table 4.3 | : | Nucleotide information of *R. rattus* based on mtDNA and nDNA genes…… | 93 |
| Table 4.4 | : | Pairwise distances between populations of *R. rattus* in Peninsular Malaysia based on mtDNA sequences……... | 94 |
| Table 4.5 | : | Pairwise distances between populations of *R. rattus* in | |

Peninsular Malaysia based on nDNA sequences……. 96

Table 4.6 : Measures of nucleotide diversity ($\pi$) and net nucleotide divergence (*Da*) among populations of *R. rattus*, based on mtDNA sequences…….. 100

Table 4.7 : Measures of nucleotide diversity ($\pi$) and net nucleotide divergence (*Da*) among populations of *R. rattus*, based on nDNA sequences…… 101

Table 4.8 : Measures of nucleotide subdivision (*Nst*) and population subdivision (*Fst*) among populations of *R. rattus*, based on mtDNA sequences…….. 106

Table 4.9 : Measures of gene flow (number of migrants, *Nm*) and geographical distance (KM) among populations of *R. rattus*, based on mtDNA sequences…… 107

Table 4.10 : Measures of nucleotide subdivision (*Nst*) and population subdivision (*Fst*) among populations of *R. rattus*, based on nDNA sequences…… 109

Table 4.11 : Measures of gene flow (number of migrants, *Nm*) and geographical distance (KM) among populations of *R. rattus*, based on nDNA sequences…….. 110

Table 4.12 : Analysis of molecular variance of *R. rattus* populations, based on mtDNA sequences…….. 112

Table 4.13 : Analysis of molecular variance of *R. rattus* populations, based on nDNA sequences…… 112

Table 4.14 : Summary statistic of mitochondrial DNA sequence variations in 13 populations of *R. rattus*……. 113

Table 4.15 : Historic demographic parameters of *R. rattus* based on mtDNA (COI, Cyt b, D-loop) gene…… 115

Table 4.16 : Summary statistic of nuclear DNA sequence variations in 13 populations of *R. rattus*……... 117

Table 4.17 : Historic demographic parameters of *R. rattus* based on nDNA (IRBP) gene…… 119

Table 4.18 : The 14 craniodental measurements (in mm) of the 67 samples of male *R. rattus* collected in Peninsular Malaysia…… 121

Table 4.19 : The 14 craniodental measurements (in mm) of the 63 samples of female *R. rattus* collected in Peninsular Malaysia……. 122

Table 4.20 : t-test for difference in mean of craniodental measurements between male and female *R. rattus*……. 123

Table 4.21 : Variable coefficients and percentage variation explained by the first three components (PC1-PC3) of the PCA on 14 standardised craniodental measurements from male and female *R. rattus*……... 126

Table 4.22 : Results of the 'a posteriori' Scheffé's test to evaluate the significant differences of craniodental measurements between ages C2, C3, and C4 for male and female *R. rattus*……... 128

Table 4.23 : Regressions of the different skull size parameters of male and female *R. rattus* in relation to the geographical factors in Peninsular Malaysia…… 129

Table 4.24 : Average percentages of pairwise genetic distances among *Leptospira* spp. and *Leptonema illini* (Outgroup), based on the 16S rRNA gene……... 138

Table 4.25 : Estimates of genetic diversity of *L. interrogans* and *L. borgpetersenii*, based on 16S rRNA and LipL32 genes……. 141

Table 4.26 : Summary statistics of genetic variability among the two leptospiral species (*L. interrogans* and *L. borgpetersenii*), based on 16S rRNA and LipL32 genes…….. 141

Table 4.27 : Results of analysis of molecular variance (AMOVA) testing genetic structure of *L. interrogans* and *L. borgpetersenii*, based on 16S rRNA gene…… 142

Table 4.28 : Results of analysis of molecular variance (AMOVA) testing genetic structure of *L. interrogans* and *L. borgpetersenii*, based on LipL32 gene……. 142

Table 4.29 : Clean reads quality metrics……. 145

Table 4.30 : Summary of genome mapping……. 145

Table 4.31 : Summary of novel transcripts…….. 146

Table 4.32 : SNP variant type summary…….. 147

Table 4.33 : Summary of gene mapping ratio……. 151

Table 4.34 : Genes and transcripts statistics…… 152

## LIST OF SYMBOLS AND ABBREVIATIONS

| | | |
|---|---|---|
| α | : | Alpha |
| β | : | Beta |
| γ | : | Gamma |
| µl | : | Microliter |
| µM | : | Micromolar |
| $\pi$ | : | Nucleotide Diversity |
| 16S rRNA | : | 16S Ribosomal RNA |
| A | : | Adenine |
| AIC | : | Akaike Information Criterion |
| AMOVA | : | Analysis of Molecular Variance |
| BI | : | Bayesian Inference |
| BLAST | : | Basic Local Alignment Search Tool |
| bp | : | Base pairs |
| C | : | Cytosine |
| cDNA | : | Complementary Deoxyribonucleic Acid |
| COI | : | Cytochrome Oxidase Subunit I |
| $C_t$ | : | Threshold Cycle |
| Cyt b | : | Cytochrome b |
| D-loop | : | Displacement Loop |
| DDBJ | : | DNA Data Bank of Japan |
| DEG | : | Differentially Expressed Gene |
| DNA | : | Deoxyribonucleic Acid |
| dNTP | : | Deoxynucleotide |
| e.g. | : | For example |

| | | |
|---|---|---|
| FDR | : | False Discovery Rate |
| FPKM | : | Fragments Per kb Per Million fragments |
| *Fst* | : | Estimate of Population Subdivision |
| G | : | Guanine |
| g | : | Gram |
| GBIF | : | Global Biodiversity Information Facility |
| GC | : | Guanine Cytosine |
| GO | : | Gene Ontology |
| *Hd* | : | Haplotype Diversity |
| IFN-γ | : | Gamma Interferon |
| IL-1β | : | Interleukin-1 Beta |
| IL-6 | : | Interleukin-6 |
| IL-10 | : | Interleukin-10 |
| IP-10 | : | Inducible Protein 10 kDa |
| IRBP | : | Interphotoreceptor Retinoid Binding Protein |
| *k* | : | Average Number of Nucleotide Differences |
| kb | : | Kilobases |
| kDa | : | Kilodalton |
| KEGG | : | Kyoto Encyclopedia of Genes and Genome |
| KM | : | Kilometre |
| K2P | : | Kimura 2-Parameter |
| LipL32 | : | The 32 kDa Lipoprotein |
| mA | : | MilliAmpere |
| MEGA | : | Molecular Evolutionary Genetics Analysis |
| mg | : | Milligram |
| $MgCl_2$ | : | Magnesium Chloride |

| | | |
|---|---|---|
| MIP-1α | : | Macrophage Inflammatory Protein-1 Alpha |
| ml | : | Millilitre |
| mm | : | Millimetre |
| mRNA | : | Messenger Ribonucleic Acid |
| MSN | : | Minimum-Spanning Network |
| mtDNA | : | Mitochondrial Deoxyribonucleic Acid |
| Mya | : | Million Years Ago |
| NCBI | : | National Centre for Biotechnology Information |
| nDNA | : | Nuclear Deoxyribonucleic Acid |
| ng | : | Nanogram |
| NGS | : | Next Generation Sequencing |
| NJ | : | Neighbor-Joining |
| nM | : | Nanomolar |
| *Nm* | : | Number of Migrants Per Generation |
| *Nst* | : | Nucleotide Subdivision |
| PCR | : | Polymerase Chain Reaction |
| qRT-PCR | : | Quantitative Real Time Polymerase Chain Reaction |
| RNA | : | Ribonucleic Acid |
| RNA-Seq | : | RNA Sequencing |
| *SD* | : | Standard Deviation |
| SNP | : | Single-Nucleotide Polymorphisms |
| *SSD* | : | Sum of Squared Deviation |
| T | : | Thymine |
| TNF-α | : | Tumor Necrosis Factor-Alpha |
| V | : | Voltage |
| WHO | : | World Health Organisation |

## LIST OF APPENDICES


Appendix A : The basic field measurements of the collected *Rattus* spp. in Peninsular Malaysia……... 255

Appendix B : List of genetic samples used to infer the phylogenetic relationships of genus *Rattus* in Peninsular Malaysia……. 256

Appendix C : List of reference sequences used to examine the prevalence of leptospiral infection on *R. rattus* in Peninsular Malaysia…… 261

Appendix D : Segregating sites (45 bp) in 2415 bp segment of concatenated mtDNA (COI, Cyt b, D-loop) sequences defining 91 haplotypes……... 262

Appendix E : Segregating sites (32 bp) in 1227 bp segment of nuclear DNA (IRBP) sequences defining 95 haplotypes…….. 266

Appendix F : Results of nucleotide BLAST showing similarity with 16S rRNA gene of *Leptospira* spp. from GenBank…… 270

Appendix G : Segregating sites of 16S rRNA and LipL32 sequences……... 273

Appendix H : Melting curve analysis and dissociation curve of MIP-1α and β-actin……. 274

Appendix I : Threshold cycle values and subsequent fold changes obtained in the expression of selected genes during leptospiral infection in the kidney……... 275

Appendix J : Threshold cycle values and subsequent fold changes obtained in the expression of selected genes during leptospiral infection in the liver……... 277

# CHAPTER 1: INTRODUCTION

## 1.1 General Introduction

Biogeographically, Malaysia is situated in the Oriental Region (Wallace, 1869) more often named the Indomalayan region (Leessmith et al., 1988; Udvardy, 1975) of the Palaeotropical division. Since a long time ago, it has been documented that the Thai-Malay Peninsula possess a transition zone connecting two zoogeographical areas (Woodruff & Turner, 2009; Hughes et al., 2003; Wallace, 1876). The Sundaic division comprises the Greater Sunda Islands and the Thai-Malay Peninsula extent to north as the Isthmus of Kra, while the Indochinese division includes ample of continental Southeast Asia and extents to south as the northern area of the Thai-Malay Peninsula (Rueda et al., 2013; Woodruff, 2010). These divisions were assumed to intersect in a transition zone, that approaches to the Isthmus of Kra, the smallest part of the Thai-Malay Peninsula bordered with an east-west division (Parnell, 2013). Dispersion and distributional boundaries proportional to this border of movement were illustrated in a series of organisms (Dejtaradol et al., 2016), even though the actual latitude of the transition line differs among diverse taxonomic groups. Due to the complexity of species distributional trends in this limited region, the Thai-Malay Peninsula was considered a biogeographical crossroads (Lohman et al., 2011).

The order Rodentia signifies approximately 42% of all mammal diversity with 2277 species (Musser & Carleton, 2005). Within Muridae family, the Rattini tribe comprises 35 genera equivalents to 167 rat species (Musser & Carleton, 2005) according to the tribal classification of the Murinae as suggested by Lecompte et al. (2008). The hyperdiverse rodent genus *Rattus* Fischer de Waldheim, 1803, comprising of at least 66 legitimate species may be the most diversified mammalian genus, is believed to have originated in tropical east Asia (Musser & Carleton, 2005). A few of these species are exceptionally

good colonisers and are invasive throughout the world in all different kinds of habitats (Lack et al., 2012; Chinen et al., 2005; Aplin et al., 2003; De Graaff, 1981).

Members of this worldwide genus are perhaps known as most destructive of all animal pests (Aplin et al., 2011) for their involvement in causing serious economic damage. Although wild *Rattus* populations can be beneficial to people, their negative impacts overweigh the advantages that they offer (Aplin et al., 2003; Twig, 1975). *Rattus* species always cause harm to agricultural plants, hence represent a predominant problem to the farmers, as their crop yields facing decreased in amount of productions (Singleton, 2003). Apart from that, rodents, mainly the ones belonging to the genus *Rattus* are considered as huge pest species that usually live in human habitations, who prefer to live inside human houses while residing off discarded human waste material (Begon et al., 2003; Singleton et al., 2003).

Mainly found in urban area, coastal habitat, and paddy field (Paramasvaran et al., 2013), *Rattus rattus* has a wide distribution (Long, 2003) and recorded among the worst invasive species globally. The species is normally known as either the house rat, black rat, ship rat, or roof rat (Musser & Carleton, 2005; Meester et al., 1986; De Graaff, 1981; Meester & Setzer, 1971). The distribution and identification of the black rat has been difficult to pin down. The emergences of the black rats are also intensely argued. Nowak (1999) suggested that this species possibly native to the Malaysian region, while Musser & Carleton (2005) specified that these rats are originated in the Indian Peninsula, since which they were dispersed worldwide. As a successful, highly adaptive, and dominant species in most regions in the world, *R. rattus* has the capability to adapt to a new area rapidly (Singleton et al., 2003). This species not only live well in forest and woodlands, it also adapts to live in and around human settlements (Clark, 1981). Rodent distributions

do depend heavily on its preferable diet and as for *R. rattus*, it feeds on almost all edible things and that explain why this species is so widespread. *Rattus rattus* has the ability to conquer mostly all terrestrial habitat types on both islands and continents.

There are huge interests in research on rodent behaviour, reproduction, and associations to human disease but not much study had been done on the morphology and genetic. The information of Murinae in Malaysia is still lacking even though several studies have been done using morphology, genetic, immunology, albumin, and karyotypic analyses (Tamrin & Abdullah, 2011). An importance or maybe a reason of this taxonomic doubt, is that several rats are tough to identify morphologically even from whole carcasses. Besides, several species of *Rattus* are difficult to identify even when the entire animal is accessible (Robins et al., 2007). Thus, a complex taxonomy of the genus *Rattus* has difficulties in morphology-based identifications. Therefore, identification of *Rattus* should be done not only through morphology but also through molecular. This is due to variation of morphological traits as it has the high rate of evolution and a rapid adaptation toward their ecological habitats in Murinae (Tamrin & Abdullah, 2011).

Aplin et al. (2011) successfully discovered *R. rattus* complex based on molecular evidence, and the phylogenetic restructuring to classify this species is still ongoing. There are actually many different species within what had been identified as *R. rattus* in the past (Pagès et al., 2010; Robins et al., 2007) including *Rattus tanezumi* Temminck, 1844 (syn. *R. diardii*) (Musser & Carleton, 2005). It is almost impossible to distinguish *R. rattus* and *R. tanezumi* based only on morphological characteristics (Aplin et al., 2003). Thus, it is clear that after the evolution of molecular analysis, the taxonomic status of many species can be re-evaluated. Without molecular advancement, it is very difficult to discriminate the species.

With the evolution of molecular approaches in investigating various species of rodents, there were many recent studies regarding new molecular methods that have been used with greater success, for example restriction fragment length polymorphism (Michaux et al., 1998; Chelomina, 1996), random amplified polymorphic DNA (RAPD) (Bellinvia et al., 1999), and sequencing of mitochondrial and/or nuclear genes (Martin et al., 2000; Serizawa et al., 2000; Chelomina, 1998; Chelomina et al., 1998). Molecular phylogenetic analysis are the best solutions to solve the problem of identification and there are recent studies have been done to elucidate the members and structure of Muroidea using slowly evolving nuclear genes (Michaux et al., 2001; Jansa & Weksler, 2004; Steppan et al., 2004). Deoxyribonucleic acid (DNA)-based methods of documentation have verified to discriminate cryptic species when genetic lineages have deviated but morphology has not (Aplin et al., 2011; Hebert et al., 2004; Wada et al., 2003).

Besides, rats are also currently known as main hosts and carriers of various zoonotic diseases mainly derived from bacteria, viruses, and parasites (Buckle & Smith, 2015). Another important matter related to rodents is the ability of those animals to serve as zoonotic vectors which actively transmitting and spreading of numerous infectious agents to humans (Khaghani, 2007). The *Rattus* spp. have also served as zoonotic vectors which may transmit diseases (Kosoy et al., 2015; Morand et al., 2015; Meerburg et al., 2009; Aplin et al., 2003; Belmain et al., 2003; De Graaff, 1981). In Malaysia, rodents belonging to the genus *Rattus* are recognised to transmit diseases such as leptospirosis, hantavirus, typhus, plague, and salmonellosis, with leptospirosis being of principal focus, as it is an emerging infectious disease worldwide (Taylor et al., 2008; LeclercMadlala & Janowski, 2004).

Leptospirosis, an emergent zoonotic disease with global circulation is triggered by spirochetes belonging to the genus *Leptospira* (Levett, 2001). Over 500000 incidents of acute leptospirosis are announced annually with more than 10% reports of fatality (WHO, 2010). Leptospires can endure for weeks in proper moist environments before bump into a new host (Grassmann et al., 2015). Reservoir hosts, typically rodents, show slight to no symptoms of infection but shed huge numbers of leptospires in their urine (Caimano et al., 2014). Transference happens when mucosal surfaces or abraded skin exposed with infected urine or urine-contaminated water or soil (Athanazio et al., 2008). Leptospires can create a variability of clinical signs in humans, varying from asymptomatic or slight fever to acute icteric (Weil's) syndrome and pulmonary bleeding (Evangelista & Coburn, 2010).

Presently, it is still unclear how leptospires survive within a reservoir host (Nally et al., 2017). Previous in vitro works have proposed that leptospires adjust their transcriptomic and proteomic profiles due to environmental signs faced during mammalian infection. However, there is no report on the gene expression of the mammalian host attacked by the leptospires (Nally et al., 2017; Caimano et al., 2014). Host-pathogen interactions can upshot in acute and adaptive immune responses counteracting to the attacker. Gaining information concerning host-pathogen interactions is essential to reveal the pathogenesis of a specific infection. Nowadays, high-throughput RNA sequencing (RNA-Seq) has superseded microarrays as a way of option for genome-wide transcriptional profiling (Filiatrault, 2011; Croucher & Thomson, 2010). RNA-Seq permits transcription to be discovered at the single nucleotide level compared to microarrays ('t Hoen et al., 2013). Studying the global transcriptome is important for understanding the molecular mechanisms associated with rats' responses to pathogen attack. RNA-Seq approach can be used to understand the transcriptome profile of rats

infected with *Leptospira* spp. The comparative expression levels of genes differentially expressed by leptospiral infection within the mammalian host under both healthy and infected conditions can be identified via this method (Nally et al., 2017; Caimano et al., 2014).

## 1.2 Problem Statements

The taxonomy of the tribe Rattini that consist of genus *Rattus* is arguably among the most complex and vague, inclusive of cryptic and sympatric species, notwithstanding being genetically defined (Aplin et al., 2011; Pagès et al., 2010; Robins et al., 2007; Musser & Carleton, 2005; Aplin et al., 2003). Their species status boundaries are still problematic and debated. The genus *Rattus* contains a large number of species, that are often difficult to tell apart morphologically. This has obscured the documentation of species distributions and has generated many instances of uncertain taxonomy. The situation is particularly difficult in tropical East Asia, the centre of their diversification and home to high diversity. The specific identification and number of species present in Peninsular Malaysia is not clear. The chance for the individuals to be misidentified was high if rat species were just recorded based on outcome from new areas, with minute or no deliberation to investigate further particulars on morphological features and molecular data. Introduced species also is supposed to create harm to the environment, human economy, or human health; with a tendency to spread to a degree depending on their adaptations to the environment. Thus, the population structure at the intraspecific level could reveal the genetic diversity of a species and give information regarding their survivability.

Moreover, rodents which are recognised as hosts for various zoonotic diseases inclusive of leptospirosis, may implies a significant risk for human health. Since *R. rattus*

are mainly known as complex species and prominent in transmission of leptospirosis, the prevalence study of the disease is very needful in order to infer the resistance and susceptibility of the species towards leptospiral pathogens. It is quite complicated to know that rats can survive during leptospirosis infection in contrast to human. Right now, there is restricted indication on the immune-related genes of wild rats which are principally detailed in leptospiral infection. The mechanisms involved in the immune response of wild rats to bacterial infection are not completely understood. Different organs might react differently during infection. It is invalid to conclude the phase of infection solely depends on a single organ. The fluctuation of gene expression levels at different organs during a healthy and unhealthy state might disclose the stage of infection of the host. Thus, the host-pathogen interactions through transcriptomic approach are crucial to be explored in order to mitigate the spreading of leptospirosis.

### 1.3 Significant of the Research

Rodents are recognised as main vectors of bacteria and pathogens where some of them triggering zoonotic diseases and symbolising a serious hazard to the human health worldwide. It is expected that the rat species in charge of disease transmission are those that living near human. A precise identification of rats at a specific level by using both morphology and molecular approaches is important to comprehend the function of each rodent species in the whole host-pathogen community in anticipating the risk of emergence or re-emergence infectious diseases. Sorting of the species through morphological of external and skull features and considering inspection of dentition characteristics of rats could offer an important outcome to infer that the samples were composed of probably diverse species. This study attempted to deliver some knowledge for genetically delimitated species based on assessment of presently existing molecular genetics and morphological data. Nonetheless, this study may be valuable to show some

significant taxonomical queries and lead to proper taxonomic names usage, from the basic background of genetic and morphological variability revealed by populations of *R. rattus* found in Peninsular Malaysia.

Since rodents are occupying various environments, thus rodents-borne pathogens relationships may vary due to the different or changing environment and this likely increase the probability of zoonotic transfer of infections. With the advancement of genomics, the functional role of genes can be determined quickly and its importance can be used for further research in host-pathogen relationships. Functional genomics approach is used to determine the functional genes; to understand the adaptation of those genes in context to disease susceptibility and resistance in a vertebrate model. Therefore, transcriptomic approaches are being applied to generate and compare the profiles of the normal and infected rodent to clarify the relationship between the differentially expressed genes in the antibacterial immunological context more comprehensively. In the present study, high-throughput deep sequencing technology is crucial in investigating the transcriptome and relative expression profiles of the *R. rattus* infected with *Leptospira* spp. to understand the molecular mechanisms of the species' immune response towards the pathogenic bacteria. These studies will formulate a conceptual framework based on relationship between host and pathogen which can lead to formation and documentation of the disease patterns.

## 1.4 Research Questions

i) How many *Rattus* species present in Peninsular Malaysia and do *Rattus rattus* sensu stricto is one of them?

ii) How is the population structure of *R. rattus* in Peninsular Malaysia based on morphological and molecular data?

iii) How is the prevalence pattern of leptospiral infection on *R. rattus* in Peninsular Malaysia?

iv) Do the differentially expressed genes (DEGs) level of selected cytokines and chemokines during leptospiral infections in kidney and liver of *R. rattus*, related to the progression of leptospirosis?

## 1.5 Research Objectives

This research aims to:

i) resolve the complexities within *Rattus rattus* in Peninsular Malaysia using molecular approach,

ii) determine the population structure of *R. rattus* in Peninsular Malaysia based on morphological and molecular data,

iii) investigate the prevalence patterns of *Leptospira* spp. infecting *R. rattus* in Peninsular Malaysia,

iv) examine the differentially expressed genes (DEGs) of selected cytokines and chemokines during leptospiral infections in kidney and liver of *R. rattus*.

# CHAPTER 2: LITERATURE REVIEW

## 2.1 Order Rodentia and Its Diversity

Rodentia is the most speciose mammalian order and comprises nearly half of all mammalian species diversity (Wilson & Reeder, 2005). Geographically widespread and vastly flexible, rodents inhabit a huge range of various ecological niches. Murinae is the highly speciose subfamily, comprises 561 species and 126 genera, within the largest family, Muridae (Musser & Carleton, 2005). Its natural distribution includes all continents except Antartica and remote islands (Buckle, 1994). The Rattini tribe incorporates 35 genera relating to 167 rodent species among Muridae rodents (Musser & Carleton, 2005). Roughly all representatives of Rattini tribe can be found in Southeast Asia, a hotspot of biodiversity (Myers et al., 2000) that confronted with a runway monetary development harming natural surroundings and health, but also a hotspot of emerging and re-emerging diseases (Forman et al., 2008; Matsui, 2005). The rodent genus *Rattus* Fischer de Waldheim, 1803, covering of at least 66 valid species is supposed to have originated in Southeast Asia and it is evaluated to be the most expanded mammalian genus (Musser & Carleton, 2005).

Species diversity of *Rattus* is particularly high in tropical east Asia, the centre of origin for this group (Chaimanee & Jaeger, 2001). Within tropical east Asia, there are two zoogeographic areas: Indochina and Sundaland, which are divided by the Isthmus of Kra on the Malaysian Peninsula. Most rodents and other mammal species are distributed on only one side of this biogeographic break (Woodruff & Turner, 2009; Francis, 2008). The members of the genus *Rattus* have radiated recently and share a conserved body plan (Rowe et al., 2011). External characters are occasionally misleading with regard to species identity, so morphological identification can be difficult (Tamrin & Abdullah, 2011; Robins et al., 2007; Musser, 1981). A precise identification of *Rattus* to the specific

level is vital in various discipline inclusive of community ecology research, the strategy and application of both conservation and pest management programmes, and the study of zoonotic disease cycles. Many latest findings have confirmed that a significant number of biological variations remain unseen through phenotypic similarity in a cryptic species. This scenario led to a confused taxonomy including excessive synonyms and species complexes. Despite the public health importance and inclusion of the intensively studied laboratory rat, the genus *Rattus* remains poorly studied in their native range. A number of taxonomic issues remain unresolved. Therefore, most of the samples were classified as a species complex.

For this reason, substantial effort has gone into developing molecular barcoding methods to differentiate the species in this genus (Li et al., 2015; Galan et al., 2012; Lu et al., 2012; Aplin et al., 2011; Robins et al., 2007). These studies have uncovered numerous deep lineages not associated with traditional scientific names, especially in the *Rattus rattus* species complex (Aplin et al., 2011; Pagès et al., 2010; Robins et al., 2010; Robins et al., 2008). The *R. rattus* complex is the most problematic groups within the genus. It has been hypothesised that the widely dispersed and comparatively common black rat, *R. rattus* represents a complicated of multiple species (Aplin et al., 2011). The number of species at particular localities and the distribution of species/lineages is not clear in many cases (Bastos et al., 2011; Pagès et al., 2011). Studies described that cryptic species within rodents may serve as important reservoir hosts in the circulation of diseases.

## 2.2 The Black Rat (*Rattus rattus*)

*Rattus rattus* Linnaeus, 1758 is a medium sized rat, approximately 225 to 445 millimetres (mm) with relatively large ears and almost a hairless tail that is always longer

than the body (Long, 2003; De Graaff, 1981). They practice this long tail to support in balance due to their nature as an avid climber that often inhabits an arboreal habitat (Grzimek, 2003; Nowak, 1999). The rats weigh between 100 and 200 gram (g), and are between 105 and 215 mm in head and body length, with a tail length of 120 to 230 mm (Francis, 2008). Males are usually longer and heavier compared to their female counterparts. This species is typically brown-black to black on the upper parts in colour with a slightly grey to white on the under parts. This species is regularly classified into subspecies relying on the colour forms which can arise in any combination of black, white, grey, and agouti. The skull and nasal bones of *R. rattus* are relatively narrow. This species has a fine covering of hair and a light skull (Grzimek, 2003).

*Rattus rattus* is discovered on every continents of the earth, even though the species is supposed to be indigenous to India and perhaps other Indo-Malayan realms. It has since been introduced worldwide (Aplin et al., 2011; Pagès et al., 2011; Matisoo-Smith & Robins, 2009; Hingston et al., 2005; Musser & Carleton, 2005; Long, 2003; De Graaff, 1981). Successfully inhabits in Peninsular Malaysia, black rat is so widespread and can be easily found in all states. The spread of the black rat which are commensal organisms by nature, has been mainly attributed to recent human migrations, trade, and the higher acclimatisation rates of these species compared with other species in this genus (Chinen et al., 2005, Aplin et al., 2003; Yiğit et al., 1998; De Graaff, 1981). It is often found near to human habitations and coastal parts as well as on large ships. Due to that, these rats are usually named ship rats.

## 2.3 Cytogenetic Analyses on *Rattus rattus* Complex

*Rattus rattus* was initially characterised as a species that comprises four major chromosomal groups; Asian group, Oceanian, Ceylonese, and Mauritian (Yosida et al.,

1985). The Asian group is characterised by a diploid number of 2n = 42 and is considered by Yosida et al. (1985) as being the basic *R. rattus* karyotype. This group has recently been elevated from the subspecies *R. rattus tanezumi* to the species *R. tanezumi*. This change in taxonomic status is primarily based on biochemical and cytogenetic differences between *R. rattus* and *R. tanezumi*, yet, specimens that have this karyotype are still considered as Asian type black rats (Robins et al., 2007; Chinen et al., 2005; Hingston et al., 2005; Musser & Carleton, 2005; Long, 2003; Baverstock et al., 1983). The Asian group is defined by a karyotype that has 13 acrocentric chromosome pairs, seven metacentric chromosome pairs, and acrocentric X and Y (found only in the male karyotype) or a pair of acrocentric X chromosomes (found only in the female karyotype) (Yosida et al., 1985). The Asian type have high C-banding and the Japanese type have low C-banding (*R. tanezumi*). The Oceanian group is characterised by having a diploid number of 2n = 38 (Matisso-Smith & Robins, 2009; Robins et al., 2007; Chinen et al., 2005; Hingston et al., 2005, Musser & Carleton, 2005; Yosida et al., 1985; Baverstock et al., 1983) and normally has nine acrocentric pairs, seven metacentric pairs, M1 and M2 marker chromosomes derived from the Robertsonian fusion of pairs 4 and 7, and 11 and 12, respectively, and an acrocentric X and Y or pair of X chromosomes (Yosida et al., 1985).

The Ceylonese group is characterised by a diploid number of 2n = 40 and is considered as being an intermediate *R. rattus* karyotype (Chinen et al., 2005; Musser & Carleton, 2005; Yosida et al., 1985; Baverstock et al., 1983). This karyotype is comprised of nine acrocentric pairs, eight metacentric pairs, one metacentric M2 marker chromosome derived from the Robertsonian fusion of pairs 11 and 12, and an acrocentric X and Y or pair of X chromosomes (Yosida et al., 1985; Baverstock et al., 1983). Studies by Baverstock et al. (1983), Yosida (1980, 1979, 1978, 1977), Yosida et al. (1974), and

Yosida et al. (1971) propose two chromosomal forms of *R. rattus* Linnaeus, 1758 arise in Sri Lanka, 2n = 40 form on the central highlands and 2n = 38 form on the coastal lowlands with an infrequent 2n = 39 form demonstrating hybridisation. The Mauritian group is characterised by a distinct 2n = 42 karyotype (Matisso-Smith & Robins, 2009; Chinen et al., 2005; Baverstock et al., 1983), which comprised nine acrocentric pairs, M1 and M2 marker chromosomes derived from the Robertsonian fusion of pairs 4 and 7, and 11 and 12, five metacentric pairs, and an acrocentric X and Y or pair of X chromosomes. The distinctive trait observable in this karyotype is that it also contains four small acrocentric pairs that are derived from the Robersonian fission of two metacentric pairs, namely metacentric pairs 14 and 18 (Yosida et al., 1985).

Within the *R. rattus* complex, observations indicate that a pattern of chromosome polymorphism exists, and that this polymorphism is either derived from one of the karyomorphs of *R. rattus* or differentiated progressively from that of the black rat (Yosida et al., 1985; Baverstock et al., 1983). It is commonly accepted that the basic karyotype of *R. rattus* has a diploid number of 2n = 42 and comprises 13 pairs of acrocentric chromosomes, seven metacentric pairs, and an acrocentric pair of XY or XX sex chromosomes (Chinen et al., 2005; Yosida et al., 1985; Baverstock et al., 1983). Variations to this basic karyotype are due to either Robertsonian fusions or fissions (Yosida et al., 1985).

## 2.4 The Taxonomy of the *Rattus* Complex and Problems Encountered During Species Classification

The classification of the genus *Rattus*, since its initial definition in 1803, has over the years presented a number of problems, especially regarding subdivisions of the genus (Musser & Carleton, 2005; De Graaff, 1981). Initial problems emerged when the great

majority of rodents that were obviously either rats or mice were referred to the genus *Mus*. This policy continued until 1881, when the erection of the sub-genus *Epimys* was proposed to cover the more typical rat, as opposed to mice (De Graaff, 1981). With the implementation of this sub-genus, only 27 forms were left in *Mus*, while 140 forms were placed in *Epimys*. However, in 1916 the subgenus *Epimys* was dropped and all members placed within the genus *Rattus*, as a means of differentiating rats from mice (De Graaff, 1981). Even though the genus *Rattus* had been in existence for a number of years and had provided a better morphological description of what a typical rat was, its taxonomic priority was overlooked (De Graaff, 1981). At one point the genus *Rattus* was reviewed to comprise of up to 550 different forms. Over the years, a number of species that were initially allocated to *Rattus* have been reallocated to genera such as *Maxomys*, *Berylmys*, *Leopoldamys*, and *Niviventer*. However, even though several species have been removed from the genus *Rattus*, the boundaries of this genus have not finally been decided (Musser & Carleton, 2005; Aplin et al., 2003). *Rattus rattus* appears to have numerous close relatives and is frequently unclear with other Asian *Rattus* species, thus making field identification highly questionable (Musser & Carleton, 2005; Aplin et al., 2003).

Other problems with the taxonomy of the genus *Rattus* are related to the fact that several groups of the species may still need to be removed from this genus (Musser & Carleton, 2005; Aplin et al., 2003). Evolutionary relationships of *Rattus* to other murid genera also left somewhat unknowable, and this problem is further compounded by the lack of attention given to this genus by taxonomists and geneticists (Chinen et al., 2005; Musser & Carleton, 2005; Aplin et al., 2003). Taxonomic revisions within the genus *Rattus* are essential, as other species that could belong to this genus may be recognised once more detailed morphological and genetic studies have been conducted. Even though several authors have strived to discover the morphological complexity among black rats,

these challenges have been largely regionally-based, with authors concentrating too much on the belly coat colour, even though this characteristic is greatly polymorphic within populations and possibly due to the genetic inheritance (Musser & Carleton, 2005; Aplin et al., 2003).

The most important discovery in understanding the *R. rattus* complex initiated with Yosida's works, from 1980 onwards, of chromosomal discrepancy (Aplin et al., 2003). He recorded 11 chromosomal forms that divide into five main clusters, namely (i) Asian black rats with 2n = 42 chromosomes, (ii) Japanese black rats with 2n = 42 chromosomes, (iii) Ceylonese black rats with 2n = 40 chromosomes, (iv) Oceanian or European black rats with 2n = 38 chromosomes, and (v) Mauritius black rats with 2n = 42, where the notation 2n refers to the diploid chromosome number (Aplin et al., 2003). In more recent taxonomic revisions by Musser & Carleton (2005), Yosida's five recognised *R. rattus* groups have been reduced to only two basic population groups, namely the Oceanian or European type having a chromosome number of 2n = 38/40, which has been assigned the name *R. rattus* by Musser & Carleton (2005), and the Asian type black rats having a chromosome number of 2n = 42, which has been assigned the name *R. tanezumi* by Musser & Carleton (2005) (Matisoo-Smith & Robins, 2009; Robins et al., 2007; Hingston et al., 2005). It is clear that further genetic, cytogenetic, and morphometric studies are still required in order to completely understand the complexity of the relationships between the various species that make up the *R. rattus* complex.

During the last century, many studies were focused on investigating the biodiversity of the family Muridae in Southeast Asia. The taxonomy of the family Muridae was revised many times during this period (Musser & Carleton, 2005, 1993; Pavlinov, 2005; Nowak, 1999; Ellerman & Morrison-Scott, 1966; Ellerman, 1949) but until now still could

not be considered as finally established. The same is true for opinions on the correspondence of forms and species and their natural ranges and ecologies. Traditionally, taxonomic status of Murinae was based on morphological characteristics. However, the taxonomic status of Murinae is poorly resolved as closely-related species in the subfamily Murinae can be morphologically similar to each other, so it is very tough to distinguish morphologically at a species and the extensive range of intraspecific level (Pimsai et al., 2014). Species belonging to the genus *Rattus* are also usually considered to be a group difficult to identify, especially in Southeast Asia and the Sunda Islands. The difficulty is caused by a similarity in general appearance, external morphology, and skull structure between many of the species and their considerable geographical and ecological variability. This difficulty in identification led to problems with species identification and sometimes gives rise to conflicting views on the identification of taxa in the fauna of the region. Morphological variation of this group has triggered in vague taxonomy with an excessive synonym (i.e. 83 synonyms for *R. rattus* and 41 for *R. norvegicus*) (Musser & Carleton, 2005).

In furtherance, different authors have significantly disagreed on the species composition of the fauna of continental Indochina. Van Peenen et al. (1969) indicated seven species within the genus *Rattus* for the region (*norvegicus*, *exulans*, *exiguus*, *rattus*, *slaideni*, *nitidus*, and *argentiventer*). Musser (1986), followed by Pavlinov (2005), in his most complete review of the mammals of Southeast Asia, stated the genus *Rattus* as also comprising seven species but in a different assemblage (*norvegicus*, *exulans*, *rattus*, *nitidus*, *argentiventer*, *remotus*, and *osgoodi*). Lunde & Son (2001) agreed with Musser's opinion on *Rattus* composition, with the addition of *R. tanezumi* considered as a separate species. Kuznetsov (2006) argued that the genus *Rattus* is even more complex and comprised ten species at total (*norvegicus*, *exulans*, *losea*, *rattus*, *koratensis*, *nitidus*,

*argentiventer*, *remotus*, *osgoodi*, and *moliculus*). If synonyms are taken into consideration, it follows that *R. exiguous* is a synonym of *R. losea* Swinhoe, 1870 and *R. sladeni* is a synonym of *R. tanezumi*, which is in accordance with Musser & Carleton (1993) (or the synonym for *R. koratensis* in agreement with Kuznetsov (2006)). Generally, it may be proposed that the number of *Rattus* species in Indochina is in the limits from seven to ten.

The phylogenetic relationships of the *Rattus* spp. are still lacking in Peninsular Malaysia. The species in the genus *Rattus* reported to be on the Malaysian Peninsula (south of the Isthmus of Kra) by Francis (2008) are the house rat/black rat (*Rattus rattus*), Malaysian wood rat (*R. tiomanicus*), the Pacific rat (*R. exulans*), the ricefield rat (*R. argentiventer*), the Norway rat/brown rat (*R. norvegicus*), and Annandale's rat (*R. annandalei*). The Annandale's rat (*R. annandalei*) has been moved to the closely related genus *Sundamys* and so is now *Sundamys annandalei* (Camacho-Sanchez et al., 2017), leaving five species of *Rattus* on the peninsula. Aplin et al. (2011) suggested that the *Rattus rattus* sensu lato present north and south of the Isthmus of Kra were Malaysian wood rats (*Rattus tiomanicus*) which they referred to as lineages *R. rattus* IV and VI, respectively. This suggests that there are not *Rattus rattus* sensu stricto on the Malay Peninsula, which would bring the number of species to four. Two of those four are widely invasive (*R. exulans* and *R. norvegicus*). The ricefield rat (*R. argentiventer*) is not currently considered invasive, but it does have a disjunct distribution north and south of the Isthmus of Kra. Global Biodiversity Information Facility (GBIF) also reports records of the Asian house rat (*Rattus tanezumi*) and a single record of *R. fuscipes* from the Malaysian Peninsula. The Asian house rat is also widely invasive. It is relevant to note that approximately half of the records for the genus *Rattus* in GBIF for the Malaysian Peninsula are for *Rattus* spp.

In the natural range of *Rattus* species which is Indochina, a robust phylogeographic revisions of black rats to date have been defined and some distribution within the region has also been recognised (Aplin et al., 2011). Aplin et al. (2011) presented the genetic examination of the settlement of mainland Asia by *R. rattus* using mitochondrial sequences. Fortunately, the great genetic and morphological diversity which was ascribed to the species *R. rattus* until a few decades ago now appears more reliable with roughly four to six different species of the *R. rattus* complex (Rrc), with six lineages (I-VI). *Rattus rattus* lineage I dispersing in India, Africa, America, Europe, and Australia while *R. rattus* lineage IV covered remote areas in East Asia, distribute on the Indochina peninsula (Aplin et al., 2011) via Indonesia and Philippines. However, lineage IV was not further distributed towards the south to Australia and Papua New Guinea where *R. rattus* lineage I was noticed. Phylogenetic investigation by Aplin et al. (2011) specified that these two lineages were principally allopatric and developing in isolation. Their revisions have involved specimens from numerous geographical regions in Asia but with the absence of specimens from Peninsular Malaysia. However, they postulated that the lineage that was hypothesised to be distributed across Peninsular Malaysia (*R. rattus cf*. lineage IV sensu Aplin et al., 2011) is not the *Rattus rattus* sensu stricto, apparently native of India. The current situation in the taxonomy of the group creates problems for ecologists, medical zoologists, and other specialists in zoology and natural conservation.

## 2.5 Address on Classification Using Morphometrics Tool

One of the elemental scopes of study is morphometrics, also known as the analysis, interpretation of shape, shape variation, and quantitative description in biology. Most systematic study that is based on the morphology of organisms need these techniques of description and comparison of shapes of structures (Rohlf, 1990). Ecological and genetic studies do involve with measurements of morphological diversity. Morphometric

techniques are important for any research field that depends on comparative morphology. The fields include systematic and evolutionary biology, biostratigraphy, and developmental biology. Few of the research in developmental biology include the studies of growth patterns within species, modularity, and evolutionary patterns (Gilbert, 2003). From palaeontology to quantitative genetics, morphometric is used in a broad range. The field of morphometrics is commonly divided into traditional or conventional morphometrics and geometric morphometrics (Rohlf & Marcus, 1993). Generally, morphometrics is the measurement of form, and it can be understood as shape, structure, and some other external appearance.

### 2.5.1 Conventional morphometrics in identification of rodents

From the earlier time, all species were discovered and classified via external morphology and skull measurements through conventional morphometric method. These methods were first used by scientist to study taxonomy for a very long time before molecular study evolved. Widely used, morphometric analysis has provided the methodological basis for geographic variation study in phenotypic morphological traits. In mammals and throughout rat species, the difference in configuration is notably found in the head and the change in the shape of the head is commonly resolved with the shape of the cranium (Odigie et al., 2017). Most examinations of non-geographic variation in small mammals are just inspected at the degree of sexual dimorphism and age variation (Abdel-Rahman, 2005). To describe the barriers of independent evolutionary units in nature, the records of geographic variant of the cranium in terms of climatic and habitat variables is vital (Renaud & Millien, 2001). Prior investigations of geographic variation in rodents described divergence in morphological features (Ben Faleh et al., 2013, 2012; Jansa et al., 2008; Bronner et al., 2007; Monteiro et al., 2003; Ventura & Lopez-Fuster, 2000).

Conventional morphometric is an efficient and cost-effective tool to be used in identifying subspecies and individuals at the population level (Francoy et al., 2008). Conventional morphometric instruments have been successful in segregating between sister species (Ben Faleh et al., 2010). For instance, Darvish et al. (2008) portrayed another types of jerboa, *Allactaga toussi* in the north-east of Iran utilising conventional morphometric information of the skull. A study by Zarei and colleagues (Zarei et al., 2013) displayed that analysis of size applying conventional methods have suggested a subspecific separation. However, Allaya et al. (2016) recommended that, in spite of the fact that populations can be recognised through a multivariate assessment of skull linear measurements, this does not consistently demonstrate the occurrence of a new different species/subspecies. So, the variation in morphology probably influenced by the local environmental factors.

Morphometric methods have been appeared to emphasis on the investigation of skull variability among rodent species (Bohoussou et al., 2014; Ben Faleh et al., 2012; Bezerra & De Oliveira, 2010; Lalis et al., 2009; Nicolas et al., 2008; Chimimba, 2001). Yet, such recent studies have generally targeted continental forms. Little is known about how rodent morphology varies biogeographically in Southeast Asia, an area composed of numerous peninsulas/islands and that is indisputably influenced by sea-level fluctuations. Introduction of non-native species is the most frequently debated issue in rat rapid expansion (Pergams et al., 2015; Pergams & Kareiva, 2009; Pergams & Lacy, 2007; Pergams et al., 2003; Pergams & Ashley, 2001; Patton et al., 1975). Pergams & Ashley (2001) conducted analysis of rapid morphological variation in *R. rattus* following introduction to the Galapagos Islands (Patton et al., 1975) by comparing cranial and skeletal characters as well as body size. The authors prove that development of total body size and both cranial and skeletal characters are larger on smaller and more isolated

islands. So far, the morphometric studies have never been performed on the craniodental of *R. rattus* in Peninsular Malaysia.

Taxonomic and evolutionary status of *R. rattus* remains poorly understood. Genus *Rattus* has vast morphological diversification and geographically widespread, thus they are incompletely described (Musser, 1981). This species is frequently involved in species complex. Individuals within *R. rattus* complex are hard to be separated using traditional morphometric due to the similarity of its morphology. There is no subspecies nominated within *R. rattus* by Musser & Carleton (2005), however previous taxonomist such as Ellerman (1941) does. Hence, various problems on the species identification of rats can be ascribed to a lack of great morphological characters and complexities. That is why many scientists in the past had the problem to identify species based on mere measurements of the body or skull and end up creating a group of the species complex. Through the discoveries of the species complex, conventional morphometric which only look on the physical measurements has its limitations to identify a species or to classify individuals into species group.

## 2.6 Molecular Analysis as a Powerful Complementary Tool

In order to resolve the problem about species classification, it is necessary to include the knowledge of genetic transmission of individual traits rather than only using different characters to identify a group of species (Palasio et al., 2017). Previous studies reveal that utilising only morphological information such as skin and skull structures are not sufficient proof to discriminate between cryptic species (Mayer et al., 2007; Baker & Bradley, 2006). Investigation via both morphological and genetic data can totally assist in evaluating species features together with their natural history (Baker & Bradley, 2006). Besides, utilising the DNA sequence information as the main informant have established

the taxonomy of the Rattini tribe for confirming group membership and evaluating taxonomic boundaries (Pagès et al., 2010). A series of successful investigations of the genetic diversity in the genus *Rattus* in the Asia Pacific region were performed recently (Aplin et al., 2011; Pagès et al., 2010; Robins et al., 2010, 2007; Chinen et al., 2005).

A substantial influence of population genetic revisions has been seen in interpreting the evolutionary processes, population, and species history. These breakthroughs have revolutionised population genetic study to such a level that this branch of knowledge which is previously debated as field of mathematics and theory, but has now develop into an explanatory science. The current expansion of coalescent and phylogenetic concept has transformed the method to examine and infer the molecular data. With the advancement in molecular techniques, genetic interpretation of populations has been much improved. It is probable to incorporate data acquired from numerical taxonomy and molecular means in order to generate a better profile of the population differentiation of a species (Quezada-Euán et al., 2007). The outcomes of such practical advances are so deep that they have altered the common population-genetics exploration (Zhang et al., 2003).

The number of DNA-based markers have greatly increased with current developments in molecular technology (Isabel et al., 1999). This increase in DNA-based markers has resulted in DNA sequencing becoming the most commonly-utilised molecular method for assessing genetic variation, as it delivers information of the nucleic acid sequence that is precise for the target DNA (Chan, 2005; Gharizadeh et al., 2003; Gülbitti & Sümer, 2003). The increase in the reliance on DNA sequencing has resulted in mitochondrial DNA (mtDNA) being identified as one of the most suitable markers in genetic studies for taxonomic re-evaluation of closely related species or populations of various species (Koh

et al., 2004). Using the analysis of short mtDNA fragments (cytochrome b (Cyt b) gene and cytochrome oxidase subunit I (COI) gene) (Robins et al., 2007; Ross et al., 2003), the possibilities to determine genetic species and estimate their proper taxonomic attributions was clearly demonstrated. A set of discrete specific genetic clades of *Rattus* have been shown (Robins et al., 2007; Chinen et al., 2005), and minimal genetic distances to segregate specific levels taxa were estimated (Pagès et al., 2010). Currently, ten specific level phylogroups have been characterised for Sundaland area, with more than 20 such groups recognised for entire Australasian region. The phylogenetic relationships among them are discussed in detail by Robins et al. (2010, 2007). More than 1000 available sequences of Cyt b, COI, and displacement loop (D-loop) genes in the mtDNA, the interphotoreceptor retinoid binding protein (IRBP) gene in the nuclear DNA (nDNA), and some other genes of wild representatives of *Rattus* were deposited in international gene databases (GenBank, NCBI, DDBJ), providing plenty of material for an accurate definition of intrageneric taxonomy based on up-to-date molecular approaches to actual natural biodiversity investigations.

The DNA barcoding method has been offered as a technique for the recognition of species based on the pairwise evolutionary discrepancy using an established marker, which is the mitochondrial DNA cytochrome oxidase I (COI) gene (Hebert et al., 2003). Based on observation, DNA barcoding has significantly smaller in term of intraspecific genetic divergences as compared to divergences between species (Robins et al., 2007). Robins et al. (2007) were the pioneer to endeavour in identifying *Rattus* species using mtDNA data mainly derived from museum tissue specimens. Nonetheless, their assumptions through DNA-barcoding and phylogeny tree-based approaches were inadequate since these approaches require well identified samples as references. Samples and tissues accessible from museums are collected by several individuals and seems very

possible in misidentifications, since species identification of rats is a tough task even for mammal experts. Furthermore, the taxonomy of the tribe Rattini is complex and frequently changed to that in use when specimens were primarily defined and recorded in museums (Robins et al., 2007).

The degree of divergence in Cyt b sequences was also suggested as a reference element in making assumptions regarding species-level discrepancies (Bradley & Baker, 2001). From the investigation of four genera of rodents, Bradley & Baker (2001) proposed that genetic distance with value less than 2% is signifying intraspecific divergence and value more than 11% is of species recognition. Yet, the confusion rises on how to deduce the values between 2% and 11%? The DNA-based species delimitation method suggested by Pons et al. (2006) depends on nucleotide data itself as the main information in constructing group membership and designate putative species without the involvement of establishing species as priors. This technique was described to be beneficial in identifying important individuals among groups whose recent taxonomy is inadequate or uncertain.

Regardless of the high diversity and abundance in Southeast Asia, only few studies have been done concerning the molecular genetics of Old World rats and mice at phylogenetic level (Gorog et al., 2004; Steppan et al., 2003). Although effectively recognised at a non-specific level by a specialist, Asian rats are frequently hard to segregate at a specific level utilising morphological or cytological measures. The extensive variety of intraspecific morphological discrepancy makes morphological features unacceptable for exact rodent species identification. This leads to an over-depiction of species and to a complicated taxonomy, hindered by an excess of equivalent words (Robins et al., 2007; Musser & Carleton, 2005). As murid rodents have high

adaptability, they evolve independently according to the landmasses and this made morphological-based classification to be more difficult. Molecular phylogenetic analysis however is less confined by this issue and previous revisions showed that the slowly evolving nuclear markers have been used to elucidate members and structure of Muroidea successfully (Steppan et al., 2004; Michaux et al., 2001).

In Malaysia, the taxonomic status of Murinae is poorly resolved until a study conducted by Tamrin & Abdullah (2011), where molecular phylogenetics and systematics of five genera (*Maxomys*, *Sundamys*, *Leopoldamys*, *Niniventer*, and *Rattus*) has been done. However, there are still other genera of rats in Malaysia that are left with no updated taxonomic status. In spite of ample revisions on the genetics of *R. rattus*, the molecular information and phylogeny of Peninsular Malaysia's *R. rattus* remain unknown. The phylogenetic position of *R. rattus* in Sundaland is also unclear. This may hold the key to reveal the biogeographic history of the black rat. The phylogenetic examinations at the subspecies level are insufficient as numerous scientists are focusing on the species classifications (Rosli et al., 2014). So, phylogenetic kinships among Malaysia's *R. rattus* have yet to be acknowledged, even with ample genetic revisions of the species globally.

Field identification of many small mammal species remains difficult, although their taxonomy and species diversity are established in the previous studies (Wilson & Reeder, 2005). Definitive identification can be made only via investigation of internal morphology (e.g. cranial and mandible). With the available of DNA sequencing, the identification of species has been rapidly increased as it solves the limitations regarding morphological-based identification. Identifications through morphology such as external measurements and skull sometimes can be misleading when considering the phenotypic variances between immature and mature phases of a species, among individuals of

unreliable species, and with the absence of phenotypic variations among cryptic species (Hebert et al., 2003). Molecular information from earlier research endorses the common incidence of cryptic mammal species that are unnoticed when solely concentrating on the morphological features (Baker & Bradley, 2006). Therefore, molecular technique serves as a very important tool in species identification.

## 2.7 Animal Mitochondrial DNA (mtDNA) and Nuclear DNA (nDNA)

Animal mitochondrial DNA (mtDNA) becomes the well-known part of eukaryotic DNA since it is very easy in purification compared to any particular piece of nuclear DNA (nDNA). Besides, mtDNA is also easier to characterise due to its smaller size and absences of the most confounding characteristics of nDNA (e.g. introns and repetitive sequences). Mitochondrial genome different from nuclear genomes in various aspects, i.e. mtDNA are maternally inherited, they encounter higher mutation rates, and not exposed to recombination (Cann et al., 1987). According to White et al. (1990), with the present of the polymerase chain reactions (PCR), it was possible to obtain mtDNA sequences directly from many taxa. Adkins et al. (2001) reported that rodents have been proven to evolve more faster compared to other placentals for some molecular markers. Therefore, mtDNA is an important molecule that has been used in comprehending the evolutionary kinships between individuals, populations, and species (Irwin et al., 1991). Saccone et al. (1999) stated that mitochondrial genome of animals is a recommended structure for investigation compared to the nuclear genome due to the absence of introns, do not undergo recombination, and its haploid mode of inheritance.

While analysis of mtDNA is convincing for genealogical and evolutionary studies of animal populations, it has several limitations. These involve the occurrence of mitochondrial pseudogenes in the nucleus that may occur at high copy numbers and be

mistakenly amplified, results in false comparisons (Gülbitti & Sümer, 2003). Mitochondrial DNA is maternally-inherited, thus biasing inferences made about species/population history using mtDNA as a genetic marker (Gülbitti & Sümer, 2003; Zhang et al., 2003). Another consideration with regards to the practising of mtDNA as a genetic marker is that the effective population size of mtDNA is only a quarter of that of nuclear autosomal sequences, thus mtDNA ancestries have a rapid lineage sorting rate and higher rate of allele loss, resulting in (i) possible over-simplification of evolutionary relationships estimated from mtDNA data, (ii) under-estimation of genetic diversity, and (iii) doubt in ancestral examination because of the increased possibility of missing links in mitochondrial haplotypes (Zhang et al., 2003). Even though a number of problems have been encountered in the use of mtDNA, its popularity in genealogical and evolutionary studies on animals has not diminished (Freeland, 2005). The popularity of mtDNA is based on its small size and highly-conserved structure (Gülbitti & Sümer, 2003), as well as the ease with which it can be manipulated, its comparatively fast mutation rate, its presumed lack of recombination, which resulting in an effectively clonal inheritance, and the fact that universal animal mitochondrial primers are readily available (Freeland, 2005). Mitochondrial markers often used to revise taxonomic issues of rodents at different taxonomic levels (Aplin et al., 2011; Robins et al., 2007).

Steppan et al. (2003) stated that most of the studies have relied primarily on Cyt b as it rapidly evolving mitochondrial genes. Because of having a rapid evolutionary rate and a clear evolutionary pattern within and between species, Cyt b is most prominent marker to be used in retrieving phylogenetic relationships among closely related taxa (Patwardhan et al., 2014). Also, Cyt b gene proved that they are effective in clarifying relationships within genera (Oshida et al., 2000; Oshida & Masuda, 2000; Kruckenhauser et al., 1999; Steppan et al., 1999). Hence, this marker has been employed for a variety of

systematic queries for deep phylogeny (Kumazawa & Nishida, 2000; Irwin et al., 1991; Lydeard et al., 1997; Cantatore et al., 1994; Meyer et al., 1990). This powerful marker also has been utilised to explore the genetic signatures of the black rat's species (Aplin et al., 2011). Besides, COI gene also is one of the favourable markers used in molecular studies. The marker is capable to code a large transmembrane protein discovered in the mitochondrion which is greatly conserved between taxa. Due to its rapid mutation rate to discriminate closely related species and since its sequence is very conserved among conspecifics, this gene is regularly utilised as a DNA barcode to recognise species (Hebert et al., 2003). Moreover, D-loop gene is a non-coding region and the most polymorphic site in the mitochondrial genome. It is a hypervariable region with no functional genes. The examination of nucleotide sequence variations of the D-loop marker is useful in elucidating the individual or ethnic identity and also the maternal relationships. The D-loop region accumulates mutations at rapid rate and very useful for ancestry studies. The D-loop gene of the mtDNA also encompasses features that control the replication of the molecule and is extremely informative. This potential shows that D-loop is suitable for the study of interspecies population structure (Liu et al., 2012). Martin et al. (2000) also stated that if the Cyt b sequence analyses are in agreement with D-loop sequence analyses, the relationships between populations inferred from DNA samples indicate stable phylogenetic relationships.

Besides, nuclear genes are most commonly used in vertebrate systematics (Steppan et al., 2003). Scientist such as Madsen et al. (2001), Matthee et al. (2001), Murphy et al. (2001), Groth & Barrowclough (1999), Stanhope et al. (1998), and Graybeal (1994), stated that most of the nuclear gene is relatively describing relationships among classes and orders. Springer et al. (2001) stated that nuclear genes were favoured because it contains more phylogenetic signal than mitochondrial genes. IRBP gene is one of the

common nuclear markers used in resolving phylogenetics and deep-level divergences of rodent species. Michaux et al. (2002) examined the phylogenetic connections between 17 extant species of Murinae via the nuclear IRBP marker and two mitochondrial genes of Cyt b and 12S ribosomal RNA (rRNA) emphasising on the genus *Apodemus* which revealed their phylogenetic positions among others. Chinen and colleagues (Chinen et al., 2005) demonstrated the genetic characterisation of two groups of black rats (*R. rattus* sensu lato) in Japan using Cyt b and IRBP genes, with the proof of introgression at various sites. IRBP gene was also used by Pagès et al. (2010) to revisit the taxonomy of the Rattini tribe. However, nuclear genes are actually has been less well evaluated for further recent discrepancies, which many of the nodes are expecting to be resolved and it has lower evolutionary rates which might not provide many characters (Steppan et al., 2004). Several molecular studies such as Huchon et al. (2002, 2000, 1999), Adkins et al. (2001), DeBry & Sagel (2001), Huchon & Douzery (2001), Murphy et al. (2001), Nedbal et al. (1996, 1994), and Catzeflis et al. (1995), have contributed a sensibly clear depiction of phylogenetic relationships of rodents and other placentals through mitochondrial and nuclear markers. Trewick (2000) and Bucklin et al. (1999) proved that DNA-based identification systems have been expanded with the involvement of the higher organisms.

## 2.8 Rodents as Carrier of Zoonotic Diseases

Rats are very active in the zoonotic transference of bacterial, viral, protozoal, and helminthic infections, and as carriers of arthropod vectors and ectoparasites such as ticks, mites, lice, and fleas. According to Pagès et al. (2010), more than 60 zoonotic diseases were hosted by rodents, e.g. leptospirosis, salmonellosis, rat bite fever, typhus, and jaundice. Nowadays, various rodent species has been demonstrated to be associated as hosts for pathogens. Genus *Rattus* is well known as human pests and mediums of diseases (Buckle, 1994), but nowadays vastly appreciated as model animals for study associated

to human health (Smits & Cuppen, 2006; Guenet, 2005). This is a serious hazard to the human health as the mobility of humans and animals are increasing, thus intensifying the risk of disease appearance. A common presumption is that the disease transmission is transferred from rats' species that live near human, but wild species far off from human residences have been demonstrated to play a key role in preserving, circulating, and disseminating pathogens (Jittapalapong et al., 2009). However, this presumption is questionable. Suzan et al. (2009) and Ostfeld & Keesing (2000) said that species diversity inside the host community also functioned in maintaining and transmitting the diseases to humans. Rodent-borne pathogens can spread both directly (through biting and water contaminated with rodent urine and feces) and indirectly (ectoparasitic arthropod vectors such as ticks and mites) (Meerburg et al., 2009). Thus, it shows how crucial to recognise rodents at specific level to acknowledge the function of each rodent species in the whole host-pathogen community in anticipating the risk of emergence or re-emergence infectious diseases.

## 2.9 Rodents and Leptospirosis

Leptospirosis is one of the utmost dominant zoonotic and emerging infectious disease promoted by the pathogenic spirochete genus *Leptospira*, a Gram-negative bacterium (Levett, 2001). Not less than 22 species of *Leptospira* were categorised via DNA-DNA hybridisation study and over 300 serovars were identified through agglutinating lipopolysaccharide (LPS) antigens (Bourhy et al., 2014). The classified species can be additionally separated into pathogenic, intermediate, and saprophytic species (Bourhy et al., 2014; Magalhaes et al., 2010). The saprophytes are supposed not to cause disease. The pathogenic *Leptospira* spp. comprise *L. interrogans*, *L. kirschneri*, *L. santarosai*, *L. weilii*, *L. alexanderi*, *L. borgpetersenii*, *L. alstoni*, *L. noguchii*, *L. kmetyii*, and *L. mayottensis*. The intermediate *Leptospira* spp. include *L. broomi*, *L. fainei*, *L. inadai*, *L.*

*licerasiae*, and *L. wolffii*. The saprophytic *Leptospira* spp. involve *L. biflexa*, *L. meyeri*, *L. wolbachii*, *L. idonii*, *L. vanthielii*, *L. terpstrae*, and *L. yanagawae* (Bourhy et al., 2014; Perolat et al., 1998).

During the early stage of infection, leptospires visible in the blood and disseminate almost all tissues and organs. They are eventually vanished from the body by the host's immune response to the infection, after an incubation period. However, leptospires may remain in the kidneys and urogenital tract of the hosts (WHO, 2003). Leptospires may continue to shed in the urine into the surroundings where they can persist for several months and occasionally even longer depending on favourable humid and warm conditions (Levett, 2001). Leptospirosis is transmitted directly or indirectly through the urine of affected animals or environments contaminated with urine of the reservoir animals (Ullmann & Langoni, 2011; Bharti et al., 2003).

The clinical presentations of the infection are tremendously varying, with asymptomatic or non-severe forms of leptospirosis with fever, headache, and muscle pain that can impulsively heal. Still, this illness can also worsen to more critical cases demanding hospitalisation in an intensive care unit with serious sepsis and several organs failures, involving liver failure, severe kidney dysfunctions, and bleedings, a demonstration symbolised the Weil's disease (Levett, 2001). Leptospirosis also triggers inflammation, initiated with higher levels of pro-inflammatory cytokines and chemokines in blood and organs of humans and animals (Papa & Kotrotsiou, 2015; Matsui et al., 2011; Chassin et al., 2009). Cytokines and chemokines are redundant secreted proteins with growth, differentiation, and activation roles that modulate and control the nature of immune responses and regulate the trafficking of immune cells and the cellular structure of immune organs.

Leptospirosis is a main health concern in rural communities of developing countries in Asia such as Thailand, Philippines, and Malaysia (Benacer et al., 2013; Tangkanakul et al., 2000). However, people in this area have little knowledge about the impact of leptospirosis (Meerburg et al., 2009). Past studies have signified that leptospirosis is a native disease in Malaysia happening in both urban and rural areas (Benacer et al., 2016). The number of reported cases has increased significantly until the Ministry of Health Malaysia gazetted leptospirosis as a notifiable disease in 2010. Even though the disease is most commonly found in tropical and rural areas, leptospirosis is evaluated as an expanding urban issue, mostly in fast growing towns in resource-limited areas with an absence of basic hygiene which promotes rodent-borne transmission (Ko et al., 1999).

A diversity of wild and domestic animals has been reported to harbour leptospires, but rodents are presumed to be the main reservoirs (Cosson et al., 2014). The isolation of zoonotic *Leptospira* spp. was initiated by Smith and co-workers who described 13 different leptospiral serogroups from rats, in the year 1961 (Smith et al., 1961). *Rattus rattus* (black rat) is a species of rodents globally distributed that is often related to leptospiral infection (Vanasco et al., 2003). In Malaysia, the earliest discovery on leptospirosis was discovered from black rats (Fletcher, 1928) and the rats are currently appraised as the primary maintenance hosts of *Leptospira* spp. and the conveyors for pathogenic leptospiral serovars. So far, up to 37 leptospiral serovars from 13 distinct serogroups have been recognised in Malaysia with over half are transmitted by rodents (Benacer et al., 2013). Roughly one-half of the pathogenic serovars associate to *L. interrogans* or *L. borgpetersenii* (Victoriano et al., 2009).

Prior research on leptospirosis in Malaysia concentrated mostly on humans and domestic animals in the zone of outbreak areas (Koay et al., 2004; Bahaman & Ibrahim,

1988). The rat's population in the world is increasing at an alarming rate, yet there is still insufficient documentation regarding this infection on urban rat populations in Malaysia. Recently, two predominant pathogenic leptospiral serovars were seen in two species; *Leptospira borgpetersenii* serovar Javanica and *Leptospira interrogans* serovar Bataviae spreading in the two dominant rat species, *R. rattus* and *R. norvegicus* in Peninsular Malaysia (Benacer et al., 2016). The additional two *Leptospira* spp. of *L. noguchii* and *L. weilli* were then discovered by Latifah et al. (2017) in wild rats from Kuala Lumpur.

## 2.10 Diagnosis of Rodent Leptospirosis

Traditional diagnostic tests, for instance microscopic agglutination test (MAT), are mainly serological and, thus, affirm the infection best at a late acute stage, when antibiotic remedy is almost less functional (Vanasco et al., 2016). Detection of leptospirosis via culture is not feasible in diagnostic laboratories due to the issue of the reagents, slow bacterial growth rate that usually takes weeks, and contamination problems in culture media (Erol et al., 2015). All these conflicts drew the awareness towards the usage of molecular methods for the discovery of rodent leptospirosis. Lately, molecular approaches such as conventional and real-time PCR are acknowledged as specific and sensitive trials for the fast detection of infection in the early stage of the disease and commonly negate the require for isolation and culture of the disease-causing organism for a valid result. These procedures can be conducted reliably on numerous templates, involving blood, kidney tissues, and urine (Hamond et al., 2014).

Rapid and confirmatory results through molecular diagnostic procedures are now succeeding serological tests in vicinities where leptospirosis is endemic (Picardeau, 2013). PCR-based investigations have been productively expanded for a broad range of

microbes, minimising the risk of false positive results introduced by carry-over contamination. The PCR has been appeared to be very practical for recognising and distinguishing organisms in instances where current culture methods have failed or have been insufficient, due to its high sensitivity, specificity, and rapid amplification (Post & Ehrlich, 2000). Amongst all molecular-based techniques for investigating microbial diversity, the PCR-based methods have been favoured by researchers. This is perhaps due to the relatively simple and fast nature of the process.

The general layout includes amplifying any desired gene for sequencing and identifying the source organism based on the sequence similarity to previously known organisms; or in case of a novel entity, finding the phylogenetic relationship between that sequence and others in order to categorise it accordingly (Kirk et al., 2004). It all started when Woese and colleagues presented that phylogenetic kinships of bacteria, and all life forms for that matter, can be conducted by differentiating a stable fragment of the genetic code (Woese, 1987; Woese et al., 1985; Woese & Fox, 1977). This new movement, following an earlier paper by Zuckerkandl & Pauling (1965), successfully launched biology into the world of "molecular chronometers" (a molecule whose sequence changes randomly in time). A molecular chronometer needs to fulfil a few criteria in order to be considered a suitable candidate for phylogenetic studies, and the ultimate choice for this purpose is found to be ribosomal RNA (rRNA) molecules (Woese, 1987). The genes that encode rRNA have a slow rate of evolutionary change and are highly conserved among all cellular life forms. Furthermore, the nucleic acid sequence and secondary structure of rRNA includes conserved regions found in all living organisms as well as variable domains containing group-specific sequence motifs which act as indicative factors (Clarridge, 2004; Buckley & Schmidt, 2002). Therefore, the rRNA-encoding genes have

proved to be very promising as biomarkers for studying microbial diversity in natural environments.

From the several segments of rRNA molecule, the genes that code for the small subunit (ssu) rRNA (i.e. 5S and 16S) and the 23S rRNA have been used more extensively for identifying bacteria (Clarridge, 2004). 16S and 23S rRNA genes are large enough to portray informative phylogenetic positions for reasonable characterisation and identification of bacteria species (Kolbert & Persing, 1999). The most commonly used sequence for bacterial taxonomy studies is the 16S rRNA gene which is comparable among all bacteria as well as the archeobacterial 16S rRNA gene and the eukaryotic 18S rRNA gene (Clarridge, 2004). This has allowed for a genetic sequence-based phylogeny relating all organisms and reconstructing the history of life (Pace, 1997). In furtherance, bacterial outer membrane proteins (OMPs), mainly those with surface-exposed regions, play vital roles in pathogen virulence mechanisms and alteration to surroundings environments, involving those discovered in the mammalian host. Thus, clarification and characterisation of the surface-exposed OMPs of *Leptospira* spp. is of strong interest in the leptospirosis area. A detailed, multi-dimensional strategy for evaluating surface exposure of leptospiral OMPs is important. One of the most studied leptospiral major outer membrane protein is the 32 kDa lipoprotein (LipL32). For screening of infection, PCR using *Leptospira* genus-specific primer of 16S rRNA gene and pathogenic *Leptospira* spp. specific primer of leptospiral major outer membrane protein LipL32 were highly exploited in earlier studies (Patel et al., 2017; Benacer et al., 2016, 2013; Plata-Luis et al., 2016; Hookey, 1992).

## 2.11 Rodents Immune Response Against Leptospiral Infection

The innate immune mechanism comprises the front line of host defence, providing an important role in primary identification and eradication of leptospires. One of the utmost crucial effector mechanisms through the first hours of infection is the initiation of the alternate pathway of the complement system (Barbosa et al., 2009; Meri et al., 2005). The acquired immune response relies on the fabrication of antibodies and the initiation of the classical pathway of the complement system since leptospires are extracellular pathogens. Almost all of the particular antibodies fabricated in leptospirosis are opposed to the lipopolysaccharides (LPS). Consequently, passive immunisation by polyclonal or monoclonal anti-LPS antibodies is capable in defence against the infection (Jost et al., 1986). In various experimental models, it is obvious that phagocytosis of *Leptospira* spp. by neutrophils and macrophages is only productive if this bacterium is opsonised by specific immunoglobulin G (IgG) (Wang et al., 1984; Banfi et al., 1982). Besides opsonisation, these antibodies may clump leptospires (Malkin, 1984) and stimulate the classical pathway of the complement system. Humoral-mediated immunity is extremely crucial in providing defence against leptospirosis in humans, pigs, guinea pigs, dogs, and hamsters (Faine et al., 1999).

As a result of infection, leptospires can be recognised by the immune system via the pathogen recognition receptors (PRRs), which identify molecular motifs named as pathogen-associated molecular patterns (PAMPs). PRRs can be classified into endocytic, signalling, and secreted receptors. Two major families of signalling PRRs play a critical part in pathogens' identification, the Toll-like receptors (TLRs) and the Nod-like receptors (NLRs) (Mogensen, 2009). Amongst TLRs, TLR2 and TLR4 are greatly studied in leptospirosis so far. LPS from Gram-negative bacteria is capable in stimulating TLR4, leading to a pro-inflammatory cytokine and chemokine dependent response (Miller et al.,

2005). Yet, leptospiral LPS which is less endotoxic compared to the Gram-negative LPS, triggers human macrophages via TLR2 rather than TLR4 (Werts et al., 2001). This diverse recognition is ascribed to the uncommon configuration of the leptospiral Lipid A component (Que-Gewirth et al., 2004), and perhaps a tactic that pathogenic leptospiral possibly use to evade the initiation of immune cells, promoting to the development of the infection in humans. The LPS is recognised by TLR2 and TLR4 in leptospires-resistant mice (Nahori et al., 2005). The activation of both TLR2 and TLR4 are definitely important to suppress leptospirosis in mice. Thus, with the lack of TLR activation, the innate immune receptors from the Nod-like family possibly in charge for activating inflammation in response to intrusive leptospires (Chassin et al., 2009).

In furtherance, immune response to leptospires has been progressively examined as it probably involved in the initiation of leptospirosis pathogenesis. The exploration of immune response does not only offer guidance for suitable vaccine formation but also for an additional clarifying mechanism of pathogenesis for this disease. A better consideration of host immune response in leptospirosis will guide to the expansion of various successful remedies and suppression of the disease (Lowanitchapat et al., 2010).

## 2.12 Host-Pathogen Transcriptomics by RNA Sequencing

The knowledge of host-pathogen interface can be established by exploring the transcriptome or the comprehensive collection of expressed RNA transcripts. Transcriptome profiling provides an approach to evaluate the comparative influence of gene products in selected cell, tissue, organism, or particular state (Mu et al., 2010). A very great technology that has transformed recent genomics study, acknowledged as 'Next Generation Sequencing' (NGS), was established in 2004. The NGS technologies have transformed biological sciences by enabling rapid discovery of genetic variation

among and within individuals efficiently (Metzker, 2010). This platform offers a rapid and low-cost process to sequence huge quantities of data with great breadth and depth. This genomics revolution has drastically reduced the cost of sequencing by > 1000 folds per base with increase outputs in an NGS platform (from less than one Gb/run in 2007 to 1500 Gb/run in 2016) (Yue & Wang, 2017). These recognised NGS platforms, e.g. the Illumina Genome Analyzer, the ABI SOLiD System, and the Roche/454 Genome Sequencer FLX Instrument have been verified as an effective and affordable tools for advanced study in various fields, for instance genome sequencing, epigenetics, metagenomics, and the disclosure of transcriptomics analysis, non-coding RNAs, and protein-binding sites (Morozova & Marra, 2008).

Besides, functional genomics provides an interesting way to gain deep understanding into the molecular basis of immune responses in species (e.g. rat), in which few data and a small number of tools (involving bacterial artificial chromosome libraries, monoclonal antibody reagents, and cell lines) are accessible. Functional genomics method has successfully been utilised to rapidly provide molecular information on the immune responses at gene level. Large scale immune related gene discoveries have been achieved through the functional genomics method (Robalino et al., 2007; Tassanakajon et al., 2006; Gross et al., 2001).

Several transcriptomic works related to leptospirosis have been implemented; yet, most of these works concentrated on genome-wide transcriptional profiling in pathogen (Caimano et al., 2014; Xue et al., 2010). So far, examinations of the host transcriptomic changes related to leptospirosis have been described in cell infection models (Xue et al., 2013; Martinez-Lopez et al., 2010). An extensive transcriptome profiling in murine kidneys-leptospires interfaces were performed recently to understand the overall changes

in renal gene expression during leptospirosis. The transcriptome study via NGS application has been effectively employed to observe the murine renal transcriptome profiles during leptospirosis, in regard to chronic kidney diseases (Chou et al., 2018).

## 2.13 Novel Findings on Cytokines and Chemokines Expressions During Leptospirosis Infection

The integration of whole-RNA sequencing (RNA-Seq) as the superior approach in revealing transcription remodelling via biological conditions (Vilariño-Güell et al., 2011) has empowered the investigation of differential gene expression related to host-pathogen interactions (Frese et al., 2013; Rumbo-Feal et al., 2013; Dötsch et al., 2012). Various research has verified the roles of cytokine initiation in immune response towards leptospires. Lowanitchapat et al. (2010) revealed the expression of cytokines in renal of hamsters affected with pathogenic *Leptospira* spp. Matsui et al. (2011) disclosed the gene expression profiles of immune mediators and histopathological discoveries in animal models of leptospirosis. In furtherance, the distinctive in vivo gene expression of major leptospiral proteins in animal models were also studied by Matsui et al. (2012). Matsui et al. (2016) examined and described the significant variances in transcript levels of cytokine and chemokine genes in the renals of leptospiral-infected animals.

Numerous plasmatic inflammatory genes are expressed during sepsis and organ dysfunction, promoting to or indicating a dysregulation of the inflammatory response towards infectious bacteria. Remarkably, there is a high expression of inflammatory cytokines, i.e. tumor necrosis factor-α (TNF-α), interleukin (IL)-1β, IL-6, and IL-10, that promotes organ dysfunction, leading to fatality in septic shock (Opal, 2007; Philippart & Cavaillon, 2007; O'Grady et al., 1999). Besides, C-C-type chemokines, for instance macrophage inflammatory protein-1α (MIP-1α)/CCL3 and MIP-1β /CCL4 or C-X-C-type

chemokines, for example IL-8/CXCL8 and gamma interferon (IFN-γ)-inducible protein 10 kDa (IP-10)/CXCL10 are also initiated throughout sepsis (Punyadeera et al., 2010; Fujishima et al., 1996). The information recommended that the initiation of cytokine probably involved in tubulointerstitial nephritis triggered by leptospirosis. Tian et al. (2006) suggested that *Leptospira* spp. may induce kidney dysfunction by increasing the extracellular matrix production. The role and distinctive expression of each cytokine and chemokine described the progression of leptospirosis. Inflammation is necessary in the rating of healthy immune response, though it is also important in the pathogenesis of various infections. The balance between pro- and anti-inflammatory mediators of cytokines and chemokines immune responses is important in protecting multiple organ failure due to inflammations, and may assist host survival.

The differentially expressed genes (DEGs) of pro- and anti-inflammatory mediators of cytokines and chemokines immune responses, TNF-α, IL-1β, IL-6, IL-10, MIP-1α, and IFN-γ were highly observed and reported in mammals experimentally infected with leptospirosis (Matsui et al., 2016; Vernel-Pauillac & Goarant, 2010; Lowanitchapat et al., 2010). These genes were previously described as most related to the immunopathogenesis of leptospirosis. To date, however, the DEGs of these cytokines and chemokines in both kidney and liver of wild rat reservoirs naturally infected with leptospirosis have been poorly reported. Deep investigation is necessary in order to elucidate the host-pathogen interactions.

Several research of leptospirosis-related interstitial nephritis and chronic infection have been reported in mice (Gomes-Solecki et al., 2017). Even though asymptomatic, carrier hosts are exposed to morphological alterations of renals as examined in the rats experimentally infected with leptospires or wild carnivores naturally infected by

leptospires which primarily demonstrating chronic interstitial nephritis (Millan et al., 2008; Monahan et al., 2008; Tucunduva de Faria et al., 2007). It has been previously shown that TLRs, Na/K-ATPase, and inducible nitric-oxide synthase assist in the innate response to leptospirosis infection (Fanton d'Andon et al., 2014; Gonc¸alves-de-Albuquerque et al., 2012; Lacroix-Lamandé et al., 2012; Chassin et al., 2009). Moreover, it has been proposed by an in vitro study that pathogenic *Leptospira* spp. may avoid the host innate immune response to infection via its resistance to the complement system (Fraga et al., 2016). The discoveries concerning cytokine and chemokine cascades might clarify the role of renal microenvironments partaking in kidney injury triggered by *Leptospira* spp.

# CHAPTER 3: METHODOLOGY

## 3. 1 Ethics and Permits Approval

This study was endorsed by the Institutional Animal Care and Use Committee, University of Malaya, Malaysia (UM IACUC) with the ethics reference no. ISB/10/06/2016/NHMI (R). The Department of Wildlife and National Parks authorised this study in collecting wild samples with permit reference number; JPHL&TN(IP): 100-34/1.24 Jld. 7(6). This study was carried out without the involvement of any endangered species.

## 3.2 Sampling and Study Area

Located north of the equator, with the latitude of roughly 1.27° and 6.72° N and the longitude of roughly 99.64° and 104.53° E of Greenwich is Peninsular Malaysia. The highest peak of Peninsular Malaysia is Mount Tahan, with altitude of 2187 metre (m). Malaysian climate is tropical, i. e. hot, humid, and rainy throughout the year. The climate is influenced by monsoonal wind systems, with the Northeast Monsoon blowing from November to March (rainy season), while the Southwest Monsoon brings rain between May and September (dry season).

In this study, 13 main cities in Peninsular Malaysia (Figure 3.1) were selected as sampling sites. The trappings of wild rats were conducted with the assistance of municipalities of each city as part of their vector control programme. Sampling was carried out in various sites (i.e. housing areas, fresh markets, and seaside) with emphasis on populated areas, as the study organisms are usually found nearby to human habitation. Trapping was conducted over a period of four days and three nights for each session at every sampling site with a total of 30 trapping sessions per all 13 locations, organised between December 2016 and February 2018.

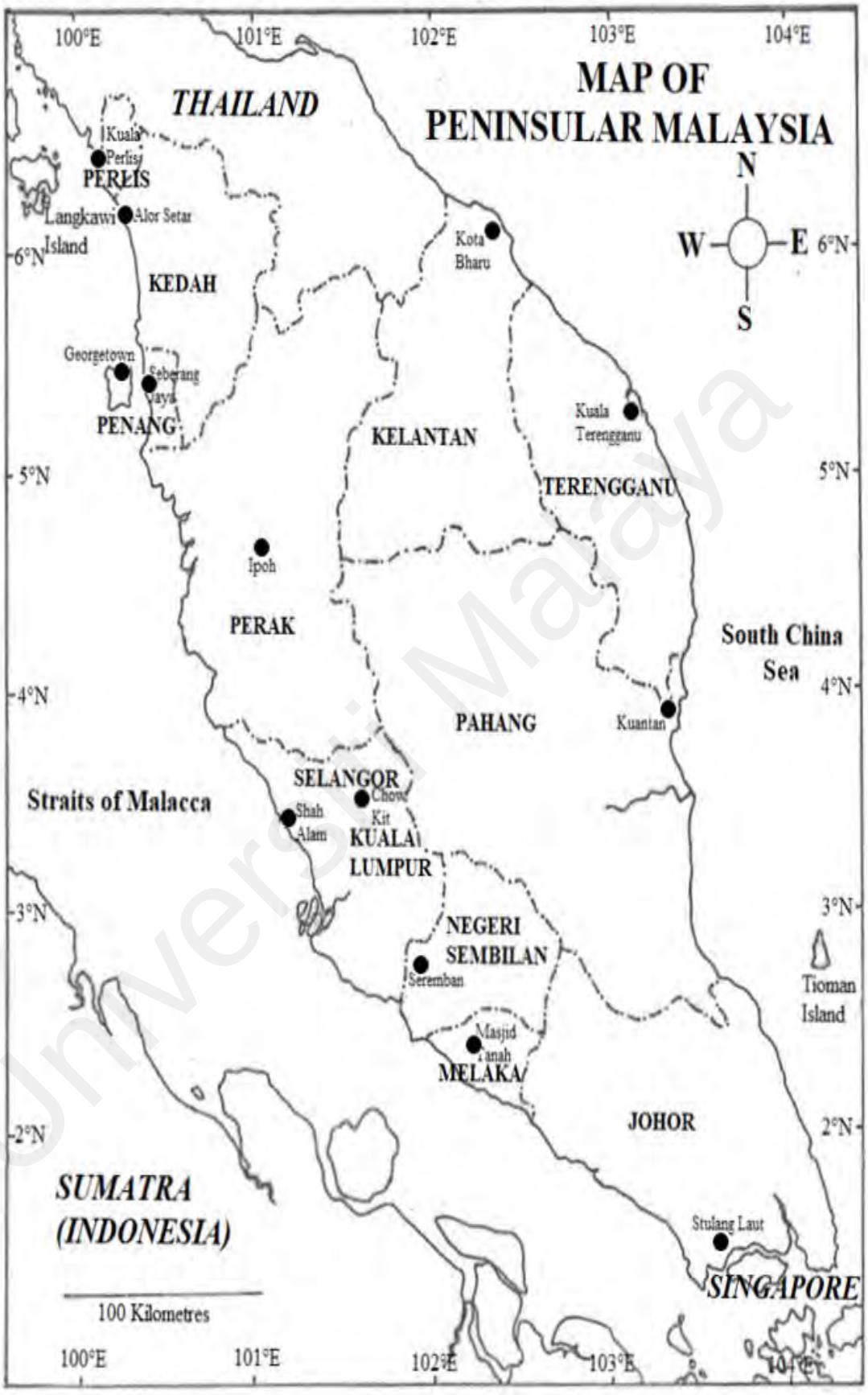


**Figure 3.1:** Sampling location of *Rattus* spp. throughout Peninsular Malaysia.

Thirty cage traps sized 28 x 15 x 12 centimetre (cm) baited with banana, coconut, or bread coated with peanut butter or sardines were placed at each site. The traps were set up in the late evening and checked early morning. The captured rats were identified as *Rattus* spp. based on their external size and features by referring to the taxonomic keys (Francis, 2008; Wilson & Reeder, 2005; Aplin et al., 2003; Corbet & Hill, 1992). Field identifications were also made based on phenotypic characteristics such as overall size of the specimen, tail length, and head-body length.

A total of 130 individuals identified as *R. rattus* (67 males (♂), 63 females (♀)) were collected throughout this study (Table 3.1). In addition, 23 individuals of several species of genus *Rattus*; *R. tanezumi* (*N* = 7: 3 ♂, 4 ♀), *R. norvegicus* (*N* = 6: 3 ♂, 3 ♀), *R. tiomanicus* (*N* = 4: 2 ♂, 2 ♀), *R. argentiventer* (*N* = 3: 1 ♂, 2 ♀), and *R. exulans* (*N* = 3: 2 ♂, 1 ♀) accidentally caught during the trapping sessions were also included in this study to infer the phylogenetic relationships of genus *Rattus* in Peninsular Malaysia. Live specimens were euthanised in a sealed container connected to a carbon dioxide ($CO_2$) tank. Dead specimens then underwent dissection procedures after information such as photos, external measurements, sex, and weight were recorded. The basic field measurements of the collected *Rattus* spp. were tabulated in Appendix A. Skinning of specimens were done and deposited at Museum of Zoology, University of Malaya for museum collections. All skulls of *R. rattus* were extracted for morphometrics study. Kidney and liver tissues were preserved in 95% (v/v) ethanol and kept in -80 °C freezer to be used in genetic study.

**Table 3.1:** The localities and samples sizes (*N*) from which *R. rattus* populations were collected in Peninsular Malaysia.

| Locality | Latitude | Longitude | Average precipitation (mm per year) | Samples sizes (*N*) |
|---|---|---|---|---|
| Kuala Perlis, Perlis | 6° 23' 52.4" N | 100° 7' 50.4" E | 2133 | 9 (2♂, 7♀) |
| Kota Bharu, Kelantan | 6° 8' 23.54" N | 102° 14' 31.93" E | 2558 | 8 (3♂, 5♀) |
| Alor Setar, Kedah | 6° 7' 29.3" N | 100° 22' 4.1" E | 2276 | 12 (4♂, 8♀) |
| Georgetown, Penang island | 5° 26' 8.3" N | 100° 18' 32.8" E | 2462 | 6 (4♂, 2♀) |
| Seberang Jaya, Penang mainland | 5° 24' 4.5" N | 100° 23' 51.1" E | 2425 | 7 (4♂, 3♀) |
| Kuala Terengganu, Terengganu | 5° 19' 48.72" N | 103° 8' 26.88" E | 2736 | 10 (6♂, 4♀) |
| Ipoh, Perak | 4° 35' 50.9" N | 101° 5' 24.4" E | 2435 | 8 (8♂, 0♀) |
| Kuantan, Pahang | 3° 48' 27.7" N | 103° 19' 33.6" E | 2887 | 12 (5♂, 7♀) |
| Chow Kit, Kuala Lumpur | 3° 9' 35.9" N | 101° 41' 49" E | 2486 | 15 (8♂, 7♀) |
| Shah Alam, Selangor | 3° 4' 23.8" N | 101° 31' 6.5" E | 2304 | 9 (5♂, 4♀) |
| Seremban, Negeri Sembilan | 2° 43' 33.2" N | 101° 56' 16.2" E | 2043 | 12 (7♂, 5♀) |
| Masjid Tanah, Melaka | 2° 21' 8.2" N | 102° 6' 32.1" E | 2101 | 10 (6♂, 4♀) |
| Stulang Laut, Johor | 1° 28' 10.3" N | 103° 46' 53.9" E | 2554 | 12 (5♂, 7♀) |

## 3.3 Species Diversity of *Rattus* (Rodentia: Muridae) on the Malay Peninsula and Their Phylogenetic Relationships Inferred from Complete Mitochondrial Cytochrome b Gene Sequences

A total of 35 individuals from the genus *Rattus*; *R. rattus* (*N* = 12), *R. tanezumi* (*N* = 7), *R. norvegicus* (*N* = 6), *R. tiomanicus* (*N* = 4), *R. argentiventer* (*N* = 3), and *R. exulans* (*N* = 3) were utilised to infer the phylogenetic relationships of genus *Rattus* in Peninsular Malaysia. For species validation and estimation of the molecular divergence time, an additional 86 DNA sequences were retrieved from GenBank; which 79 of them are representatives of genus *Rattus* from various countries, six of them belongs to genus *Mus*, and one sequence of *Leopoldamys sabanus* served as an outgroup (Appendix B).

### 3.3.1 Extraction of genomic DNA

Genomic DNA from liver tissue of the 35 individuals of wild rats were extracted using the modified innuPREP DNA Mini Kit (Analytik Jena). 50 milligram (mg) tissue samples were weighed and put into 2 millilitre (ml) sample tube. A total of 400 microlitre (µl) lysis solution TLS buffer and 25 µl Proteinase K were added into the sample tube for DNA lysis. The samples were vortexed for five seconds. The samples were then incubated in an incubator at a temperature of 50 °C for one hour. After incubation, the samples were centrifuged at 12000 revolutions per minute (rpm) for one minute. The supernatants were transferred into 1.5 ml microcentrifuge tube. Next, 400 µl of binding solution TBS buffer were mixed into the samples for DNA binding. The samples were vortexed for 15 seconds. Spin filter were added to a receiver tube. Samples were added to the spin filter. The mixtures were centrifuged at 12000 rpm for two minutes. The filtrates were discarded and the spin filter were removed to a new receiver tube.

For washing purposes, 500 µl washing solution HS buffer were added and centrifuged at 12000 rpm for one minute. 750 µl washing solution MS buffer were added and centrifuged at 12000 rpm for one minute. The filtrate was then discarded. Spin filter were added to a new receiver tube and centrifuged at a maximum speed for two minutes to remove the ethanol. Finally, for elution process, spin filter was added to an elution tube. 40 µl elution buffer were added and incubated for three minutes at room temperature. The samples were centrifuged again at a speed of 8000 rpm for one minute.

### 3.3.2 Determination of DNA quality and concentration

The quality of extracted DNA has been tested using 1% (w/v) agarose gel electrophoresis. Five µl of DNA and two µl of loading dye were mixed and loaded into wells of 1% (w/v) agarose gel, added with gel stain. Five µl of 1 kilobase (kb) DNA ladder

was loaded into the first well to quantify the DNA. Electrophoresis were done in 1 X TAE buffer (0.04 Molar (M) Tris acetate pH 8.5 and 0.001 M Ethylenediaminetetraacetic acid (EDTA)) at a voltage of 100 Volt (V) and current of 200 milliampere (mA) for 30 minutes. Digital image of agarose gel was visualised under ultraviolet (UV) light in transilluminator machine using Alpha Imager Gel Documentation System (Siber Hegner, Germany). To check the concentration of the isolated genomic DNA, a NanoDrop™ 2000 Spectrophotometer (Thermo Fisher Scientific, USA) was used. The purity of isolated genomic DNA was determined by the A260/A280 nanometres (nm) ratio using 1:10 (DNA: buffer) dilution. This value is best between 1.8 and 2.0 (Sambrook & Russell, 2001). A ratio of A260/A230 nm can be calculated to quantify the presence of such contaminating agents, when best results are achieved with values greater than 1.5 (Sambrook & Russell, 2001).

### 3.3.3 DNA amplification by the polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) were employed in order to amplify the targeted locus by using Mastercycler® Nexus (Eppendorf North America, Inc.). PCR was conducted using the PCR reagents included 12.5 µl of 2 X Power Taq PCR MasterMix (BioTeke, Beijing), 0.5 µl of each 10.0 micromolar (µM) of forward and reverse primer, 1.0 µl of genomic DNA template, and 10.5 µl of ultrapure water added up to the final volume of 25.0 µl, as suggested by the manufacturer. PCR was performed using a set of established primers correspond to the Cyt b gene. The complete Cyt b sequences (with 1140 bp) were amplified using established forward primer L14723 (5′-ACCAATGACATGAAAAATCATCGTT-3′) and reverse primer H15915 (5′-TCTCCATTTCTGGTTTACAAGAC-3′) as in Irwin et al. (1991). The profile for the PCR amplification were as follows: initial denaturation for four minutes at 94 °C, continued with 35 cycles of denaturation for 30 seconds at 94 °C, annealing for 30

seconds at 54 °C, an extension for one minute at 72 °C, and a final extension stage for ten minutes at 72 °C, before holding-up to 10 °C after the reaction completed.

### 3.3.4 Gel electrophoresis

The product of PCR amplification went through electrophoresis by using 2% (w/v) agarose gel. The gel was prepared by adding 0.5 g of agarose powder (Cambrex Bioscience Rocklan, Inc, ME USA) with 25 ml 1 X TAE buffer. 1 X TAE buffer was used in gel preparation and acts as a buffer in the electrophoresis tank. Mixture of agarose powder and 1 X TAE buffer were put into a microwave oven for three minutes. 2.5 µl of SYBR® Safe DNA gel stain (BioTeke, Beijing) were added into 25 ml of agarose gel. The gel mixture was poured into a gel tray containing sample comb and were kept for 30 minutes to ensure the gel becomes solid.

Next, the sample comb was removed carefully from the gel tray to ensure the wells of the gel were in a good condition. The gel tray was put into an electrophoresis tank containing 1 X TAE buffer. Five µl of PCR products and two µl of loading dye were mixed and loaded into wells of 2% (w/v) agarose gel. Five µl of 100 base pair (bp) DNA ladder were loaded into the wells for identification of the DNA fragments according to their sizes. Electrophoresis were done at a voltage of 80 V, a current of 180 mA for 45 minutes. Then, gel was removed from the electrophoresis tank. Digital image of agarose gel was visualised under ultraviolet (UV) light in transilluminator machine using Alpha Imager Gel Documentation System (Siber Hegner, Germany).

### 3.3.5 Purification of PCR products

Purification of PCR products were done using the innuPREP PCRpure Kit (Analytik Jena). For the binding of PCR fragments, 500 µl of binding buffer and five µl of PCR

products were added into 1.5 ml microcentrifuge tube. The sample were vortexed shortly. Spin filter were added to a receiver tube. Mixed sample were added into the spin filter. The sample were centrifuged at 12000 rpm for two minutes. For elution of PCR fragments, spin filter was added to an elution tube. Ten µl of elution buffer were added into the spin filter and incubated for three minutes at room temperature. Finally, the sample were centrifuged at 8000 rpm for one minute.

### 3.3.6 DNA sequencing

The PCR product that has undergone a purification process were sent to Apical Scientific Sdn. Bhd. (Seri Kembangan, Selangor, Malaysia) together with a set of forward and reverse primers for nucleotide sequencing purposes. Sequencing using Applied Biosystems 3730XL Genetic Analyzer was done to obtain DNA sequences of the targeted gene locus by using Sanger Sequencing method.

### 3.3.7 Sequence editing and alignment

In the editing process, two methods were used; visual editing and editing using a computer software. DNA sequencing results obtained were proofread by checking the chromatogram file for confirmation, to ensure there were no errors on bases interpretation done by the sequencing machines. Both aligned forward and reverse sequences were edited using Molecular Evolutionary Genetics Analysis Version 7.0 (MEGA7) (Kumar et al., 2016). To ensure the targeted species locus sequences were obtained, the edited sequences were validated using sequence similarity searches using GenBank BLASTn application. Geographic localities of matching sequences were noted. The MEGA7 ClustalW multiple alignment algorithms (Chenna et al., 2003) was then used to align all the sequences, together with selected DNA sequences and an outgroup sequence obtained from GenBank. The stop codon (*) was checked in the translated protein sequences.

### 3.3.8 Sequence and phylogenetic analysis

In the sequence analysis, MEGA7 was heavily exploited, as well as DNA Sequence Polymorphism Version 5.10.1 (DNASP v5) (Librado & Rozas, 2009). Cyt b sequences from each sampled species from across their range was downloaded from GenBank and compared to the sequences generated in this study. The edited and aligned sequences were analysed using MEGA7 software to verify the percentage of nucleotide compositions, sequence divergences, and pairwise analysis calculated using Kimura 2-Parameter (K2P) algorithm model. DNASP v5 was used to generate haplotype data within and between species. Network 5.0.0.1 (Bandelt et al., 1999) was used to generate a minimum-spanning network (MSN) to illustrate the haplotype relationships. A minimum spanning network for each species was constructed, and the geographic origin of the haplotypes is indicated.

Phylogenetic analysis involves the construction of phylogenetic tree. The phylogenetic trees were built using distance-based (neighbor-joining (NJ) tree) and Bayesian inference (BI). Separate software programmes were used to build the phylogenetic tree, specifically, MEGA7 for the NJ tree and MrBayes for BI (Huelsenbeck & Ronquist, 2001). The Kimura 2-parameter model was used in NJ tree reconstructions tested with a bootstrap value of 1000. The best substitution model for Cyt b sequences was selected using Modeltest version 3.7 software (Posada & Crandall, 1998) through the Akaike information criterion (AIC) requirements and used in the Bayesian inference using MrBayes software.

To construct BI, the best model for the sequences selected was the Hasegawa-Kishino-Yano model with invariant sites plus gamma distribution, HKY+I+G (-lnL = 5429.9136; AIC = 10871.8271) under Akaike information criterion (AIC), with invariable sites proportion of 0.5749, gamma shape parameter of 1.0623, and base frequencies of 0.3163

for Adenine (A), 0.3224 for Cytosine (C), 0.1059 for Guanine (G), and 0.2553 for Thymine (T). Metropolis-coupled Markov chain Monte Carlo (MCMC) was run with 2350000 generations, and the tree was sampled every 1000 generations. A split frequency probability (*P*) of 0.009861 was obtained from the run of MrBayes. The first 25% of the trees acquired in the analysis were discarded as burn-in, and a 50% majority-rule consensus for the remaining trees was constructed. The posterior probabilities (*PP*) were summarised for each branch.

### 3.3.9 Molecular clock analyses

For an estimation of the molecular divergence time, six additional sequences of house mouse, *Mus musculus* (GenBank accession numbers: EF108340-EF108345) were included into the analysis to assimilate a calibration timepoint from the murine fossil evidence. The root of the tree was calibrated using the *Mus-Rattus* divergence time of 12.3-11.0 million years ago (Mya). The divergence time interval has been highly exploited in the divergence time estimations of mammal (Donoghue & Benton, 2007). BEAUti v1.7.5 and BEAST v1.7.5 (Drummond & Rambaut, 2007) were used in order to construct a Bayesian phylogenetic analysis with molecular dating estimation. The MCMC analyses involved a randomly generated starting tree, an uncorrelated lognormal relaxed clock with HKY+G+I substitution model on data partitioned by the three codon positions, a speciation with Yule process for the tree-prior, and the programme's default prior distributions for model parameters and statistics. The calibration setting was served as a lognormal distribution (offset = 11.0, mean = 0.413, standard deviation [*SD*] = 1.0). The calculation of the effective sample sizes (ESSs) to estimate the convergences was handled using Tracer v1.5 (Rambaut, 2009), and the first 10% generations were scrapped as burn-in. The ESSs value > 200 were ensured in the whole analysis. MCMC analyses were

generated for 50000000 steps, with posterior sampling was undertaken every 5000 steps after a burn-in of 10% steps.

### 3.4 Population Structure of *Rattus rattus* in Peninsular Malaysia Based on Mitochondrial and Nuclear DNA Data

A total of 130 individuals of *R. rattus* trapped throughout this study (as listed in Table 3.1) were enrolled to deduce the population structure of *R. rattus* from Peninsular Malaysia. The genomic DNA from liver tissue of all the 130 individuals of *R. rattus* were extracted using the modified innuPREP DNA Mini Kit (Analytik Jena). The procedures for extraction of the genomic DNA is described in section 3.3.1. The DNA quality and concentration were also checked following the procedures mentioned in section 3.3.2.

Next, PCR amplification were employed in order to amplify the targeted locus in the mitochondrial DNA (mtDNA) and nuclear DNA (nDNA), by using Mastercycler® Nexus (Eppendorf North America, Inc.). PCR reactions were generated using the 2 X Power Taq PCR Master Mix (BioTeke, Beijing), which has extreme speed, high accuracy, and a very high yield in reduced times. The list of primers used to deduce the population structure of *R. rattus* from Peninsular Malaysia is shown in Table 3.2. Mitochondrial markers of cytochrome oxidase subunit I (COI), cytochrome b (Cyt b), displacement loop (D-loop), and nuclear marker of interphotoreceptor retinoid binding protein (IRBP) used in this study were successfully established in the previous rodents' study. The PCR profile and the PCR chemical components used for amplification of mtDNA and nDNA genes is shown in Table 3.3 and Table 3.4, respectively.

**Table 3.2:** Details of primers used to deduce the population structure of *R. rattus* from Peninsular Malaysia.

| Genes | Primers | Nucleotide sequence (5'-3') | Fragment Length (bp) | Annealing Temperature (°C) | Original Publication |
|---|---|---|---|---|---|
| COI | BatL5310 | CCTACTCRGCCATTTTACCTATG | 711 | 48 | Robins et al. (2007) |
| | R6036R | ACTTCTGGGTGTCCAAAGAATCA | | | |
| Cyt b | L14723 | ACCAATGACATGAAAAATCATCGTT | 1140 | 54 | Irwin et al. (1991) |
| | H15915 | TCTCCATTTCTGGTTTACAAGAC | | | |
| D-loop | EGL4L | CCACCATCAACACCCAAAG | 564 | 60 | Robins et al. (2007) |
| | RJ3R | CATGCCTTGACGGCTATGTTG | | | |
| IRBP | I2 | ATCCCCTATGTCATCTCCTACYTG | 1227 | 52 | Pagès et al. (2010) |
| | J1 | CGCAGGTCCATGATGAGGTGCTCCG TGTCCTG | | | |

**Table 3.3:** PCR profile used for amplification of mitochondrial and nuclear genes.

| Stage | Temperature (°C) | Time | Number of cycles |
|---|---|---|---|
| Pre-denaturation | 94 | 4 minutes | 1 |
| Denaturation | 94 | 30 second | 35 |
| Annealing | 48-60 | 30 seconds | |
| Extension | 72 | 1 minute | |
| Post extension | 72 | 10 minutes | 1 |
| Hold | 10 | ∞ | |

**Table 3.4:** PCR chemical components used for amplification of mitochondrial and nuclear genes.

| Components | Volume (µl) | Final Concentration |
|---|---|---|
| 2 X Power Taq Master Mix (mixture of DNA polymerase, reaction buffers, $MgCl_2$, and dNTPs) | 12.5 | 1 X |
| Ultrapure water | 10.5 | - |
| Forward primer (10.0 µM) | 0.5 | 0.2 µM |
| Reverse primer (10.0 µM) | 0.5 | 0.2 µM |
| DNA template | 1.0 | ~ ng (nanogram) |
| Total | 25.0 | - |

The product of PCR amplification went through electrophoresis by using 2% (w/v) agarose gel, as stated in section 3.3.4. Purification of PCR products were done using the innuPREP PCRpure Kit (Analytik Jena), as indicated in section 3.3.5. The PCR product that has undergone a purification process were continued with the nucleotide sequencing, together with a set of forward and reverse primers.

DNA sequencing results obtained were proofread by checking the chromatogram file for confirmation, to ensure there were no errors on bases interpretation done by the sequencing machines. Both aligned forward and reverse sequences were edited using MEGA7. To ensure the targeted species locus sequences were obtained, the edited sequences were validated using sequence similarity searches using GenBank BLASTn application. The MEGA7 ClustalW multiple alignment algorithms was then used to align all the sequences. For protein coding genes (COI, Cyt b, and IRBP), the stop codon (*) was checked in the translated protein sequences. For the mtDNA genes of COI, Cyt b, and D-loop, SeaView Version 4.7 (Gouy et al., 2010) was used to concatenate all the three genes sequences, sum up to a total of 2415 bp used in the analysis. The nDNA gene of IRBP was analysed separately. The aligned sequences were then analysed at three different levels, namely, sequence analysis, phylogenetic analysis, and population genetics analysis.

### 3.4.1 Sequence analysis

The edited and aligned sequences were analysed using MEGA7 software to verify the percentage of nucleotide compositions, sequence divergences, transition to transversion ratio, and pairwise analysis. Kimura 2-Parameter (K2P) algorithm with a bootstrap value of 1000 was used for pairwise analysis. DNASP v5 was used to estimate several measures of DNA sequence variation within and between populations. Sequence analysis are vital

in estimating a few key end results, such as single-nucleotide polymorphisms (SNPs), net nucleotide divergence (*Da*), and nucleotide diversity ($\pi$).

### 3.4.2 Phylogenetic analysis

Phylogenetic analysis involves the construction of phylogenetic tree. The phylogenetic trees of both mtDNA and nDNA genes were built using distance-based (neighbor-joining (NJ)) tree. Specifically, MEGA7 software programme was used to build the NJ phylogenetic tree. The Kimura 2-parameter model was used in NJ tree reconstructions tested with a bootstrap value of 1000. A single species of *Leopoldamys sabanus* was used as an outgroup to root the phylogenetic tree.

### 3.4.3 Population genetics analysis

Evaluations of population genetic parameters such as haplotype diversity (*Hd*), average number of nucleotide differences (*k*), segregating sites (*S*), nucleotide subdivision (*Nst*) (Lynch & Crease, 1990), population subdivision (*Fst*) (Hudson et al., 1992), and the number of migrants per generation (*Nm*) (Hudson et al., 1992) were also estimated from the mtDNA and nDNA dataset using DNASP v5. The neutrality tests were examined by employing Tajima's test of neutrality, *D* (Tajima, 1989), Fu & Li's *D** and *F** (Fu & Li, 1993), and Fu's *Fs* (Fu, 1997). Tajima's *D* test compares the average number of pairwise nucleotide differences (*k*) between haplotypes in a sample expected from the number of segregating sites (*S*). Fu & Li's $D^*$ and $F^*$ and Fu's *Fs* were employed to test for deviation of sequence variation from evolutionary neutrality. Fu's *Fs* is depended on the probability of the observed number of haplotypes occurring under conditions of neutrality, while Fu & Li's $D^*$ and $F^*$ compare estimates of theta based on mutations in internal and external branches of a genealogy.

Population bottlenecks and expansions, and mutational rate heterogeneity may all result in a Poisson distribution of substitution differences between pairs of haplotypes (Slatkin & Hudson, 1991). Therefore, historic demographic and mismatch distribution analysis (Rogers, 1995; Rogers & Harpending, 1992) was performed using Arlequin Version 3.5 with 1000 permutations (Excoffier & Lischer, 2010). The population genetic structure was analysed with an analogue of *F*-statistic ($F_{ST}$) to investigate the amount of genetic variability segregated between and within populations (Excoffier et al., 1992), implemented in analysis of molecular variance (AMOVA) (Schneider et al., 2000) in Arlequin Version 3.5. The significance of variance components was assessed using 1000 permutations. Arlequin Version 3.5 was also used to conduct Mantel test (Mantel, 1967), with 1000 permutations to observe the effect of isolation-by-distance (IBD) by evaluating the analogy between geographical distance and genetic differentiation among populations. The parsimony criterion was applied to rebuild the haplotype relationships, assuming that differences at any given site between two randomly drawn haplotypes were unlikely to have arisen from more than one mutational step (Alexandrino et al., 2002). Network 5.0.0.1 was used to generate a minimum-spanning network (MSN) to illustrate the haplotype relationships.

## 3.5 Morphometric Analysis of Skull Characteristics of *Rattus rattus*

Morphometric analysis was undertaken to examine the intraspecific variation of the skull characters in *R. rattus* from Peninsular Malaysia. Conventional morphometrics was employed to elucidate the skull variation between specimens.

### 3.5.1 Skinning

Skinning process of the rats started through the incision line at the belly part using scissor. The top layer of skin of the specimen was hold by using tweezers. After that, the

skin from the lower part to the middle part of the body were cut. By using scalpel, the membrane was cut slowly from top layer skin of hindfoot. Then, the top of the tailbone was carefully cut by avoiding deep-cut to ensure the tail attached or to avoid cutting the tail accidentally. The skin that has been cut was removed slowly. Next, the skin in forefoot and hindfoot were removed. Finally, skin at the head part, especially at the ear part and the nose part was cut slowly and carefully, as it is the crucial part and used for morphological identification. After done with skinning, the flesh was removed out and the inner skin of rat were cleaned with Sodium borate or borax powder to remove the remaining flesh at the skin and air-dried. Finally, cotton wool was inserted into the sample skin and sewed.

### 3.5.2 Skull extraction

Extraction of skull was carefully handled to ensure the skulls were in a good condition and to avoid broken. Both 70% (v/v) ethanol and distilled water are the most important medium to soak the skulls. Both of them were used frequently in all the following steps. Firstly, most of the flesh at the head part were removed. Tissue was removed with a variety of scalpels, knives, and other tools, such as dental instruments, depending on the size of the head. It is not necessary to remove all the tissue, but the processing time becomes shorter once more flesh was removed. When tissues were being removed, carefully not to scrape too hard and scratch the bone surfaces. At the final stage, the brain tissue was removed. Once all the flesh was completely removed, the skulls were air-dried. The skulls were labelled and kept in specimen bottles with some naphthalene to avoid attacks by insects and prevent damage.

### 3.5.3 Conventional morphometrics

The cranial and mandible characters were measured for the 130 skulls of *R. rattus* using the digital vernier caliper (Mitutoyo Co., Kanagawa, Japan) and values were rounded to the nearest 0.01 mm. The exact biological age of the specimens could not be identified in this study since these rats were wild-caught during trapping sessions. Hence, age classes were described on the basis of molar wear stages following Ben Faleh et al. (2013) and interpreted as (Figure 3.2): Stage C0: No upper M3 is erupted; Stage C1: Upper M3 erupted but not worn; Stage C2: Cusps still visible on all molars and link between first and second lobe of the upper M3 is very narrow; Stage C3: The longitudinal link between first and second lobe of upper M3 is larger and generally wider; Stage C4: The upper M3 displays nearly the total fusion of the first and second lobe of the longitudinal link that is very wide but it remains visible on the other molars cusps; Stage C5: No more cusps were visible and the longitudinal link is so wide that lobes appeared fused on all the molars.

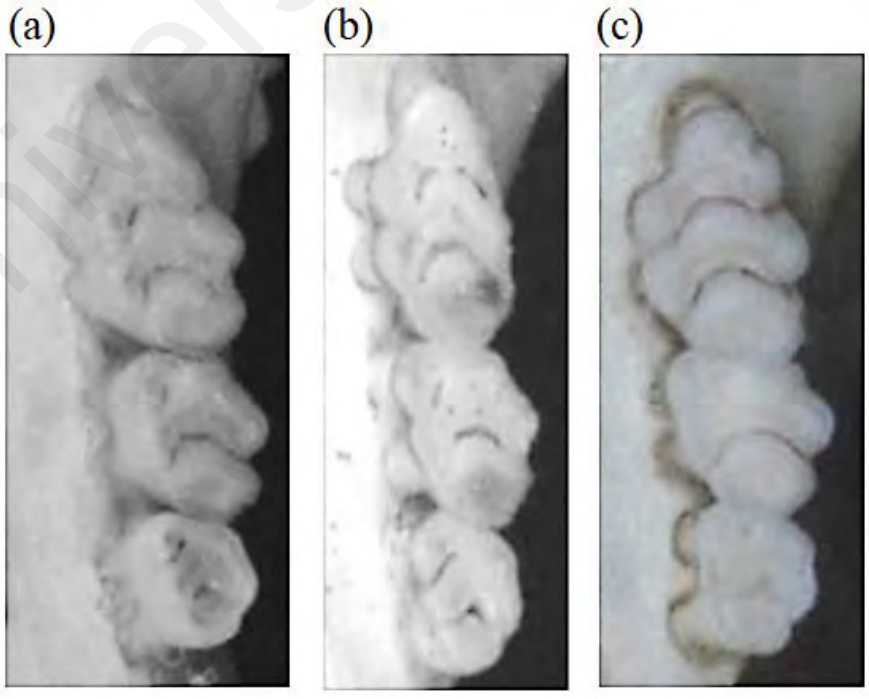


**Figure 3.2:** Different age classes of *R. rattus* described on the basis of molar wear stages (M3) ─ (a-c) corresponding to stage C2, C3, and C4, respectively.

Twenty quantitative characters of rat skulls were measured based on four different parts of measurements referring to Musser & Newcomb (1983) and Musser et al. (2006) (Figure 3.3): i) dorsal – occipitonasal length, or the greatest length of the skull (ONL), length of rostrum (LR), breadth of rostrum (BR), zygomatic breadth (ZB), breadth of braincase (BBC), interorbital breadth (IB); ii) ventral – length of diastema (LD), length of incisive foramina (LIF), breadth of incisive foramina (BIF), breadth of first upper molar (BM1), length of bony plate (LBP), length of auditory bulla (LB), postpalatal length (PPL), breadth of mesopterygoid fossa (BMF), breadth across palate at first molars (BBP); iii) lateral – breadth of zygomatic plate (BZP), crown length of maxillary molar row (CLM1-3), height of braincase (HBC); and iv) mandible – length of mandible (ML), length of mandible toothrow (M1-M3).

### 3.5.4 Statistical analysis

Statistical analyses were executed using Morpho (Schlager, 2017), maps (Richard et al., 2018), MASS (Venables & Ripley, 2002), and mapdata (Kahle & Wickham, 2013) packages in R version 3.5.1 (R Core Team, 2015). To avoid multicollinearity, the correlation matrix was computed and the highly correlated variables (correlation coefficient above 0.85) such as LR, IB, BIF, PPL, ZB, and BR were removed, and the morphometric variables examined reduced to 14 variables.

The analysis of variance (ANOVA) and multivariate analysis of variance (MANOVA) were conducted to test the influence of sexual dimorphism and age on the size of the skull. Regressions were performed on the log of skull size versus geographical factors such as latitude, longitude, and precipitation for males and females separately since sexual dimorphism was present.

The significance of differences between males and females were investigated using t-test for every craniodental measurement considered. The ‘a posteriori’ Scheffé test (Scheffé, 1959) was performed as a post hoc procedure to study the significance of difference in ages for each of the 14 morphometric variables. General trends of the skull sizes were studied via principal component analysis (PCA) for the 14 standardised craniodental measurements of males and females separately due to the presence of sexual dimorphism. In addition, the morphometric difference of the specimens was assessed via linear discriminant analysis (LDA) where the specimens were considered as clades.

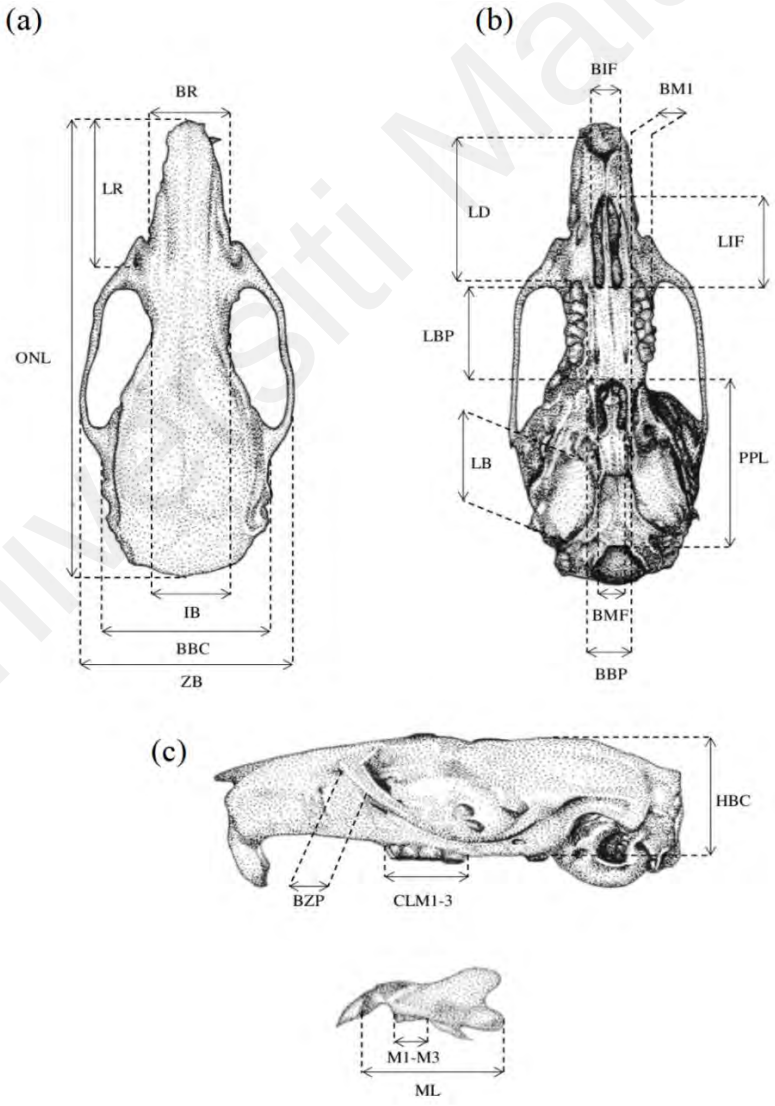


**Figure 3.3:** Craniodental measurements of *R. rattus* based on (a) dorsal, (b) ventral, and (c) lateral views used in morphometric analyses.

### 3.6 Prevalence and Genetic Divergence of *Leptospira* spp. in *Rattus rattus*

Since rats are often related to leptospiral transmission, a total of 130 individuals of *R. rattus* trapped throughout this study (as listed in Table 3.1) were exploited to presume the prevalence and genetic divergence of *Leptospira* spp. in *R. rattus*. The genomic DNA from both kidney and liver tissues of all the 130 individuals of *R. rattus* were extracted using the modified innuPREP DNA Mini Kit (Analytik Jena). The procedures for extraction of the genomic DNA is described in section 3.3.1. The DNA quality and concentration were also checked following the procedures mentioned in section 3.3.2.

PCR was performed on both kidney and liver tissues of each individual using two sets of established primers to amplify 412 and 240 bp fragment correspond to 16S rRNA and LipL32 genes, respectively. The 16S rRNA primer is a *Leptospira* genus-specific primer, while the LipL32 is mainly for the pathogenic *Leptospira* spp. The list of primers used to presume the prevalence and genetic divergence of *Leptospira* spp. in *R. rattus* is shown in Table 3.5. Three reference strains of *L. interrogans* serovar Lai, *L. interrogans* serovar Copenhageni, and *L. borgpetersenii* serovar Javanica serve as positive controls, while bacteria of *Escherichia coli* acts as the negative control. Both control groups were used along in every PCR screening to eliminate most potential confounding results. The PCR profile and the PCR chemical components used for leptospiral DNA screening is shown in Table 3.6 and Table 3.7, respectively. The PCR protocol designed for both 16S rRNA and LipL32 genes were similar.

**Table 3.5:** Details of primers used to presume the prevalence and genetic divergence of *Leptospira* spp. in *R. rattus*.

| Genes | Primers | Nucleotide sequence (5'-3') | Fragment Length (bp) | Annealing Temperature (°C) | Original Publication |
|---|---|---|---|---|---|
| 16S rRNA | Lep 1 | GGAACTGAGACACGGTCCAT | 412 | 58 | Tansuphasiri et al. (2006) |
| | Lep 2 | GCCTCAGCGTCAGTTTTAGG | | | |
| LipL32 | LipL32-2F | TGGCTATCTCCGTTGCACTC | 240 | 58 | Backstedt et al. (2015) |
| | LipL32-2R | CCCATTTCAGCGATTACGGC | | | |

**Table 3.6:** PCR profile used for leptospiral DNA screening.

| Stage | Temperature (°C) | Time | Number of cycles |
|---|---|---|---|
| Pre-denaturation | 94 | 4 minutes | 1 |
| Denaturation | 94 | 30 seconds | 35 |
| Annealing | 58 | 30 seconds | |
| Extension | 72 | 30 seconds | |
| Post extension | 72 | 10 minutes | 1 |
| Hold | 10 | ∞ | |

**Table 3.7:** PCR chemical components used for leptospiral DNA screening.

| Components | Volume (µl) | Final Concentration |
|---|---|---|
| 2 X Power Taq Master Mix (mixture of DNA polymerase, reaction buffers, $MgCl_2$, and dNTPs) | 12.5 | 1 X |
| Ultrapure water | 10.5 | - |
| Forward primer (10.0 µM) | 0.5 | 0.2 µM |
| Reverse primer (10.0 µM) | 0.5 | 0.2 µM |
| DNA template | 1.0 | ~ ng |
| Total | 25.0 | - |

The PCR products underwent electrophoresis using 2% (w/v) agarose gel in 1 X TAE buffer, as stated in section 3.3.4. Purification of PCR products were also done using the innuPREP PCRpure Kit (Analytik Jena), as indicated in section 3.3.5. The positive

screening PCR product that has undergone a purification process were continued with the nucleotide sequencing, together with a set of forward and reverse primers.

DNA sequencing results obtained were proofread by checking the chromatogram file for confirmation, to ensure there were no errors on bases interpretation done by the sequencing machines. Both aligned forward and reverse sequences were edited using MEGA7. To ensure the targeted species locus sequences were obtained, the edited sequences were validated using sequence similarity searches using GenBank BLASTn application. For protein coding gene of LipL32 gene, the stop codon (*) was checked in the translated protein sequences.

### 3.6.1 Sequence and phylogenetic analysis

The 16S rRNA gene from 22 *Leptospira* spp. were used as reference sequences with the GenBank accession numbers as follows: FJ154562, JQ988846, DQ848350, JQ988861, KP739780, AY631880, DQ483058, AY631883, AB279549, AB758746, KT338879, EF612284, EF025496, AY796065, AY631896, AY631885, AB721966, AY631878, NR115294, AF157070, NR115297, NR115293, and AY631879. A single 16S rRNA sequence of *Leptonema illini* (AY714984) was used as an outgroup to root the phylogenetic tree. For LipL32 gene, only the sequences of *L. interrogans* (KY356961) and *L. borgpetersenii* (KT338940 and KT338941) were included. The list of reference sequences used in the analysis is shown in Appendix C. The nucleotide sequences of leptospires acquired from GenBank and those obtained from the kidney, liver, and positive control samples after the editing process were aligned accordingly using the ClustalW multiple alignment algorithms of MEGA7. The sequences were analysed using MEGA7 software to verify the percentage of nucleotide compositions, sequence divergences, transition to transversion ratio, and pairwise analysis.

Phylogenetic tree was built in MEGA7 using distance-based neighbor-joining (NJ) tree. Kimura 2-Parameter model was used and tested with a bootstrap value of 1000. Evaluations of genetic differentiation criteria such as nucleotide diversity ($\pi$), net nucleotide divergence (*Da*), haplotype diversity (*Hd*), average number of nucleotide differences (*k*), nucleotide subdivision (*Nst*), population subdivision (*Fst*), the number of migrants per generation (*Nm*), and segregating sites were estimated from the DNA dataset using DNASP v5. The genetic divergence was analysed using Arlequin Version 3.5. AMOVA was executed using *F*-statistic to investigate the amount of genetic variability segregated between and within species. The significance of variance components was assessed using 1000 permutations. Network 5.0.0.1 was used to generate a minimum-spanning network (MSN) to illustrate the haplotype relationships.

## 3.7 Transcriptome Study of *Rattus rattus* Infected with *Leptospira* spp.

After screening for leptospiral infections using PCR, one control and one infected with *Leptospira* spp. of both kidney and liver samples were used as references in transcriptome resequencing. The total ribonucleic acid (RNA) of all the four reference samples were isolated, and the complementary DNA (cDNA) library construction and sequencing were conducted using BGISEQ-500 platform (Beijing Genome Institute, Shenzhen, China).

### 3.7.1 Isolation of total RNA

Total RNA from liver and kidney tissues were isolated using *TransZol* Up Plus RNA Kit (TransGen Biotech Co., Beijing). Prior to the isolation procedure, working surfaces and other needed apparatus such as pipettes were treated with 10% (v/v) chlorox, 70% (v/v) ethanol, and RNase Zap in order to promote a RNase-free environment. Besides, all the necessary equipment such as mortars, pestles, forceps, surgical scissors, and stainless-

steel lab spoons which had been autoclaved were pre-chilled by placing them in a laboratory freezer for one hour.

A starting materials of 50 mg tissue samples were weighed and the frozen samples were quickly transferred into a sterilised and pre-chilled mortar with liquid nitrogen. The samples were grinded thoroughly to powder. Liquid nitrogen was added whenever needed to ensure complete grind of samples. Incomplete grind can affect RNA yield and quality. The tissue powder was transferred to a 1.5 ml microcentrifuge tube. One ml of *TransZol* Up was added to per 50-100 mg tissue. Tissue samples were homogenised with a homogeniser and were repeatedly pipetted up and down. The samples were incubated at room temperature for five minutes. Next, 0.2 ml of chloroform were added for per ml *TransZol* Up. The tube was vigorously shaken for 30 seconds, and incubated at room temperature for three minutes. The samples were then centrifuged at 10000 relative centrifugal force (rcf) for 15 minutes at 4 °C. The mixture separates into a lower pink organic phase, an interphase, and a colourless upper aqueous phase which contains the RNA. The volume of the aqueous upper phase is 50%-60% volume of *TransZol* Up reagent. The colourless, upper phase containing the RNA were transferred to a fresh RNase-free tube (to avoid DNA contamination from interphase, a portion of aqueous phase can be left). Equal volume of absolute ethanol was added (precipitates may be seen at this state). The samples were mixed gently by inverting tube.

Then, all following centrifugation steps were carried out at room temperature. The resulting solution and precipitates were transferred together to a spin column. The samples were centrifuged at 12000 rcf for 30 seconds at room temperature. The flow-through were discarded (this step was repeated if the volume of lysate is more than the spin column can hold). For cleaning purposes, 500 µl of CB9 cleaning buffer were added

to the spin column. The samples were centrifuged at 12000 rcf for 30 seconds. The flow through were discarded and the step was repeated once. For washing purposes, 500 µl of WB9 washing buffer were added into the spin column. The samples were then centrifuged at 12000 rcf for 30 seconds at room temperature. The flow through were discarded and the step was repeated once. The samples were later centrifuged at 12000 rcf for two minutes at room temperature in order to completely remove remaining ethanol and then the column matrix was air-dried for ten minutes. The spin column was placed into a clean 1.5 ml of RNase free tube. Finally, 50 µl of RNase-free Water were added into the spin column matrix and incubated at room temperature for one minute. The samples were finally centrifuged at 12000 rcf for one minute to elute RNA. The isolated RNA was then stored at -80 °C.

### 3.7.2 Determination of RNA quality and concentration

The quality of isolated RNA has been tested using 1% (w/v) agarose gel electrophoresis. Five µl of RNA and two µl of loading dye were mixed and loaded into wells of 1% (w/v) agarose gel, added with gel stain. Five µl of 1 kb RNA ladder was loaded into the first well to quantify the RNA. Electrophoresis were done in 1 X TAE buffer (0.04 M Tris acetate pH 8.5 and 0.001 M EDTA) at a voltage of 100 V and current of 200 mA for 30 minutes. Alpha Imager Gel Documentation System (Siber Hegner, Germany) was used to view the digital image of agarose gel under ultraviolet (UV) light in transilluminator machine. Two obvious band sized 2 kb and 4 kb shows the good quality of isolated RNA. A NanoDrop™ 2000 Spectrophotometer (Thermo Fisher Scientific, USA) was used to check the concentration of the isolated RNA. The purity of the isolated RNA was determined by the A260/A280 nm and A260/A230 nm. The integrity of RNA was checked using Agilent 2100 Bioanalyzer (Agilent Technologies, USA) to quantify the RNA concentration, RNA Integrity Number (RIN) value, 28S/18S,

and the fragment length distribution. RNA integrity number (RIN) of > 8.0 is considered good and shows the RNA still intact and have not yet degraded.

### 3.7.3 RNA sequencing (RNA-Seq) library construction

The first step in the transcriptome workflow involves purification using poly-T oligo-attached magnetic beads of the poly-A containing messenger RNA (mRNA) molecules. The mRNA is fragmented into small pieces after purification using divalent cations at elevated temperatures. Using reverse transcriptase and random primers, the cleaved RNA fragments are copied into first strand cDNA. This is followed by second strand cDNA synthesis using DNA Polymerase I and RNase H. These cDNA fragments then have the addition of a single 'A' base and subsequent ligation of the adaptor. PCR amplification purifies and enriches the products. Then, Qubit quantified the PCR yield and pooled samples to create a single strand DNA circle (ssDNA circle) that gave the final library. To enlarge the fluorescent signals at the sequencing process, DNA nanoballs (DNBs) were generated with the ssDNA circle by rolling circle replication (RCR). The DNBs were loaded into the patterned nanoarrays and for the following data analysis study, pair-end reads of 100 bp were read through on the BGISEQ-500 platform. For this step, the BGISEQ-500 platform uses the Combinational Probe-Anchor Synthesis Sequencing Method to combine DNA nanoball-based nanoarrays and stepwise sequencing. The transcriptome experimental workflow was described in Huang et al. (2017).

### 3.7.4 Bioinformatics workflow

The low-quality reads were first filtered to get the clean reads. Then, these clean reads were mapped to reference genome, followed by novel gene prediction, calling for Single Nucleotide Polymorphism (SNPs) and insertion or deletion (INDELs), and detection of

gene-splicing. Finally, differentially expressed genes (DEGs) were identified between samples, and clustering analysis and functional annotations were conducted.

#### 3.7.4.1 Sequencing reads filtering

An internal software of SOAPnuke were used to filter reads. Reads with adaptors, reads in which unknown bases (N) are more than 5%, and low-quality reads were removed (low-quality reads have been defined as the percentage of base which quality is less than 15 is greater than 20% in a read). The remaining reads are referred to as clean reads after filtering and stored in FASTQ format (Cock et al., 2010).

#### 3.7.4.2 Genome mapping

*Rattus norvegicus* was used as a reference genome in this study. Hierarchical Indexing for Spliced Alignment of Transcripts (HISAT) was used for the mapping step. HISAT is an analysis software which is faster, sensitive, and high accuracy. The mapping method was demonstrated in Kim et al. (2015).

#### 3.7.4.3 Novel transcripts prediction

StringTie (Pertea et al., 2015) was used to reconstruct transcripts and Cuffcompare was used from Cufflinks tools (Trapnell et al., 2012) to compare reconstructed transcripts with reference annotations. Class code types ‘u’, ‘i’, ‘o’, and ‘j’ (u-unknown, intergenic transcript; i-a transfrag falling entirely within a reference intron; o-generic exonic overlap with a reference transcript; and j-potentially novel isoform (fragment): at least one splice junction is shared with a reference transcript) were selected as novel transcripts. To predict the coding potential of novel transcripts, Coding Potential Calculator (CPC) (Kong et al., 2007) were used. To obtain a complete reference, the coding novel transcripts were merged with reference transcripts, and the downstream analysis was

based on this reference. StringTie is a much faster and accurate software for transcriptome assembly, compared to Cufflinks software (Kim et al., 2015). The pipeline for transcriptome assembly based on reference was illustrated in Martin & Wang (2011).

#### 3.7.4.4 SNP and INDEL detection

With genome mapping result, Genome Analysis Toolkit (GATK) (McKenna et al., 2010) was used to identify the SNP and INDEL for each sample. The final SNP and INDEL were obtained in VCF format, after filtering out the unreliable sites. The pipeline for detection SNP and INDEL in RNA-Seq was described in McKenna et al. (2010).

#### 3.7.4.5 Differentially splicing gene detection

Differential isoform relative abundance is important to be distinguished from differential isoform expression. Changes in relative isoform abundance, irrespective of the change in expression, indicate a mechanism related to splicing. On the other hand, measurable changes in isoform expression across samples can occur, without necessarily changing relative abundance, which may indicate a mechanism related to transcription. rMATS (Shen et al., 2014) was used to detect differentially splicing genes (i.e. differential isoform relative abundance between samples), a computational tool for detecting differential alternative splicing events from RNA-Seq data, calculating isoform inclusion and isoform skipping. The statistical model of Multivariate Analysis of Transcript Splicing (MATS) calculates the *P*-value and false discovery rate (FDR) which is the difference between two conditions in the isoform ratio of a gene. A gene with FDR < 0.05 is defined as a significant differentially splicing gene (DSG).

#### 3.7.4.6 Gene expression analysis

Bowtie2 mapped clean reads to reference (Langmead & Salzberg, 2012), and RSEM calculated the level of gene expression (Li & Dewey, 2011). RSEM is a software package to estimate gene and isoform expression levels from RNA-Seq data. Then, Pearson correlation between all samples was calculated using cor, hierarchical clustering between all samples was performed using hclust, and the diagrams were drawn using ggplot2 with functions of R.

#### 3.7.4.7 Differentially expressed gene detection

PossionDis was used to detect Differentially Expressed Genes (DEGs). PossionDis is based on the Poisson distribution, performed as described by Audic & Claverie (1997).

#### 3.7.4.8 Hierarchical clustering analysis of DEGs

Hierarchical DEGs clustering was conducted using pheatmap, a function of R. The intersection and union DEGs were performed between them, for cluster more than two groups, respectively.

#### 3.7.4.9 Gene ontology analysis of DEGs

With the result of the Gene Ontology (GO) annotation, DEGs were classified according to the official classification, and functional enrichment of GO was also performed using phyper, a function of R. Then, for each *P*-value, false discovery rate (FDR) was calculated, generally the terms that are defined as significantly enriched by FDR not greater than 0.01.

#### 3.7.4.10 Pathway analysis of DEGs

With the annotation result of the Kyoto Encyclopedia of Genes and Genome (KEGG), DEGs were classified to the official classification, and functional enrichment of the pathway was also performed using phyper, a function of R. Then, for each *P*-value, false discovery rate (FDR) was calculated, generally the terms that are defined as significantly enriched by FDR not greater than 0.01.

### 3.7.5 Reverse transcription and first strand complementary DNA (cDNA) synthesis

Since the infection of leptospirosis mostly affected both kidney and liver, validation of the BGISEQ-500 platform sequencing results were performed on three biological replicates, two cohorts of rodents: infected and non-infected with *Leptospira* spp., on both kidney and liver tissues using cDNA. After total RNA isolation, cDNA for each sample were synthesised from an equal amount of total RNA (~ five µg) using *TransScript*® One-Step gDNA Removal and cDNA Synthesis SuperMix kit following the manufacturer's protocol (TransGen Biotech Co., Beijing). For the first step, five µl of RNA template with concentration of 1000 ng/µl were added into two µl of RNase-free water and one µl of oligo $(dT)_{18}$ primer. The mixture was then incubated at 65 °C for five minutes – a PCR thermocycler was used for this incubation step. Then, the mixture was put on ice for two minutes. For the second step, one µl of genomic DNA (gDNA) remover, ten µl of reaction mix, and one µl of enzyme mix were added into the mixtures. Finally, the mixtures were incubated at three different states; 25 °C for ten minutes, 42 °C for 15 minutes, and to inactivate enzyme at 85 °C for five seconds, subsequently. The final cDNA product was stored at -20 °C. The chemicals components for first strand cDNA synthesis and reverse transcription profile were described in Table 3.8 and Table 3.9, respectively.

**Table 3.8:** First strand cDNA synthesis components.

| Components | Volume (µl) | Final Concentration |
|---|---|---|
| Total RNA template | 5 | ~ 5000 ng |
| Oligo $(dT)_{18}$ | 1 | 0.5 µg (microgram) |
| 2 X TS Reaction Mix | 10 | 1 X |
| *TransScript*® RT/RI Enzyme Mix | 1 | |
| gDNA Remover | 1 | |
| RNase-free Water | 2 | |
| Total | 20 | |

**Table 3.9:** Reverse transcription profile.

| Stage | Temperature (°C) | Time | Remarks |
|---|---|---|---|
| Incubation | 65 | 5 minutes | Mixture of RNA template, RNase-free water, and oligo $(dT)_{18}$ primer |
| Incubation (on ice) | 0 | 2 minutes | |
| Annealing | 25 | 10 minutes | Additional of gDNA remover, reaction mix, and enzyme mix |
| Reverse-transcription step | 42 | 15 minutes | |
| Inactivation of reaction | 85 | 5 seconds | |

### 3.7.6 Primer design and optimisation for quantitative RT-PCR (qRT-PCR) amplification

Validation of the BGISEQ-500 platform sequencing results involved the choosing of six differentially expressed *R. rattus* immune genes; tumor necrosis factor-α (TNF-α), interleukin (IL)-1β, IL-6, IL-10, macrophage inflammatory protein-1α (MIP-1α/C-C-type chemokine ligand 3, CCL3), and gamma interferon (IFN-γ)-inducible protein 10 kDa (IP-10/C-X-C-type chemokine ligand 10, CXCL10) for quantitative RT-PCR analysis (qRT-PCR).

Specific primers for qRT-PCR amplification of all the six genes were designed based on the conserved regions of each gene correspond to the several published mRNA sequences of brown rat, *Rattus norvegicus* and house mouse, *Mus musculus* available at the GenBank (https://www.ncbi.nlm.nih.gov), using the Primer3 software (http://primer3.ut.ee/). Several parameters were taken into consideration, including melting temperature, hairpin loop, self-complementary, and GC content. All the primers were produced by Integrated DNA Technologies (IDT, Singapore). Primers optimisation were done prior to use in qRT-PCR by identifying the optimal annealing temperature of the primers. A PCR gradient was carried out using *R. rattus* cDNA as template. The range of temperature was started from 5 °C below the calculated temperature of the primer melting temperature (*Tm*) to 5 °C above the calculated temperature of the primer melting temperature. Primers sequences, expected fragment length, and melting temperature are shown in Table 3.10.

**Table 3.10:** Details of primers designed for qRT-PCR.

| Genes | Primers | Nucleotide sequence (5'-3') | Fragment Length (bp) | Melting Temperature (°C) |
|---|---|---|---|---|
| TNF-α | TNF-αF | CTCAGCCTCTTCTCATTCCTG | 198 | 56 |
| | TNF-αR | GGTTTGCTACGACGTGGGCTA | | |
| IL-1β | IL-1βF | GCTTCCTTGTGCAAGTGTCTGAAG | 183 | 56 |
| | IL-1βR | GAGATTTGAAGCTGGATGCTCTCA | | |
| IL-6 | IL-6F | ACAGAAGGAGTGGCTAAGGACCA | 128 | 56 |
| | IL-6R | GTCCACAAACTGATATGCTTAGGC | | |
| IL-10 | IL-10F | AATAACTGCACCCACTTCCCAGT | 166 | 56 |
| | IL-10R | AAGGCTTGGCAACCCAAGTAACC | | |
| MIP-1α | MIP-1αF | ACCAGCAGCCTTTGCTCCCA | 187 | 56 |
| | MIP-1αR | CTCAAGCCCCTGCTCTACAC | | |
| IP-10 | IP-10F | GCACCATGAACCCAAGTGCTG | 198 | 56 |
| | IP-10R | GTGGCAATGATCTCAACATGCG | | |
| β-actin | β-actinF | TTCGCCATGGATGACGATATCGCT | 139 | 56 |
| | β-actinR | ACCATCACACCCTGGTGCCTA | | |

PCR chemical components and PCR profile used for optimisation of designed primers are shown in Table 3.11 and Table 3.12, respectively. The PCR products were then run on 2% (w/v) agarose gel electrophoresis on 1 X TAE buffer, voltage of 80 V, current of 180 mA, for 45 minutes. Only the temperatures that signified a single band with approximate expected size without primer dimer were chosen as annealing temperature, to be used in qRT-PCR. Afterwards, the PCR products with respective annealing temperature chosen previously, underwent the nucleotide sequencing process. After sequenced, a BLAST check was conducted against the nucleotide database at NCBI in order to confirm the obtained sequence.

**Table 3.11:** PCR chemical components used for optimisation of designed primers.

| **Components** | **Volume (µl)** | **Final Concentration** |
|---|---|---|
| 2 X Power Taq Master Mix (mixture of DNA polymerase, reaction buffers, $MgCl_2$, and dNTPs) | 12.5 | 1 X |
| Distilled water ($ddH_2O$) | 10.5 | - |
| Forward primer (10.0 µM) | 0.5 | 0.2 µM |
| Reverse primer (10.0 µM) | 0.5 | 0.2 µM |
| cDNA template | 1.0 | ~ ng |
| Total | 25.0 | - |

**Table 3.12:** PCR profile used for optimisation of designed primers.

| **Stage** | **Temperature (°C)** | **Time** | **Number of cycles** |
|---|---|---|---|
| Pre-denaturation | 94 | 4 minutes | 1 |
| Denaturation | 94 | 30 seconds | 35 |
| Annealing | 51-61 | 30 seconds | |
| Extension | 72 | 30 seconds | |
| Post extension | 72 | 10 minutes | 1 |
| Hold | 10 | ∞ | |

### 3.7.7 Quantification of mRNA expression on healthy and leptospiral-infected *Rattus rattus* by qRT-PCR

The expression of tumor necrosis factor-α (TNF-α), interleukin (IL)-1β, IL-6, IL-10, macrophage inflammatory protein-1α (MIP-1α/C-C-type chemokine ligand 3, CCL3), and gamma interferon (IFN-γ)-inducible protein 10 kDa (IP-10/C-X-C-type chemokine ligand 10, CXCL10) in healthy and unhealthy kidney and liver tissues were measured by quantitative real-time PCR (qRT-PCR) using a real-time PCR machine (Mx3005P qPCR System, Agilent Technologies, Inc., Santa Clara, CA, USA). SYBR Green was used as fluorescent dye for detection and ROX was the reference dye used to normalise fluorescent reporter signal. In addition, the housekeeping gene of Beta-actin (β-actin) was used as an internal control to normalise the gene expression. Each target gene along with reference gene and no template control (NTC) (primer control) was analysed in one plate for qRT-PCR. Moreover, each sample and NTC was analysed in three technical replicates (triplicates) to normalise the result.

The SensiFAST™ SYBR® Lo-ROX Kit (Bioline USA Inc.) were used for all qRT-PCR reactions. The components and their corresponding volume are shown in Table 3.13. Prior to perform qRT-PCR reactions, the synthesised cDNA of both healthy and infected kidney and liver tissues was diluted to 10-fold in nuclease-free water. Briefly, the total reaction volumes were 20 µl, included four µl of diluted cDNA (approximate concentration was 100 ng/µl), 0.4 µl of each forward and reverse primer (10 µM), ten µl of 2 X SensiFAST SYBR® Lo-ROX Mix, and 5.2 µl of nuclease-free water. The procedure of mixing the chemical components was done in a biosafety cabinet to avoid contamination.

**Table 3.13:** The chemical components used in qRT-PCR amplification.

| Components | Volume (µl) | Final Concentration |
|---|---|---|
| 2 X SensiFAST SYBR® Lo-ROX Mix | 10.0 | 1 X |
| 10 µM forward primer | 0.4 | 200 nM (nanomolar) |
| 10 µM reverse primer | 0.4 | 200 nM |
| Nuclease-free water | 5.2 | - |
| cDNA template | 4.0 | ~ 100 ng |
| Total | 20.0 | - |

For plate setup, the area proportional to the wells of the tubes were selected on the computer screen. ROX and FAM were selected for fluorescence dyes, and ROX was selected as the reference dye. Next, the thermal profile was set up. The thermal cycling conditions were as follows: step 1, 95 °C for two minutes; step 2, 40 cycles of 95 °C for five seconds and 60 °C for 30 seconds. The combined annealing/extension temperature at 60 °C was ideal for all the designed primers. Lastly, a melting curve analysis was performed at the end of the run following the real time instrument manufacturer's instructions. The specificity of the qRT-PCR amplification was verified through a melting curve analysis by generating a dissociation curve. Table 3.14 shows the thermal profile of the SYBR Green real-time PCR programme.

**Table 3.14:** Thermal profile of the SYBR Green real-time PCR.

| Stage | Temperature (°C) | Time | Number of cycles |
|---|---|---|---|
| Polymerase activation | 95 | 2 minutes | 1 |
| Denaturation | 95 | 5 seconds | 40 |
| Annealing/Extension | 60 | 30 seconds | |
| Melting curve analysis | 95 | 1 minute | 1 |
| | 55 | 30 seconds | 1 |
| | 95 | 30 seconds | 1 |
| Hold | 10 | ∞ | |

### 3.7.8 qRT-PCR statistical analysis

Cycle threshold ($C_t$) values were calculated by the inbuilt Applied Biosystems® 7500 SDS software (Thermo Fisher Scientific, USA). $C_t$ value of the PCR amplification curve of the target gene was compared with the $C_t$ value of the internal reference gene to obtain $\Delta C_t$, which was used to conduct relative quantitative analysis on the expression level of the six immune genes by using the comparative $C_t$ method (also referred to as the $2^{-\Delta\Delta Ct}$ method) (Schmittgen & Livak, 2008). Gene expression of target genes was estimated, and the fold change was calculated using the formula below. In addition, the gene expression calculation was done using Microsoft Excel 2011.

Expression ratio: $2^{-\Delta\Delta Ct}$ = Normalised expression ratio (Schmittgen & Livak, 2008)

$$\Delta\Delta C_t = \Delta C_{t\,(\text{sample})} - \Delta C_{t\,(\text{calibrator})}$$

$$= \text{treatment}\,(C_{t\,(\text{gene of interest})} - C_{t\,(\text{reference gene})}) -$$

$$\text{control}\,(C_{t\,(\text{gene of interest})} - C_{t\,(\text{reference gene})})$$

# CHAPTER 4: RESULTS

In this study, the genomic DNA that have been extracted using commercial kit were used for PCR amplification. Figure 4.1 illustrates the extraction result of genomic DNA from several samples of *Rattus rattus* used in this study.

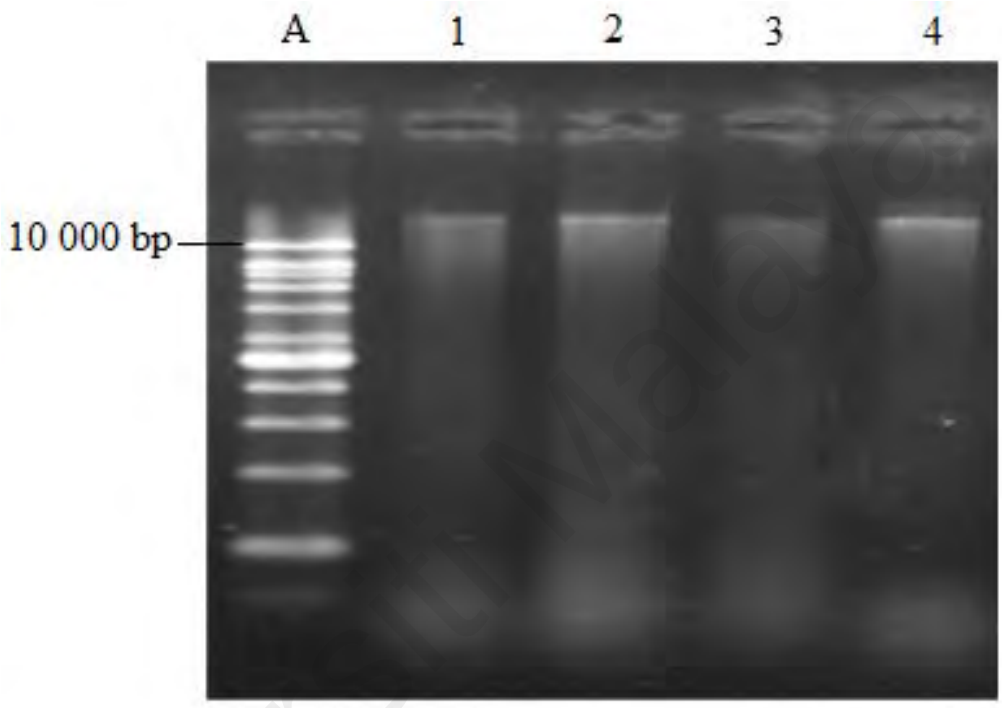


**Figure 4.1:** Extraction results of genomic DNA from several samples of *Rattus rattus*, on 1% agarose gel. (Well A = 1 kb DNA ladder; Well 1-4 = *R. rattus* samples).

## 4.1 Species Diversity of *Rattus* (Rodentia: Muridae) on the Malay Peninsula and Their Phylogenetic Relationships Inferred from Complete Mitochondrial Cytochrome b Gene Sequences

All genomic DNA from the 35 individuals of *Rattus* spp. were successfully extracted, amplified, and sequenced (GenBank accession numbers: MH818016-MH818047, MK754068-MK754070). A total of 1140 bp of Cyt b gene sequences were obtained after the final editing stage of all genetic samples used in this study. No insertions, deletions, or stop codons were discovered. Analysis of all 114 sequences of genus *Rattus* shows a total of 831 (72.9%) conserved sites and 309 (27.1%) variable sites. Parsimony-

informative sites and singleton sites showed 290 bp and 19 bp, respectively. The finding of 25.4% informative sites suggested that Cyt b as a high resolution and informative marker in the mitochondrial DNA (mtDNA). The nucleotide composition analysis of all nucleotide sequences shows that Adenine (A) has the highest composition in the sequences with the average of 30.8%, followed by Cytosine (C) with the average of 28.7%, Thymine (T) with the average of 28.1%, and Guanine (G) with the average of 12.4%.

To resolve the complexity within genus *Rattus* in Peninsular Malaysia, the samples were pooled together with an additional of several representatives from various countries. The grouping of the samples was done following previous studies (Aplin et al., 2011; Pagès et al., 2010; Robins et al., 2007; Musser & Carleton, 2005; Aplin et al., 2003) with a slight modification (see Table 4.1). The results indicated a very informative resolution within this genus. Nine main groups were derived and interpreted as: *R. rattus* lineage I, *R. rattus* lineage III, *R. rattus* lineage IV, *R. tanezumi*, *R. tiomanicus*, *R. baluensis*, *R. exulans*, *R. argentiventer*, and *R. norvegicus*. The *R. rattus* complex portrayed a broad biogeographic and phylogeographic patterns within the lineages. Lineage I shows the wide distribution, with representation from various continents; e.g. Asia, Europe, Oceania, Africa, North America, and South America. Lineage III was discovered in rats from Nepal and Pakistan, while lineage IV includes the rats from Southeast Asia (i.e. Peninsular Malaysia, Philippines, southern Thailand, Indonesia, Vietnam, Cambodia, Laos), and Sri Lanka. Interestingly, samples of *R. rattus* from Sri Lanka were separated in two different lineages, lineage I and lineage IV. In this study, *R. tanezumi*, *R. tiomanicus*, *R. baluensis*, *R. exulans*, *R. argentiventer*, and *R. norvegicus* were treated in their specific taxa.

**Table 4.1:** Correlation between the mtDNA lineage terminology and taxonomic usage in this study with that implemented in several previous studies.

| **mtDNA lineage** | **Associated karyotypes** | **Aplin et al. (2003)** | **Musser & Carleton (2005)** | **Robins et al. (2007)** | **Pagès et al. (2010)** | **Aplin et al. (2011)** | **This study** |
|---|---|---|---|---|---|---|---|
| Rrc LI | Oceanian *R. rattus* 2n = 38 | *R. rattus* | *R. rattus* | *rattus* clade | R1 = *R. rattus* | *R. rattus* Lineage I | *R. rattus* lineage I |
| Rrc LII | Asian *R. rattus* part 2n = 42 | *R. tanezumi* northern | *R. tanezumi* part | *tanezumi* clade | R2 = *R. tanezumi* | *R. rattus* Lineage II | *R. tanezumi* |
| Rrc LIII | Asian *R. rattus* part ? | Uncategorised | *R. tanezumi* part | Not sampled | Not sampled | *R. rattus* Lineage III | *R. rattus* lineage III |
| Rrc LIV | Asian *R. rattus* part 2n = 40 | *R. tanezumi* southern | *R. tanezumi* part | *diardi* clade | R3 | *R. rattus* Lineage IV | *R. rattus* lineage IV |
| Rrc LV | *R. losea* part 2n = 42 | *R. losea* western | *R. losea* part | Not sampled | R4 = *R. losea* | *R. sakeratensis* | Not sampled |
| Rrc LVI | *R. rattus tiomanicus* 2n = 42 | *R. tiomanicus* | *R. tiomanicus* | *R. tiomanicus* | R5 = *R. tiomanicus* | *R. tiomanicus* + *R. baluensis* | *R. tiomanicus* + *R. baluensis* |

Throughout this study, a total of six different species in the genus *Rattus* were identified in Peninsular Malaysia: black rat (*Rattus rattus cf*. lineage IV sensu Aplin et al., 2011), Asian house rat (*R. tanezumi*), brown rat (*R. norvegicus*), Pacific rat (*R. exulans*), ricefield rat (*R. argentiventer*), and Malaysian wood rat (*R. tiomanicus*). As suggested by Aplin et al. (2011), the black rat *Rattus rattus* sensu stricto from lineage I of the Indian holotype was not observed in this study. The results confirmed the occurrence of both *R. rattus cf*. lineage IV sensu Aplin et al., 2011 and *R. tanezumi* in Peninsular Malaysia, supported with 6.8% K2P pairwise genetic distance. These data are analysed in comparison with sequences from other vouchered specimens available in the literature. The species *R. tanezumi*, *R. norvegicus*, and *R. exulans* are widespread invasive and most are likely invasive on the Malay Peninsula. It is possible that there are only three species in the genus *Rattus* which are native to the Malay Peninsula: the Malaysian wood rat (*Rattus tiomanicus*), the ricefield rat (*R. argentiventer*), and an unnamed species (*Rattus rattus cf*. lineage IV sensu Aplin et al., 2011).

Single-nucleotide polymorphisms (SNP) analysis conducted on the Cyt b sequences excluding the outgroup revealed 309 segregating sites. Eighty haplotypes were defined for the genus *Rattus*. Minimum-spanning network (MSN) was generated with the haplotype data obtained to demonstrate the relationships of genus *Rattus* (Figure 4.2). The network analysis revealed that the haplotypes are unique to each species, as all of them lacked a common haplotype. Eighteen haplotypes were detected for *R. tanezumi* (H_1-H_18), 13 haplotypes for *R. rattus* lineage I (H_19-H_31), three haplotypes for *R. rattus* lineage III (H_32-H_34), and nine haplotypes for *R. rattus* lineage IV (H_35-H_43). Four haplotypes were identified from *R. tiomanicus* (H_44-H_47), three from *R. baluensis* (H_48-H_50), as well as six from *R. argentiventer* (H_51-H_56). Also, 11 haplotypes were seen for *R. exulans* (H_57-H_67) and 13 for *R. norvegicus* (H_68-H_80).

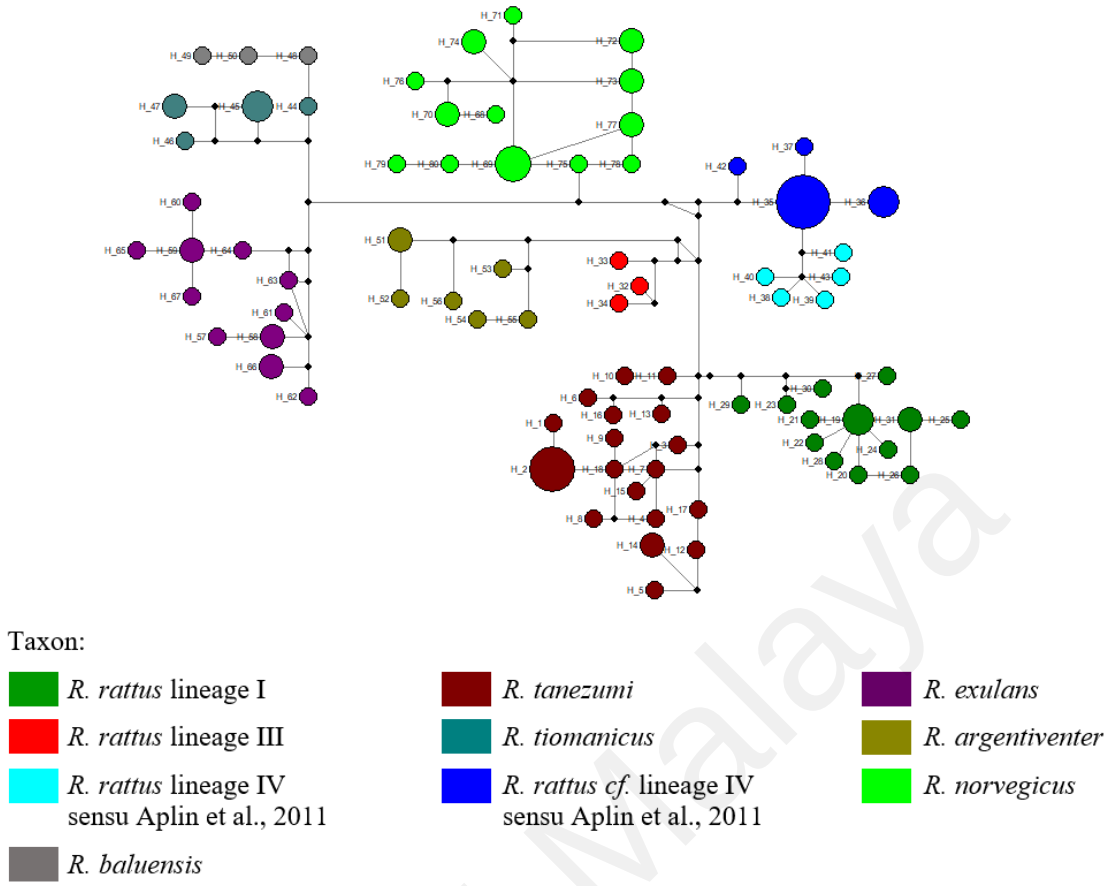


**Figure 4.2:** Minimum-spanning network of genus *Rattus* haplogroup. The circle size of inter-specific *Rattus* haplotypes is proportional to the frequency of the haplotype.

Each captured species in Peninsular Malaysia had two or three different haplotypes (Figure 4.3). The brown rat, black rat, Pacific rat, Asian house rat, ricefield rat, and Malaysian wood rat haplotypes matched haplotypes previously reported in these species in distant populations by 99.2%-100%. The brown rat shared a haplotype with China (Liu et al., 2017); the black rat shared a haplotype with Philippines (Lack et al., 2012); the Pacific rat haplotypes were 99.7% identical to animals in Papua New Guinea (Robins et al., 2008); the Asian house rat were 99.9% match to animals from Vietnam (Lack et al., 2012); the ricefield rat were 99.7% match to animals from Vietnam (Balakirev & Rozhnov, 2012); and the Malaysian wood rat haplotypes were 99.2% identical to animals in Thailand (Pagès et al., 2010). This haplotype sharing or very close relationship between haplotypes suggests a very recent common ancestry.

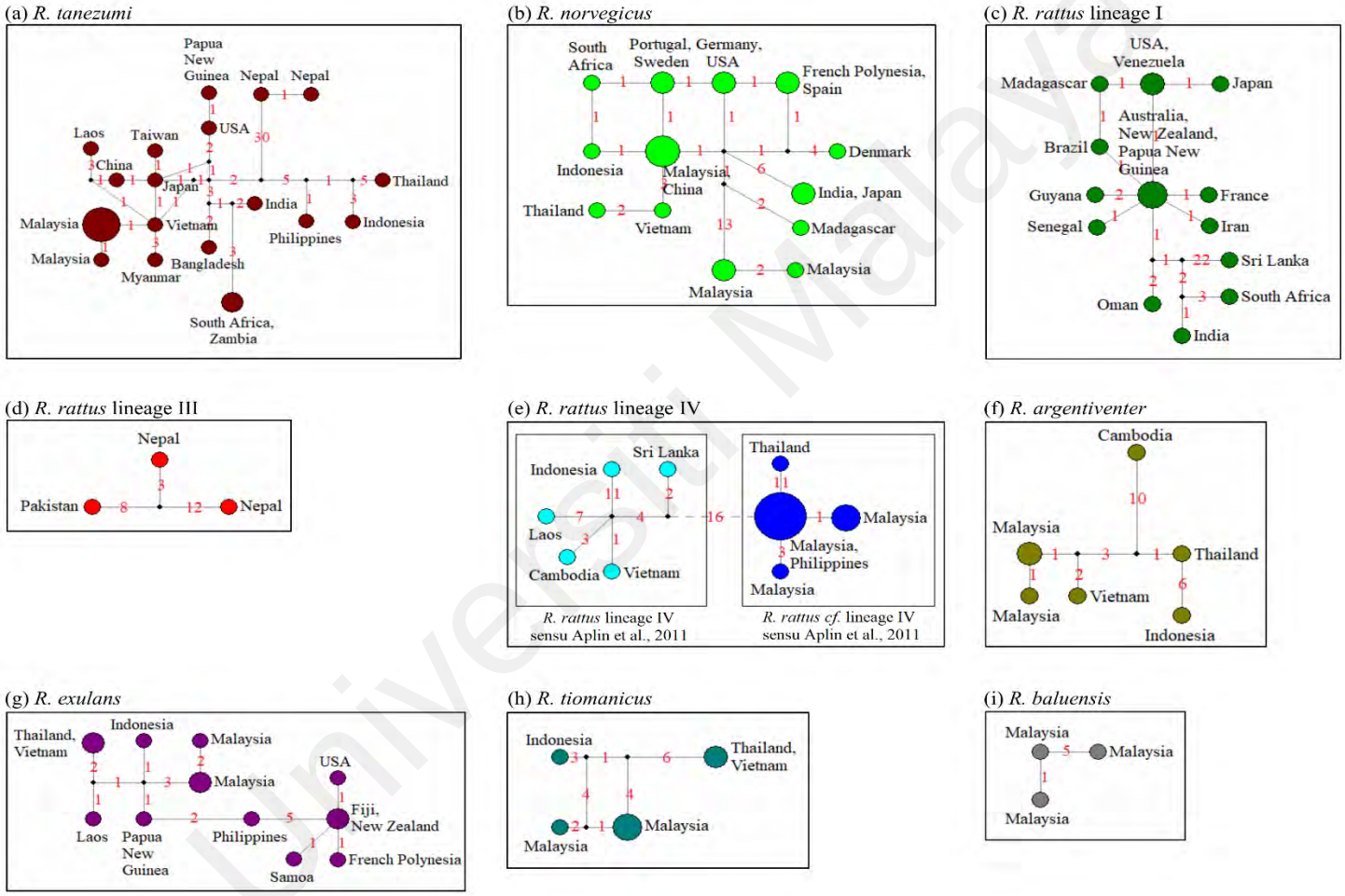


**Figure 4.3:** Minimum-spanning network of *Rattus* spp. haplogroup with specific geographic origin of the haplotypes. The number of mutations is indicated on each branch. The circle size of intra-specific *Rattus* haplotypes is proportional to the frequency of the haplotype.

The average percentages of pairwise genetic distances among *Rattus* species and outgroup calculated using K2P algorithm model are shown in Table 4.2. From the analysis, interspecific genetic distance revealed the closest relationships between *R. baluensis* and *R. tiomanicus* which is 1.9%. *Rattus tanezumi* group which is always associated with *R. rattus* group demonstrated higher interspecific genetic distance between them. The relationships of *R. tanezumi* with *R. rattus* lineage I, *R. rattus* lineage III, and *R. rattus* lineage IV are in the range of 3.9%, 4.6%-4.7%, and 5.7%-6.9%, respectively. However, *R. norvegicus* shows the most distant relationships from other species of genus *Rattus*, which is ranging from 11.5% to 13.6%. Species of genus *Rattus* exhibited the highest value of a pairwise distance with *L. sabanus* between 14.7%-17.2%, which were highly expected as they were separated in two different genera.

For intraspecific genetic relationships within *R. tanezumi*, the rats from Peninsular Malaysia group is very close in genetic distance to Asia, Africa, America, and Oceania groups, which is only 0.9%. In furtherance, the nearest relationships within *R. rattus* is between *R. rattus cf.* lineage IV sensu Aplin et al., 2011 (rats of Peninsular Malaysia, Philippines, and southern Thailand) to *R. rattus* lineage IV sensu Aplin et al., 2011 (rats of Indonesia, Vietnam, Cambodia, Laos, and Sri Lanka), with the value of 2.1%. The genetic distance of *R. rattus* lineage I to *R. rattus* lineage III and *R. rattus* lineage IV are considered high, which are 4.1% and 5.3%-6.2%, respectively. From the calculation of pairwise distances, it can be concluded that the genetic distances of Cyt b sequence can explain the relationship between the species and genus level.

**Table 4.2:** Average percentages of pairwise genetic distances among *Rattus* species and *Leopoldamys sabanus* (Outgroup), based on the Kimura 2-Parameter model.

| Taxon | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) *R. tanezumi* (Peninsular Malaysia) | - | | | | | | | | | | | |
| (2) *R. tanezumi* (Asia+Africa+America+Oceania) | 0.9 | - | | | | | | | | | | |
| (3) *R. rattus* lineage I (Asia+Africa+America+Oceania+Europe) | 3.9 | 3.9 | - | | | | | | | | | |
| (4) *R. rattus* lineage III (Asia: Himalaya) | 4.6 | 4.7 | 4.1 | - | | | | | | | | |
| (5) *R. baluensis* | 5.3 | 5.5 | 5.2 | 5.0 | - | | | | | | | |
| (6) *R. tiomanicus* | 5.7 | 5.8 | 5.0 | 5.0 | 1.9 | - | | | | | | |
| (7) *R. rattus* lineage IV sensu Aplin et al., 2011 | 5.7 | 5.8 | 5.3 | 5.4 | 5.1 | 5.0 | - | | | | | |
| (8) *R. rattus cf.* lineage IV sensu Aplin et al., 2011 | 6.8 | 6.9 | 6.2 | 6.0 | 4.8 | 4.7 | 2.1 | - | | | | |
| (9) *R. argentiventer* | 9.7 | 10.0 | 9.7 | 10.0 | 9.6 | 9.9 | 9.9 | 10.8 | - | | | |
| (10) *R. exulans* | 10.6 | 11.1 | 10.7 | 11.4 | 10.0 | 10.6 | 10.7 | 10.9 | 11.4 | - | | |
| (11) *R. norvegicus* | 13.4 | 13.5 | 12.5 | 12.6 | 11.5 | 11.7 | 12.1 | 11.9 | 13.3 | 13.6 | - | |
| (12) *L. sabanus* | 16.8 | 17.1 | 16.2 | 15.7 | 15.4 | 15.2 | 15.4 | 14.7 | 17.2 | 17.0 | 16.1 | - |

For phylogenetic analysis, the Kimura 2-Parameter model was used in NJ tree reconstructions, tested with a bootstrap value of 1000. Eight well-defined clades of *Rattus* species were successfully resolved from the NJ tree (Figure 4.4), referred to the clades of *R. tanezumi*, *R. rattus* lineage I, *R. rattus* lineage III, *R. tiomanicus*-*R. baluensis*, *R. rattus* lineage IV, *R. argentiventer*, *R. exulans*, and *R. norvegicus*. The emergence of *R. tanezumi* clade with the bootstrap value of 95%, comprised of two subclades which includes *R. tanezumi* from Asia (including Peninsular Malaysia), Africa, America, and Oceania rats in the first subclade and another subclade covered *R. tanezumi* from Nepal. *Rattus rattus* lineage I and *R. rattus* lineage III formed their own monophyletic clade with 99% and 100% of bootstrap values, respectively. Surprisingly, a strong bootstrap value of 100% supported the disclosure of *R. tiomanicus* and *R. baluensis* as sister taxa in a monophyletic clade. This presentation is highly anticipated, where both species were comparable with the lowest genetic distance value between them. The depiction of *R. rattus* lineage IV clade with bootstrap values of 100% supports the appearance of two subclades; *R. rattus cf.* lineage IV sensu Aplin et al., 2011 from Peninsular Malaysia, Philippines, and southern Thailand on the one hand and *R. rattus* lineage IV sensu Aplin et al., 2011 of Indonesia, Laos, Vietnam, Cambodia, and Sri Lanka on the other hand. From the visualisation of the NJ tree, both *R. rattus* from Sri Lanka were separated in two different clades, *R. rattus* lineage I and *R. rattus* lineage IV clade. The remaining monophyletic clades consisted solely of one species each, namely, *R. argentiventer* clade, *R. exulans* clade, and *R. norvegicus* clade.

Similar to NJ, BI tree (Figure 4.5) also revealed the separation of eight branches among the ingroups. Clade A represents *R. norvegicus*, clade B exhibits *R. exulans*, clade C comprises of *R. argentiventer*, and clade D displays *R. rattus* lineage IV. Focusing on the clade E, a single monophyletic clade was observed with two subclades portrays *R.*

*tiomanicus* and *R. baluensis*, each supported with 1.00 posterior probability values. The remaining monophyletic clades of F, G, and H, with a posterior probability of 1.00, grouped all *R. rattus* lineage III, *R. rattus* lineage I, and *R. tanezumi*, respectively. Overall, the Bayesian posterior probability value pattern underpins the clustering of each appointed species as a strong group.

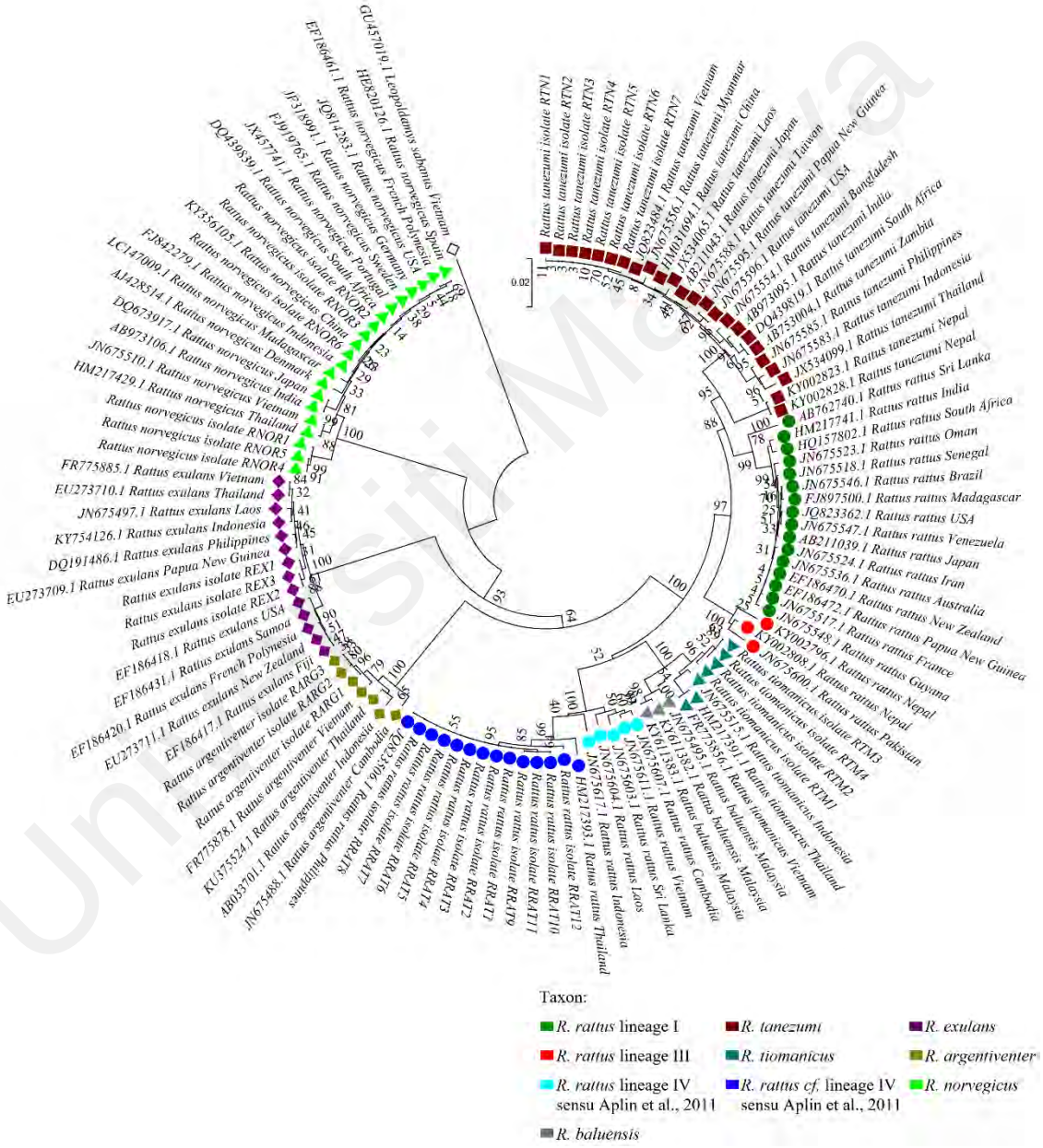


**Figure 4.4:** The neighbor-joining (NJ) phylogenetic tree of *Rattus* species estimated using the Kimura 2-parameter algorithm and 1000 bootstrap replications. The optimal tree with the sum of branch length = 0.6378 is shown and bootstrap values are indicated on the branches. The bar represents 0.02 nucleotide substitutions per alignment position.

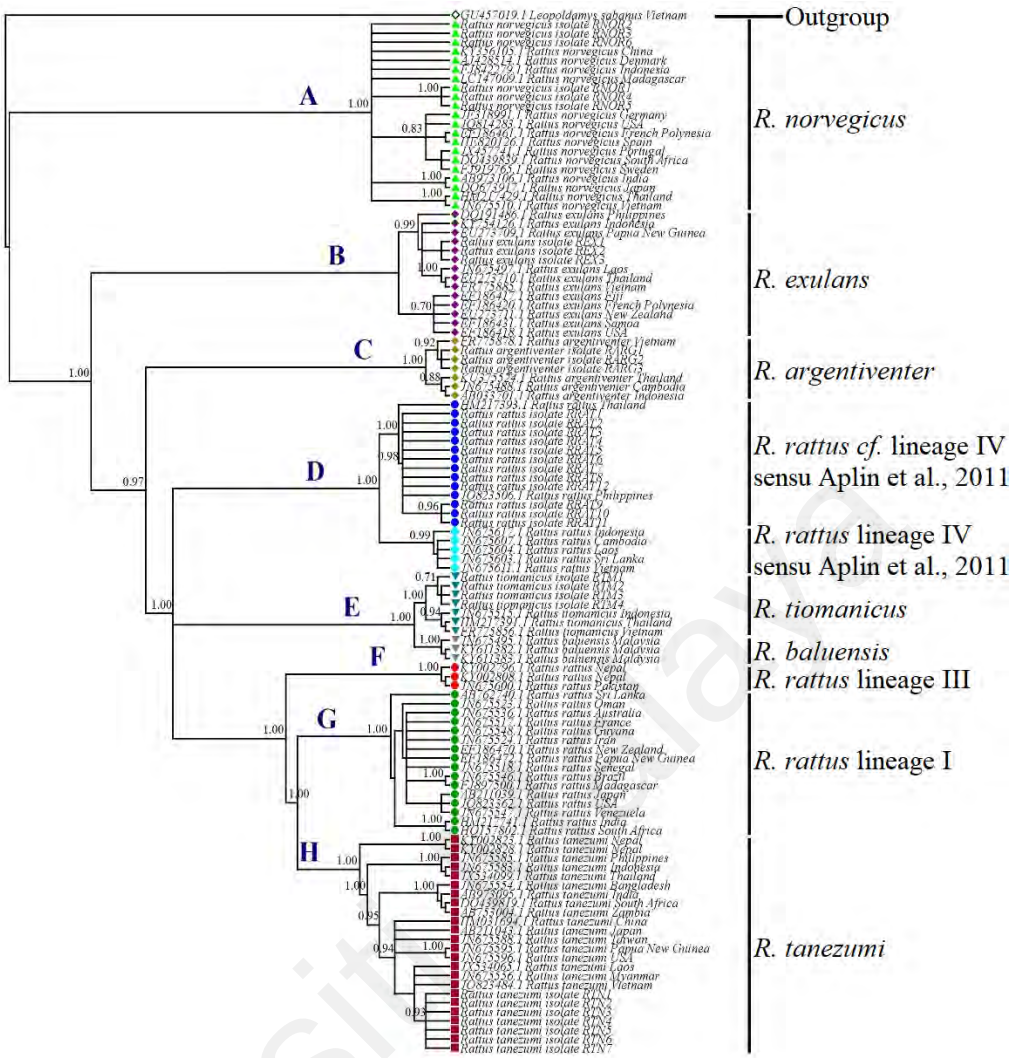


**Figure 4.5:** Bayesian inference (BI) of the 50% majority rule consensus tree of the *Rattus* species, rooted with outgroup. Bayesian posterior probability (*PP*) are accordingly indicated on the branches.

Phylogenetically, the NJ and BI trees shared nearly the similar tree topologies. The clustering patterns of both NJ and BI phylogenetic trees are parallel with results from genetic distance. Essentially, the tree revealed monophyly of the genus *Rattus* with respect to the outgroup, *L. sabanus*. NJ and BI analysis revealed *R. norvegicus* group is the earliest to split among the ingroups and may be an ancestor for the Malaysia's *Rattus*. Surprisingly, *R. rattus* and *R. tanezumi* groups which are very confusing in morphology did show significant genetic separation in phylogenetic analysis with strongly supported and consistent clade formation in NJ and BI phylogenetic trees. Profoundly, both trees

also identified the close relationship of *R. tiomanicus* and *R. baluensis*. Furthermore, the resolution of the individual in the clades represents the theoretical phylogeography pattern. For instance, *R. rattus* populations from Peninsular Malaysia were separated with rats from Southeast Asia countries (e.g. Indonesia, Laos, Vietnam, and Cambodia), which were highly expected.

The results of the molecular divergence time estimation (Figure 4.6) disclosed the *Mus-Rattus* divergence occurred around 11.75 Mya. The initial diversification of *Rattus* estimated at 2.83 Mya, with the earliest divergence of *R. norvegicus* at the basal. The interspecific divergences of *Rattus* group emerged until 0.41 Mya, for *R. tiomanicus-R. baluensis*. Diversification within each of the species of genus *Rattus* estimated to have commenced around 0.21-0.33 Mya. The result of BEAST analysis portrayed comparatively great patterns of species clustering during the late Pleistocene for *Rattus* species. In short, the separation of *R. rattus* lineage IV, *R. tiomanicus*, and *R. baluensis* from *R. rattus* lineage I, *R. rattus* lineage III, and *R. tanezumi* began at 0.91 Mya. The divergence of *R. rattus* lineage III from *R. tanezumi* and *R. rattus* lineage I evidently started around 0.62 Mya. Additionally, *R. tanezumi* and *R. rattus* lineage I last shared a common ancestor approximately at 0.48 Mya. Both *R. tiomanicus* and *R. baluensis* clearly splitted from *R. rattus* lineage IV around 0.70 Mya. The segregation of *R. rattus cf.* lineage IV sensu Aplin et al., 2011 with *R. rattus* lineage IV sensu Aplin et al., 2011 obviously commenced around 0.39 Mya. The emergence of *R. exulans* and *R. argentiventer* clades apparently initiated around 0.65 Mya. Effectively, the *Rattus* group were diverged with the outgroups of *Leopoldamys* nodes with an estimation divergence time of 3.97 Mya.

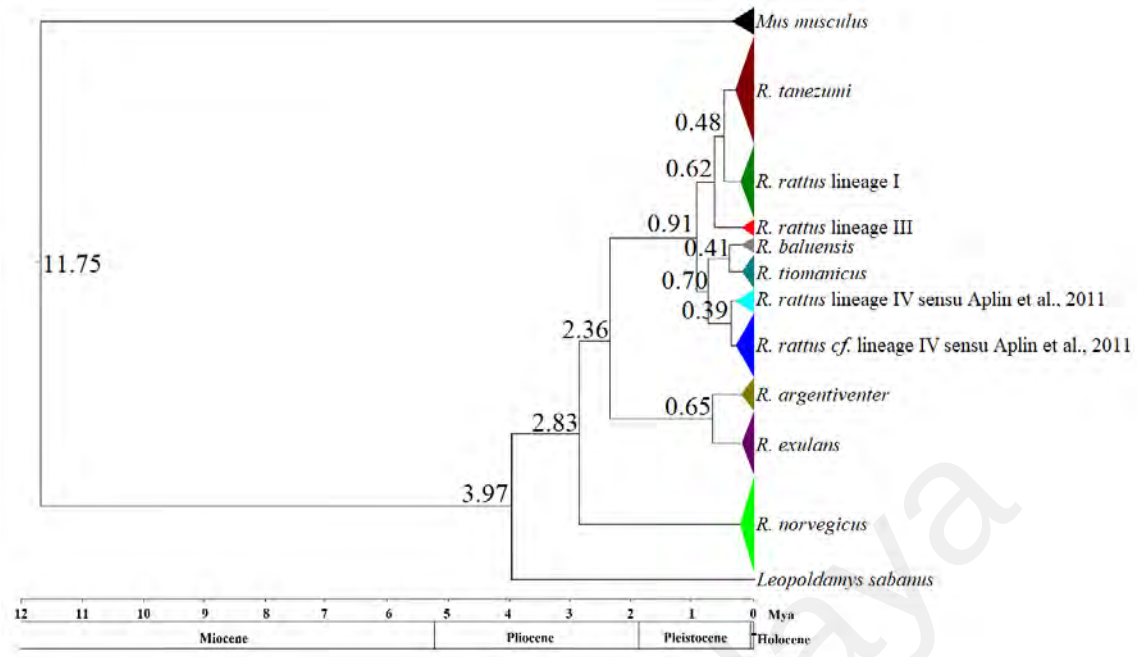


**Figure 4.6:** Result of BEAST analysis for the *Rattus* species based on Cyt b gene. The divergence times is indicated by the node numbers.

## 4.2 Population Structure of *Rattus rattus* in Peninsular Malaysia Based on Mitochondrial and Nuclear DNA Data

In elucidating the population structure of *R. rattus*, PCR was performed using four molecular markers. Mitochondrial markers involved cytochrome oxidase I (COI), cytochrome b (Cyt b), and displacement loop (D-loop), while nuclear marker represents by interphotoreceptor retinoid binding protein (IRBP).

### 4.2.1 Sequence analysis

Mitochondrial DNA (mtDNA) of COI, Cyt b, and D-loop sequences were concatenated for analysis. A single nuclear DNA (nDNA) of IRBP sequences were analysed separately. Both mtDNA and nDNA sequences were analysed using MEGA7. A total of 2415 bp of mtDNA gene and 1227 bp of nDNA gene sequences were obtained after the final editing stage of all genetic samples used in this study. No insertions, deletions, or stop codons (for the protein coding gene) were observed.

Analysis of mtDNA gene of all 130 sequences of *R. rattus* shows a total of 2370 (98.14%) conserved sites and 45 (1.86%) variable sites. Parsimony-informative sites and singleton sites showed 45 bp and 0 bp, respectively. The transition-transversion rates analysis were also conducted which shows that the mtDNA sequences undergoes higher transition rates as compared to transversion rates. The mtDNA nucleotide pair analysis indicates that the number of transition and transversion are 15 and two, respectively resulting in Ti/Tv ratio of 6.0. It is widely known that transition and transversion are the type of substitution mutations, in which transition is interchanges of two-ring purines (A-G) or one-ring pyrimidines (C-T). A transversion is interchanges of a purine for pyrimidine bases, which consequently include an exchange of one-ring and two-ring structures or vice versa. The nucleotide composition analysis of all mtDNA nucleotide sequences shows that Adenine (A) has the highest composition in the sequences with the average of 30.4%, followed by Thymine (T) with the average of 29.3%, Cytosine (C) with the average of 26.7%, and Guanine (G) with the average of 13.7%.

For nDNA gene, a total of 1195 (97.39%) conserved sites and 32 (2.61%) variable sites were observed. Parsimony-informative sites and singleton sites showed 32 bp and 0 bp, respectively. The nDNA sequences also undergoes higher transition rates as compared to transversion rates. The nDNA nucleotide pair analysis signified that the number of transition and transversion are nine and two, respectively resulting in Ti/Tv ratio of 4.5. The nucleotide composition analysis of all nDNA nucleotide sequences shows that Guanine (G) has the highest composition in the sequences with the average of 30.3%, followed by Cytosine (C) with the average of 29.7%, Adenine (A) with the average of 20.3%, and Thymine (T) with the average of 19.7%. Table 4.3 shows nucleotide information of *R. rattus* based on mtDNA and nDNA genes.

**Table 4.3:** Nucleotide information of *R. rattus* based on mtDNA and nDNA genes.

| Character | mtDNA | nDNA |
|---|---|---|
| Total character | 2415 | 1227 |
| Conserved sites | 2370 | 1195 |
| Variable sites | 45 | 32 |
| Parsimony informative sites | 45 | 32 |
| Parsimony uninformative sites | 0 | 0 |
| Percentage of informative character (%) | 1.86 | 2.61 |
| Number of transition (Ti) | 15 | 9 |
| Number of transversion (Tv) | 2 | 2 |
| Ratio of Ti/Tv | 6 | 4.5 |

The pairwise distance of both mtDNA and nDNA sequences were calculated using Kimura 2-Parameter algorithm model. The pairwise distances between populations of *R. rattus* in Peninsular Malaysia based on mtDNA sequences is shown in Table 4.4. From the analysis, it shows that the genetic distance between populations of *R. rattus* in Penang island and Terengganu were the most distant from other populations which is 1.1%. The closest relationship is between populations of Penang mainland and Perak which is 0.2%. Altogether, the genetic distances within populations of *R. rattus* in Peninsular Malaysia based on mtDNA sequences are considered low as the pairwise distance value is still less than 1.1%. Populations of *R. rattus* and outgroup species of *L. sabanus* exhibited high value of pairwise distances, ranging from 26.6%-27.1% which were highly expected as they were separated in two different genera.

**Table 4.4:** Pairwise distances between populations of *R. rattus* in Peninsular Malaysia based on mtDNA sequences, via Kimura 2-Parameter model.

| **Populations** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) *R. rattus* Perlis | - | | | | | | | | | | | | | |
| (2) *R. rattus* Kedah | 0.5 | - | | | | | | | | | | | | |
| (3) *R. rattus* Penang mainland | 0.7 | 0.5 | - | | | | | | | | | | | |
| (4) *R. rattus* Penang island | 0.7 | 0.6 | 0.5 | - | | | | | | | | | | |
| (5) *R. rattus* Perak | 0.7 | 0.4 | 0.2 | 0.5 | - | | | | | | | | | |
| (6) *R. rattus* Selangor | 0.9 | 0.8 | 0.8 | 0.7 | 0.8 | - | | | | | | | | |
| (7) *R. rattus* Kuala Lumpur | 0.9 | 0.8 | 0.7 | 0.7 | 0.7 | 0.7 | - | | | | | | | |
| (8) *R. rattus* Negeri Sembilan | 0.8 | 0.7 | 0.8 | 0.9 | 0.6 | 0.7 | 0.8 | - | | | | | | |
| (9) *R. rattus* Melaka | 0.9 | 0.9 | 0.7 | 0.6 | 0.8 | 0.7 | 0.6 | 1.0 | - | | | | | |
| (10) *R. rattus* Johor | 1.0 | 0.7 | 0.3 | 0.7 | 0.4 | 0.7 | 0.7 | 0.8 | 0.7 | - | | | | |
| (11) *R. rattus* Kelantan | 0.6 | 0.7 | 0.9 | 0.9 | 0.8 | 0.6 | 0.7 | 0.8 | 0.9 | 0.8 | - | | | |
| (12) *R. rattus* Terengganu | 0.7 | 0.8 | 1.0 | 1.1 | 1.0 | 0.7 | 0.8 | 1.0 | 0.8 | 0.8 | 0.5 | - | | |
| (13) *R. rattus* Pahang | 0.9 | 0.9 | 0.9 | 1.0 | 1.0 | 0.7 | 0.7 | 1.0 | 0.8 | 0.7 | 0.5 | 0.4 | - | |
| (14) *Leopoldamys sabanus* | 26.8 | 26.8 | 26.7 | 26.6 | 26.7 | 27.0 | 26.8 | 26.9 | 26.8 | 26.8 | 27.1 | 27.0 | 26.9 | - |

In furtherance, the pairwise distances between populations of *R. rattus* in Peninsular Malaysia based on nDNA sequences is shown in Table 4.5. From the analysis, it shows that the populations of Penang island and Terengganu have the highest value of pairwise distance as compared to other populations which is 1.2%. These two populations have a remote connection and the pattern of distance segregation is similar to the result obtained in mtDNA sequences. The lowest genetic distance within populations are shown by Kelantan and Terengganu *R. rattus* with the value of 0.5%, indicating that the individuals within this population have the closest relationships. This result is contradicted with the results of mtDNA sequences, which shows Penang mainland and Perak *R. rattus* have a minute segregation. However, the genetic distances within populations of *R. rattus* in Peninsular Malaysia based on nDNA sequences are also considered low as the pairwise distance value is still less than 1.2%. From the analysis including the outgroup, the distance relationship between *R. rattus* with *L. sabanus* was the most distant which is in the range of 5.0%-5.2%. Moreover, *L. sabanus* come from different genus of *Rattus* which is *Leopoldamys*. Therefore, *L. sabanus* is very suitable to be the outgroup in this study. From the pairwise distances, it can be generally concluded that the genetic distances of mtDNA and nDNA sequences can explain the relationship between the population and generic level.

Single-nucleotide polymorphisms (SNP) analysis conducted on the mtDNA and nDNA sequences excluding the outgroup revealed 45 and 32 segregating sites, respectively. The segregating sites for mtDNA and nDNA sequences are shown in Appendix D and Appendix E, respectively.

**Table 4.5:** Pairwise distances between populations of *R. rattus* in Peninsular Malaysia based on nDNA sequences, via Kimura 2-Parameter model.

| **Populations** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) *R. rattus* Perlis | - | | | | | | | | | | | | | |
| (2) *R. rattus* Kedah | 0.8 | - | | | | | | | | | | | | |
| (3) *R. rattus* Penang mainland | 0.9 | 0.8 | - | | | | | | | | | | | |
| (4) *R. rattus* Penang island | 0.9 | 1.0 | 1.0 | - | | | | | | | | | | |
| (5) *R. rattus* Perak | 0.8 | 0.7 | 0.8 | 0.9 | - | | | | | | | | | |
| (6) *R. rattus* Selangor | 0.9 | 0.9 | 1.0 | 1.1 | 1.0 | - | | | | | | | | |
| (7) *R. rattus* Kuala Lumpur | 1.0 | 1.0 | 1.1 | 1.0 | 1.0 | 1.0 | - | | | | | | | |
| (8) *R. rattus* Negeri Sembilan | 1.0 | 0.8 | 0.9 | 1.1 | 0.9 | 1.0 | 1.0 | - | | | | | | |
| (9) *R. rattus* Melaka | 0.9 | 0.9 | 1.0 | 1.1 | 1.0 | 0.9 | 1.0 | 1.0 | - | | | | | |
| (10) *R. rattus* Johor | 1.0 | 1.0 | 1.1 | 1.1 | 1.0 | 0.9 | 1.0 | 1.0 | 0.9 | - | | | | |
| (11) *R. rattus* Kelantan | 0.9 | 0.8 | 0.9 | 1.0 | 0.9 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | - | | | |
| (12) *R. rattus* Terengganu | 0.9 | 0.8 | 1.0 | 1.2 | 0.9 | 0.8 | 0.9 | 0.9 | 0.8 | 0.8 | 0.5 | - | | |
| (13) *R. rattus* Pahang | 0.9 | 0.9 | 1.0 | 1.0 | 0.9 | 0.9 | 0.9 | 0.9 | 0.8 | 0.8 | 0.6 | 0.6 | - | |
| (14) *Leopoldamys sabanus* | 5.0 | 5.1 | 5.0 | 5.1 | 5.1 | 5.1 | 5.1 | 5.2 | 5.1 | 5.0 | 5.0 | 5.1 | 5.1 | - |

For mtDNA sequences, 91 haplotypes were defined based on 130 sequences of *R. rattus* from 13 different populations of Peninsular Malaysia. There is a single shared haplotype (H_18) and 90 unique haplotypes. A single shared haplotype with the highest samples frequencies is representing populations of *R. rattus* from Kedah, Penang mainland, and Perak. These three cities are situated in the northern region of Peninsular Malaysia. Besides, each population of *R. rattus* has more than one unique haplotype. The population with the highest number of haplotypes is Kuala Lumpur, with 12 unique haplotypes. Minimum-spanning network (MSN) was generated with the haplotype data of mtDNA sequences obtained to illustrate the relationships of *R. rattus* populations in Peninsular Malaysia (Figure 4.7). The haplotype network illustrated the homogenise populations of *R. rattus* in Peninsular Malaysia, with no population disjunction observed.

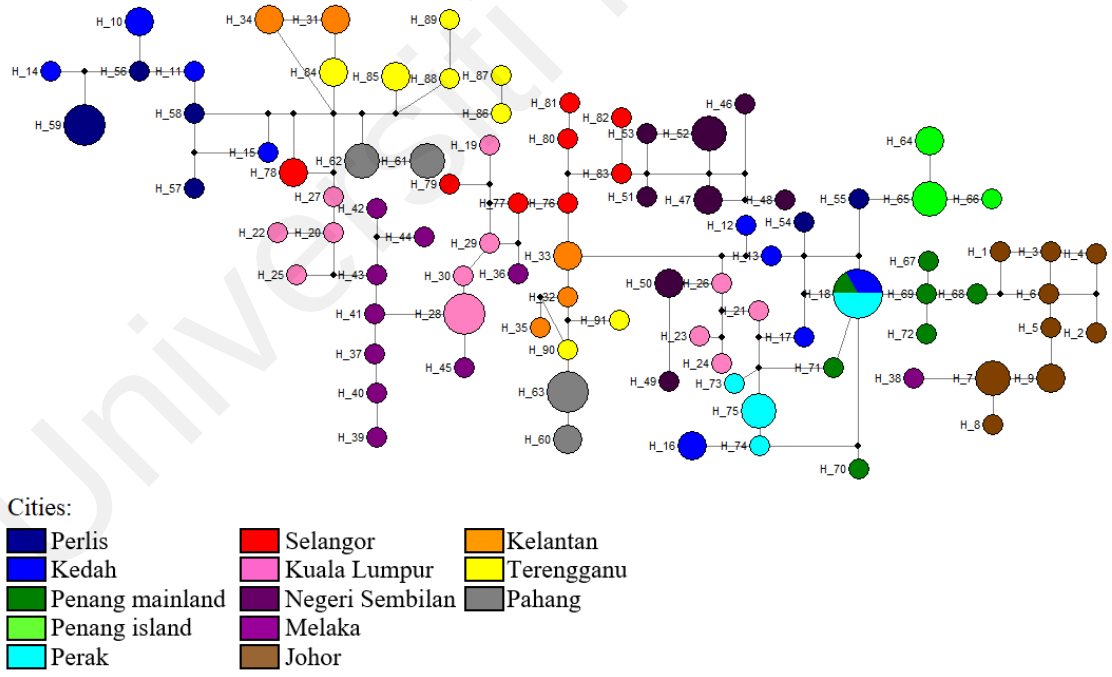


**Figure 4.7:** The minimum-spanning network (MSN) generated illustrating the relationships of *R. rattus* from different locality, based on mtDNA sequences. Each circle represents a haplotype, and the diameter is scaled to the haplotype frequency.

For nDNA sequences, 95 haplotypes were defined based on 130 sequences of *R. rattus* from 13 different populations of Peninsular Malaysia. There are six shared haplotypes (H_1, H_8, H_10, H_28, H_49, H_68) and 89 unique haplotypes. The shared haplotypes are symbolising populations of *R. rattus* from northern region of Perlis-Kedah, Perlis-Penang mainland, Penang mainland-Perak, southern region of Melaka-Johor, and east coast region of Kelantan-Terengganu. In addition, each population of *R. rattus* has more than one unique haplotype. The population with the highest number of haplotypes is Pahang, with 11 unique haplotypes. Minimum-spanning network (MSN) was generated with the haplotype data of nDNA sequences obtained to illustrate the relationships of *R. rattus* populations in Peninsular Malaysia (Figure 4.8). The haplotype network demonstrated the mixed populations of *R. rattus* in Peninsular Malaysia. There was no clear separation of populations based on theoretical phylogeography pattern.

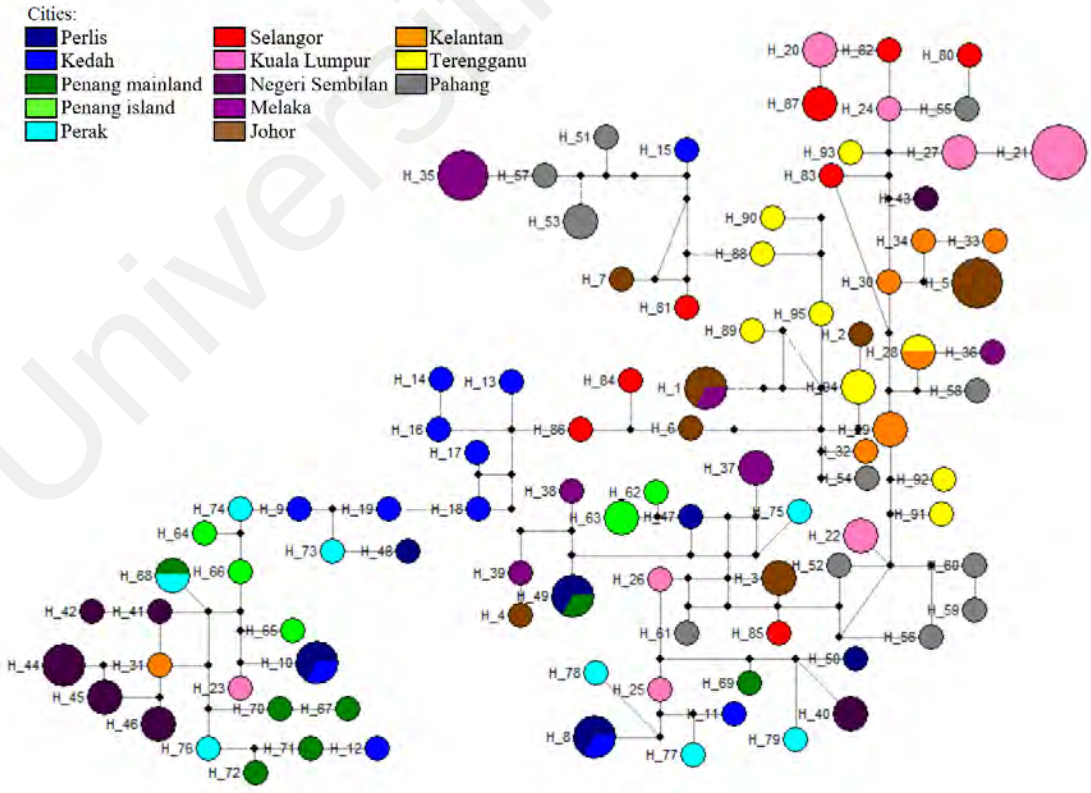


**Figure 4.8:** The minimum-spanning network (MSN) generated illustrating the relationships of *R. rattus* from different locality, based on nDNA sequences. Each circle represents a haplotype, and the diameter is scaled to the haplotype frequency.

The analysis for nucleotide diversity ($\pi$) and net nucleotide divergence (*Da*) were conducted based on the origin of samples according to their locality. The analysis based on mtDNA sequences (Table 4.6) portrayed that π value was the highest between Penang island-Pahang populations with the value of 0.00692. The results are consistent with *Da*, which the highest *Da* value is also between Penang island-Pahang populations with the value of 0.00825. The lowest values, 0.00179 for π and 0.00061 for *Da*, were found consistent between Penang mainland-Perak populations. The measures of nucleotide diversity ($\pi$) and net nucleotide divergence (*Da*) among populations of *R. rattus* based on nDNA sequences is shown in Table 4.7. For nDNA sequences, the highest π value of 0.00966 was observed in populations of Penang Island-Selangor. Also, the similar populations portrayed the highest value of *Da* of 0.00310. Populations of Kelantan-Terengganu described the lowest values of $\pi = 0.00603$ and $Da = 0.00001$.

**Table 4.6:** Measures of nucleotide diversity ($\pi$) and net nucleotide divergence (*Da*) among populations of *R. rattus*, based on mtDNA sequences.

| **Populations** | Perlis ($N = 9$) | Kedah ($N = 12$) | Penang mainland ($N = 7$) | Penang island ($N = 6$) | Perak ($N = 8$) | Selangor ($N = 9$) | Kuala Lumpur ($N = 15$) | Negeri Sembilan ($N = 12$) | Melaka ($N = 10$) | Johor ($N = 12$) | Kelantan ($N = 8$) | Tereng-ganu ($N = 10$) | Pahang ($N = 12$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) Perlis | - | 0.00433 | 0.00500 | 0.00448 | 0.00443 | 0.00628 | 0.00686 | 0.00577 | 0.00669 | 0.00588 | 0.00466 | 0.00502 | 0.00568 |
| (2) Kedah | 0.00076 | - | 0.00453 | 0.00502 | 0.00399 | 0.00652 | 0.00661 | 0.00584 | 0.00683 | 0.00530 | 0.00542 | 0.00625 | 0.00635 |
| (3) Penang mainland | 0.00492 | 0.00172 | - | 0.00323 | 0.00179 | 0.00603 | 0.00572 | 0.00521 | 0.00549 | 0.00239 | 0.00580 | 0.00642 | 0.00573 |
| (4) Penang island | 0.00459 | 0.00320 | 0.00327 | - | 0.00302 | 0.00513 | 0.00572 | 0.00564 | 0.00506 | 0.00393 | 0.00576 | 0.00659 | 0.00692 |
| (5) Perak | 0.00460 | 0.00141 | 0.00061 | 0.00377 | - | 0.00554 | 0.00573 | 0.00447 | 0.00578 | 0.00266 | 0.00507 | 0.00658 | 0.00596 |
| (6) Selangor | 0.00484 | 0.00357 | 0.00503 | 0.00385 | 0.00497 | - | 0.00586 | 0.00556 | 0.00587 | 0.00490 | 0.00525 | 0.00571 | 0.00524 |
| (7) Kuala Lumpur | 0.00505 | 0.00301 | 0.00330 | 0.00381 | 0.00402 | 0.00168 | - | 0.00649 | 0.00536 | 0.00536 | 0.00596 | 0.00614 | 0.00562 |
| (8) Negeri Sembilan | 0.00504 | 0.00335 | 0.00482 | 0.00642 | 0.00397 | 0.00319 | 0.00394 | - | 0.00602 | 0.00550 | 0.00550 | 0.00656 | 0.00640 |
| (9) Melaka | 0.00549 | 0.00408 | 0.00387 | 0.00354 | 0.00522 | 0.00239 | 0.00070 | 0.00566 | - | 0.00480 | 0.00643 | 0.00617 | 0.00565 |
| (10) Johor | 0.00736 | 0.00412 | 0.00150 | 0.00545 | 0.00262 | 0.00402 | 0.00336 | 0.00580 | 0.00357 | - | 0.00504 | 0.00531 | 0.00451 |
| (11) Kelantan | 0.00304 | 0.00246 | 0.00596 | 0.00653 | 0.00546 | 0.00242 | 0.00285 | 0.00422 | 0.00456 | 0.00556 | - | 0.00407 | 0.00413 |
| (12) Terengganu | 0.00360 | 0.00409 | 0.00715 | 0.00821 | 0.00813 | 0.00337 | 0.00320 | 0.00610 | 0.00410 | 0.00567 | 0.00141 | - | 0.00353 |
| (13) Pahang | 0.00567 | 0.00502 | 0.00681 | 0.00825 | 0.00776 | 0.00338 | 0.00286 | 0.00640 | 0.00396 | 0.00459 | 0.00235 | 0.00100 | - |

*Above the diagonal: percentage of nucleotide diversity between geographic samples ($\pi$).
Below the diagonal: nucleotide divergence between pairs of samples (*Da*).

**Table 4.7:** Measures of nucleotide diversity ($\pi$) and net nucleotide divergence (*Da*) among populations of *R. rattus*, based on nDNA sequences.

| **Populations** | Perlis ($N = 9$) | Kedah ($N = 12$) | Penang mainland ($N = 7$) | Penang island ($N = 6$) | Perak ($N = 8$) | Selangor ($N = 9$) | Kuala Lumpur ($N = 15$) | Negeri Sembilan ($N = 12$) | Melaka ($N = 10$) | Johor ($N = 12$) | Kelantan ($N = 8$) | Tereng-ganu ($N = 10$) | Pahang ($N = 12$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) Perlis | - | 0.00722 | 0.00809 | 0.00827 | 0.00777 | 0.00846 | 0.00867 | 0.00831 | 0.00825 | 0.00867 | 0.00766 | 0.00775 | 0.00796 |
| (2) Kedah | 0.00053 | - | 0.00748 | 0.00826 | 0.00737 | 0.00836 | 0.00887 | 0.00775 | 0.00829 | 0.00857 | 0.00738 | 0.00754 | 0.00782 |
| (3) Penang mainland | 0.00160 | 0.00062 | - | 0.00880 | 0.00793 | 0.00924 | 0.00890 | 0.00813 | 0.00907 | 0.00925 | 0.00789 | 0.00807 | 0.00814 |
| (4) Penang island | 0.00149 | 0.00184 | 0.00161 | - | 0.00857 | 0.00966 | 0.00865 | 0.00904 | 0.00937 | 0.00942 | 0.00823 | 0.00825 | 0.00849 |
| (5) Perak | 0.00040 | 0.00013 | 0.00005 | 0.00060 | - | 0.00899 | 0.00896 | 0.00822 | 0.00885 | 0.00902 | 0.00789 | 0.00781 | 0.00795 |
| (6) Selangor | 0.00140 | 0.00150 | 0.00220 | 0.00310 | 0.00122 | - | 0.00887 | 0.00865 | 0.00870 | 0.00869 | 0.00744 | 0.00774 | 0.00789 |
| (7) Kuala Lumpur | 0.00264 | 0.00303 | 0.00283 | 0.00193 | 0.00238 | 0.00190 | - | 0.00871 | 0.00860 | 0.00878 | 0.00767 | 0.00791 | 0.00794 |
| (8) Negeri Sembilan | 0.00233 | 0.00130 | 0.00156 | 0.00243 | 0.00120 | 0.00175 | 0.00247 | - | 0.00894 | 0.00878 | 0.00752 | 0.00772 | 0.00803 |
| (9) Melaka | 0.00135 | 0.00164 | 0.00235 | 0.00244 | 0.00141 | 0.00087 | 0.00161 | 0.00258 | - | 0.00859 | 0.00770 | 0.00752 | 0.00761 |
| (10) Johor | 0.00203 | 0.00201 | 0.00269 | 0.00263 | 0.00172 | 0.00083 | 0.00183 | 0.00212 | 0.00098 | - | 0.00748 | 0.00754 | 0.00759 |
| (11) Kelantan | 0.00247 | 0.00191 | 0.00248 | 0.00273 | 0.00181 | 0.00057 | 0.00156 | 0.00183 | 0.00139 | 0.00069 | - | 0.00603 | 0.00637 |
| (12) Terengganu | 0.00240 | 0.00070 | 0.00267 | 0.00271 | 0.00153 | 0.00104 | 0.00184 | 0.00199 | 0.00089 | 0.00070 | 0.00001 | - | 0.00629 |
| (13) Pahang | 0.00239 | 0.00208 | 0.00245 | 0.00287 | 0.00147 | 0.00102 | 0.00156 | 0.00221 | 0.00073 | 0.00049 | 0.00036 | 0.00031 | - |

*Above the diagonal: percentage of nucleotide diversity between geographic samples ($\pi$).
Below the diagonal: nucleotide divergence between pairs of samples (*Da*).

### 4.2.2 Phylogenetic analysis

The phylogenetic trees of both mtDNA and nDNA sequences were built using distance-based (neighbor-joining (NJ)) tree. The NJ phylogenetic trees were generated using Kimura 2-Parameter with 1000 bootstrap replication. In the phylogenetic analysis, a single sequence of *Leopoldamys sabanus* was used as an outgroup to root the phylogenetic tree. Based on the concatenated mtDNA sequences of COI, Cyt b, and D-loop, the tree topology showed separation of ingroups (*R. rattus*) and outgroup (*L. sabanus*) with 78% bootstrap value (Figure 4.9). For NJ tree reconstructions using nDNA sequences of IRBP, the monophyly of *R. rattus* was discovered with respect to outgroup, *L. sabanus* (Figure 4.10). Phylogenetically, both trees demonstrated a mixed grouping of *R. rattus* from 13 populations of Peninsular Malaysia, which can be seen through the tree topologies and supported with inconsistent bootstrap value for both trees. The resolution of the populations of *R. rattus* in the clades are homogeneous and not geographically partitioned. The topologies did not show distinct groupings corresponding to sampling localities. There are several unresolved relationships illustrated in the tree topologies with bootstrap value less than 20%.

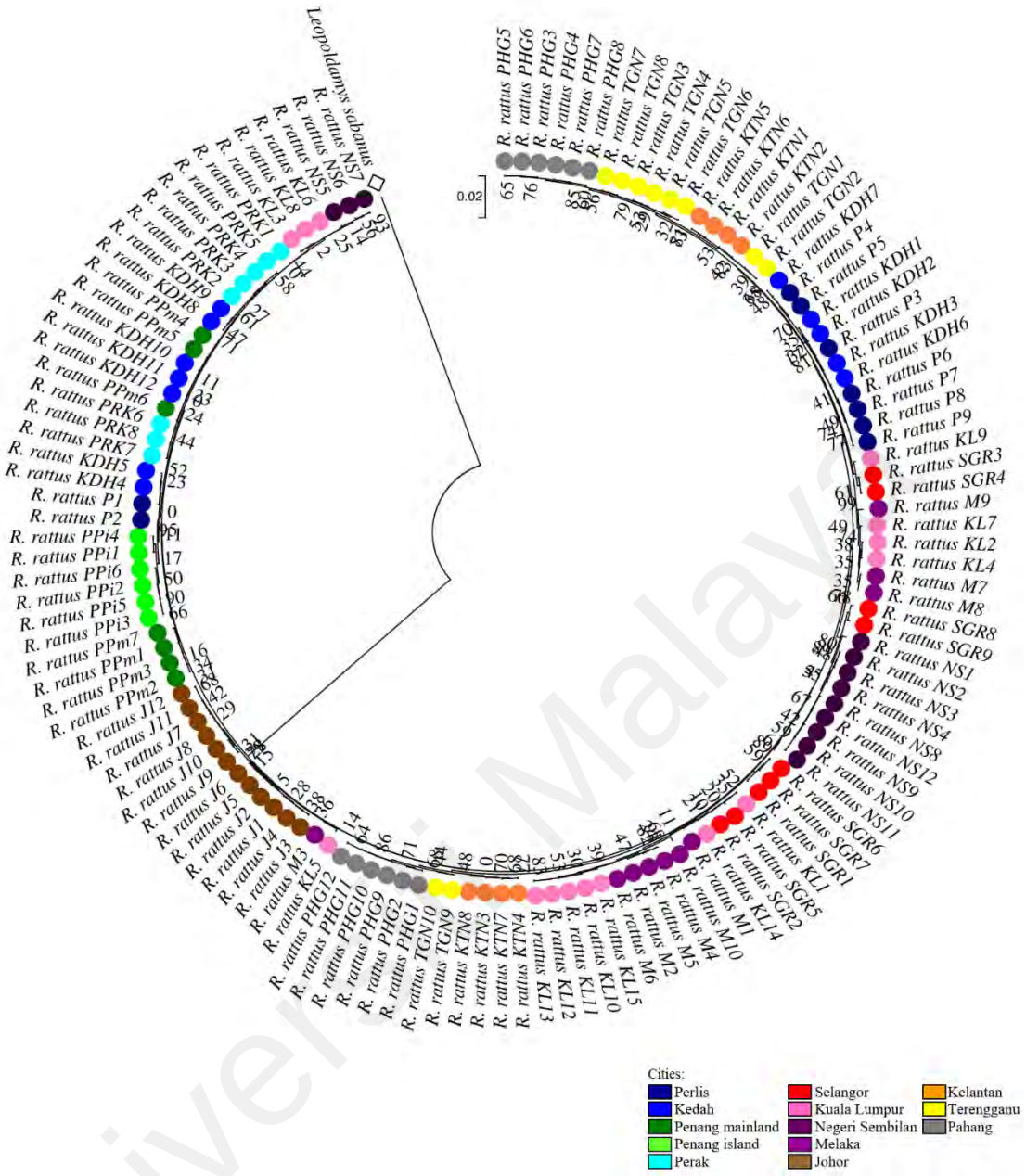


**Figure 4.9:** The neighbor-joining (NJ) phylogenetic tree of *R. rattus* species based on mtDNA sequences. The optimal tree with the sum of branch length = 0.3557 is shown and bootstrap values are indicated on the branches. The bar represents 0.02 nucleotide substitutions per alignment position.

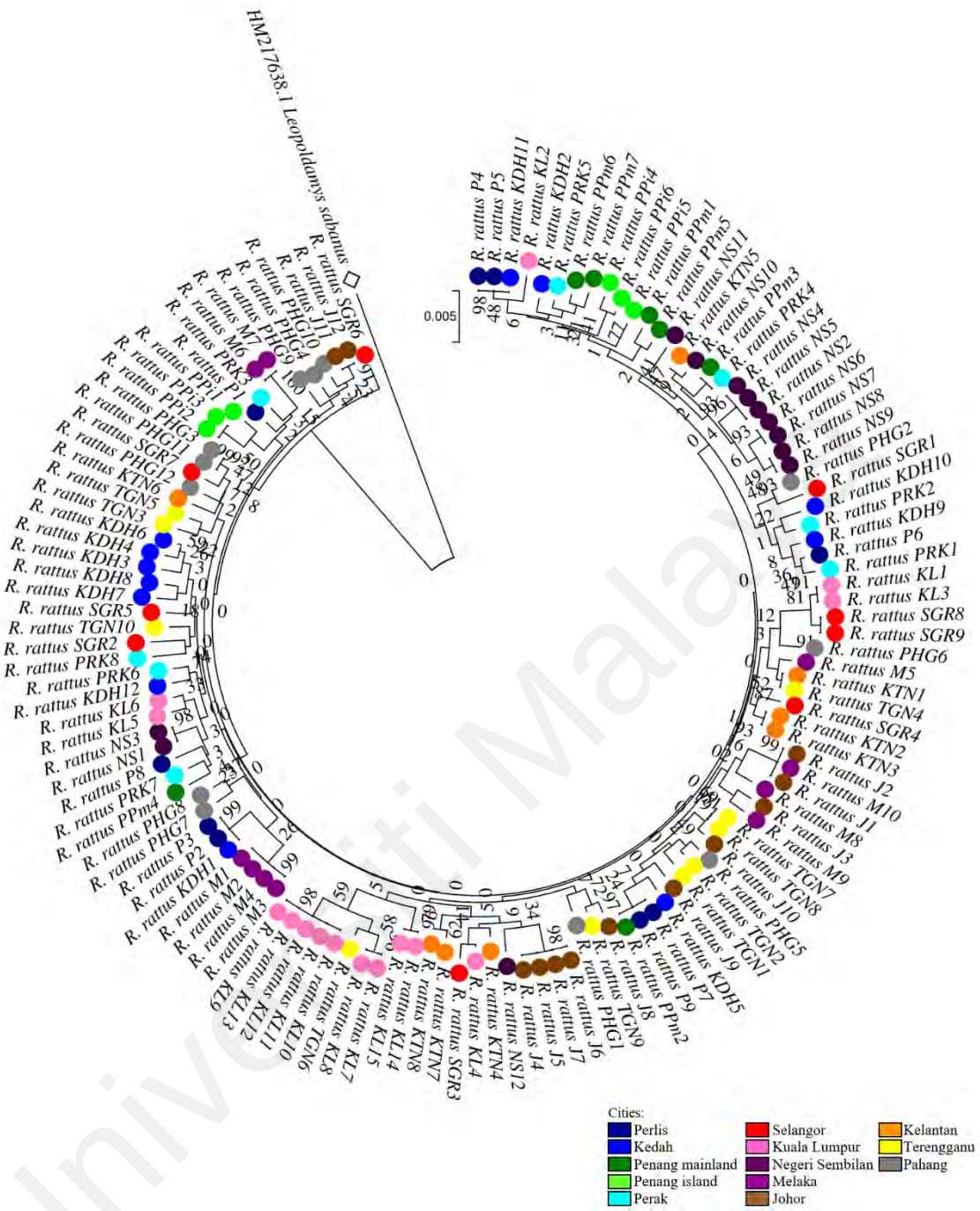


**Figure 4.10:** The neighbor-joining (NJ) phylogenetic tree of *R. rattus* species based on nDNA sequences. The optimal tree with the sum of branch length = 0.2640 is shown and bootstrap values are indicated on the branches. The bar represents 0.005 nucleotide substitutions per alignment position.

### 4.2.3 Population genetics analysis

Evaluations of population genetic parameters such as nucleotide subdivision (*Nst*), population subdivision (*Fst*), the number of migrants per generation (*Nm*), haplotype diversity (*Hd*), average number of nucleotide differences (*k*), and segregating sites (*S*) were also evaluated from the mtDNA and nDNA dataset using DNASP v5. Genetic differentiation (*Nst*, *Fst*, and *Nm*) values were calculated to further elucidate the relationships among *R. rattus* populations in 13 different localities in Peninsular Malaysia. *Nst* analysis can be utilised to evaluate a population's subdivision at the nucleotide level (Bouga et al., 2005), with 0 = no population subdivision and 1 = complete population division. *Fst* is the probability that two random gametes drawn from two populations are identical by descent and relative to gametes taken from the whole populations. Only *Fst* values > 0.25 greatly stipulate a genetic differentiation of populations (Lowe et al., 2004). Theoretically, populations are expected to genetically diverge over time when the *Nm* value is < 1, and they are expected to retain gene flow when *Nm* is > 1. Hence, the *Nm* value will be inversely proportional to *Nst* and *Fst*.

Table 4.8 shows the evaluations of nucleotide subdivision (*Nst*) and population subdivision (*Fst*) among populations of *R. rattus*, based on mtDNA sequences. The measures of gene flow (*Nm*) and geographical distance, kilometre (KM) among populations of *R. rattus*, based on mtDNA sequences is shown in Table 4.9. Both *Nst* and *Fst* analysis outcomes demonstrated populations of Kuala Lumpur-Melaka exhibited the lowest *Nst* and *Fst* of 0.10965 and 0.12725, respectively. The division of Penang island-Pahang populations also revealed the highest *Nst* and *Fst*, at 0.81661 and 0.81576, respectively. Here, *Nm* analysis validated both *Nst* and *Fst* results, as populations of Kuala Lumpur-Melaka (with geographical distance of 123 KM) had the highest *Nm*, with the value of 4.07. *Nm* explains the average number of individuals per generation migrating

**Table 4.8:** Measures of nucleotide subdivision (*Nst*) and population subdivision (*Fst*) among populations of *R. rattus*, based on mtDNA sequences.

| **Populations** | Perlis ($N = 9$) | Kedah ($N = 12$) | Penang mainland ($N = 7$) | Penang island ($N = 6$) | Perak ($N = 8$) | Selangor ($N = 9$) | Kuala Lumpur ($N = 15$) | Negeri Sembilan ($N = 12$) | Melaka ($N = 10$) | Johor ($N = 12$) | Kelantan ($N = 8$) | Tereng-ganu ($N = 10$) | Pahang ($N = 12$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) Perlis | - | 0.16647 | 0.67720 | 0.70383 | 0.70365 | 0.56701 | 0.55401 | 0.61739 | 0.59497 | 0.77157 | 0.49843 | 0.53718 | 0.67205 |
| (2) Kedah | 0.16625 | - | 0.34223 | 0.52605 | 0.32843 | 0.43456 | 0.37517 | 0.45061 | 0.46533 | 0.56809 | 0.38062 | 0.50240 | 0.57459 |
| (3) Penang mainland | 0.67631 | 0.34151 | - | 0.69492 | 0.32681 | 0.61126 | 0.48041 | 0.64746 | 0.54397 | 0.47072 | 0.70004 | 0.73360 | 0.75020 |
| (4) Penang island | 0.70317 | 0.52566 | 0.69434 | - | 0.78611 | 0.57961 | 0.54641 | 0.74372 | 0.55543 | 0.81098 | 0.75283 | 0.79002 | 0.81661 |
| (5) Perak | 0.70299 | 0.32800 | 0.32644 | 0.78561 | - | 0.63989 | 0.55951 | 0.64100 | 0.64799 | 0.67217 | 0.71746 | 0.78782 | 0.80664 |
| (6) Selangor | 0.56586 | 0.43349 | 0.61016 | 0.57889 | 0.63905 | - | 0.25398 | 0.44465 | 0.34155 | 0.56888 | 0.38175 | 0.45988 | 0.48223 |
| (7) Kuala Lumpur | 0.55286 | 0.37416 | 0.47968 | 0.54600 | 0.55883 | 0.25332 | - | 0.47516 | 0.10965 | 0.49614 | 0.39944 | 0.42514 | 0.41667 |
| (8) Negeri Sembilan | 0.61656 | 0.44991 | 0.64652 | 0.74296 | 0.64042 | 0.44378 | 0.47421 | - | 0.58483 | 0.70163 | 0.55833 | 0.64345 | 0.67738 |
| (9) Melaka | 0.60346 | 0.48009 | 0.56481 | 0.58172 | 0.66331 | 0.36231 | 0.12725 | 0.59140 | - | 0.53630 | 0.53571 | 0.50666 | 0.51920 |
| (10) Johor | 0.77061 | 0.56717 | 0.47026 | 0.81035 | 0.67168 | 0.56798 | 0.49554 | 0.70072 | 0.55839 | - | 0.69893 | 0.69938 | 0.68505 |
| (11) Kelantan | 0.49771 | 0.37994 | 0.69902 | 0.75200 | 0.71671 | 0.38095 | 0.39865 | 0.55755 | 0.55113 | 0.69804 | - | 0.29754 | 0.44085 |
| (12) Terengganu | 0.53635 | 0.50127 | 0.73253 | 0.78912 | 0.78691 | 0.45896 | 0.42428 | 0.64238 | 0.52031 | 0.69846 | 0.29707 | - | 0.24824 |
| (13) Pahang | 0.67107 | 0.57353 | 0.74917 | 0.81576 | 0.80580 | 0.48136 | 0.41603 | 0.67637 | 0.53995 | 0.68424 | 0.44030 | 0.24802 | - |

*Above the diagonal: nucleotide subdivision (*Nst*).
Below the diagonal: population subdivision (*Fst*).

**Table 4.9:** Measures of gene flow (number of migrants, *Nm*) and geographical distance (KM) among populations of *R. rattus*, based on mtDNA sequences.

| **Populations** | Perlis ($N = 9$) | Kedah ($N = 12$) | Penang mainland ($N = 7$) | Penang island ($N = 6$) | Perak ($N = 8$) | Selangor ($N = 9$) | Kuala Lumpur ($N = 15$) | Negeri Sembilan ($N = 12$) | Melaka ($N = 10$) | Johor ($N = 12$) | Kelantan ($N = 8$) | Tereng-ganu ($N = 10$) | Pahang ($N = 12$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) Perlis | - | 40 | 115 | 125 | 227 | 400 | 400 | 454 | 522 | 678 | 230 | 348 | 453 |
| (2) Kedah | 2.51 | - | 82 | 92 | 187 | 360 | 360 | 414 | 482 | 639 | 209 | 319 | 416 |
| (3) Penang mainland | 0.24 | 0.96 | - | 10 | 125 | 291 | 292 | 347 | 415 | 580 | 227 | 311 | 377 |
| (4) Penang island | 0.21 | 0.45 | 0.22 | - | 135 | 300 | 301 | 357 | 425 | 590 | 237 | 321 | 387 |
| (5) Perak | 0.21 | 1.02 | 1.03 | 0.14 | - | 176 | 173 | 213 | 295 | 456 | 210 | 238 | 261 |
| (6) Selangor | 0.38 | 0.65 | 0.32 | 0.36 | 0.28 | - | 22 | 60 | 127 | 307 | 347 | 307 | 216 |
| (7) Kuala Lumpur | 0.40 | 0.84 | 0.54 | 0.42 | 0.39 | 1.47 | - | 54 | 123 | 297 | 333 | 288 | 194 |
| (8) Negeri Sembilan | 0.31 | 0.61 | 0.27 | 0.17 | 0.28 | 0.63 | 0.55 | - | 68 | 247 | 377 | 317 | 195 |
| (9) Melaka | 0.34 | 0.58 | 0.42 | 0.40 | 0.27 | 0.97 | 4.07 | 0.36 | - | 187 | 434 | 360 | 215 |
| (10) Johor | 0.15 | 0.38 | 0.56 | 0.12 | 0.24 | 0.38 | 0.51 | 0.21 | 0.43 | - | 542 | 434 | 264 |
| (11) Kelantan | 0.50 | 0.82 | 0.22 | 0.16 | 0.20 | 0.81 | 0.75 | 0.40 | 0.44 | 0.22 | - | 131 | 282 |
| (12) Terengganu | 0.43 | 0.50 | 0.18 | 0.13 | 0.14 | 0.59 | 0.68 | 0.28 | 0.49 | 0.22 | 1.18 | - | 170 |
| (13) Pahang | 0.25 | 0.37 | 0.17 | 0.11 | 0.12 | 0.54 | 0.70 | 0.24 | 0.46 | 0.23 | 0.64 | 1.52 | - |

*Above the diagonal: geographical distance (KM) among populations of *R. rattus*.
Below the diagonal: gene flow (number of migrants, *Nm*) among populations of *R. rattus*.

between populations; hence, there appear to be at least four individuals migrating between Kuala Lumpur and Melaka per generation, retaining their gene flow as a result. The *Nm* value for Penang island-Pahang populations (with geographical distance of 387 KM) were the lowest, at 0.11, signifying that the gene flow between these two populations was cut-off over time.

The evaluations of *Nst* and *Fst* among populations of *R. rattus*, based on nDNA sequences is shown in Table 4.10. The measures of gene flow (*Nm*) and geographical distance (KM) among populations of *R. rattus*, based on nDNA sequences is shown in Table 4.11. Both *Nst* and *Fst* analysis also outcomes demonstrated populations of Penang mainland-Perak revealed the lowest *Nst* and *Fst*, at 0.00133 and 0.00123, respectively. The division of Penang island-Pahang populations exhibited the highest *Nst* and *Fst* of 0.29532 and 0.29437, respectively. Also, *Nm* analysis proved both *Nst* and *Fst* results, as populations of Penang mainland-Perak (with geographical distance of 125 KM) had the highest *Nm*, with the value of 407.50. There seem to be at least 407 individuals migrating between Penang mainland and Perak per generation, preserving their gene flow as a result. The *Nm* value for Penang island-Pahang (with geographical distance of 387 KM) populations were the lowest, at 1.20, suggesting a minute gene flow occurring between both populations.

The population genetic structure was analysed with an analogue of *F*-statistic, implemented in analysis of molecular variance (AMOVA), via Arlequin Version 3.5. From the analysis, the concatenated mtDNA sequences of COI, Cyt b, and D-loop demonstrated a statistically significant population subdivision of *R. rattus*, with the value of 55.02% variation among populations and 44.98% variation within populations. The fixation indices ($F_{ST}$) of 0.55 was observed in the entire populations of *R. rattus* from

**Table 4.10:** Measures of nucleotide subdivision (*Nst*) and population subdivision (*Fst*) among populations of *R. rattus*, based on nDNA sequences.

| **Populations** | Perlis ($N = 9$) | Kedah ($N = 12$) | Penang mainland ($N = 7$) | Penang island ($N = 6$) | Perak ($N = 8$) | Selangor ($N = 9$) | Kuala Lumpur ($N = 15$) | Negeri Sembilan ($N = 12$) | Melaka ($N = 10$) | Johor ($N = 12$) | Kelantan ($N = 8$) | Tereng-ganu ($N = 10$) | Pahang ($N = 12$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) Perlis | - | 0.07085 | 0.18059 | 0.16324 | 0.05031 | 0.15360 | 0.26639 | 0.24733 | 0.15246 | 0.21295 | 0.28248 | 0.27019 | 0.26194 |
| (2) Kedah | 0.07064 | - | 0.07885 | 0.19507 | 0.01776 | 0.16394 | 0.27957 | 0.15522 | 0.18034 | 0.21157 | 0.23403 | 0.20193 | 0.23699 |
| (3) Penang mainland | 0.17977 | 0.07872 | - | 0.16837 | 0.00133 | 0.21537 | 0.27137 | 0.17394 | 0.23161 | 0.25553 | 0.27341 | 0.28220 | 0.25821 |
| (4) Penang island | 0.16298 | 0.19474 | 0.16795 | - | 0.06790 | 0.22536 | 0.19510 | 0.28570 | 0.23050 | 0.24320 | 0.28186 | 0.27436 | 0.29532 |
| (5) Perak | 0.05018 | 0.01761 | 0.00123 | 0.06798 | - | 0.12785 | 0.23205 | 0.13459 | 0.14819 | 0.17476 | 0.20764 | 0.17736 | 0.16682 |
| (6) Selangor | 0.15315 | 0.16325 | 0.21460 | 0.22469 | 0.12729 | - | 0.19209 | 0.18238 | 0.09577 | 0.09143 | 0.07501 | 0.12535 | 0.12027 |
| (7) Kuala Lumpur | 0.26562 | 0.27880 | 0.27063 | 0.19501 | 0.23125 | 0.19142 | - | 0.25003 | 0.17072 | 0.18867 | 0.19014 | 0.21214 | 0.18041 |
| (8) Negeri Sembilan | 0.24689 | 0.15528 | 0.17376 | 0.28511 | 0.13440 | 0.18203 | 0.24968 | - | 0.25325 | 0.21722 | 0.22201 | 0.23088 | 0.24304 |
| (9) Melaka | 0.15235 | 0.18006 | 0.23112 | 0.23018 | 0.14794 | 0.09573 | 0.17064 | 0.25278 | - | 0.10838 | 0.16852 | 0.11232 | 0.09130 |
| (10) Johor | 0.21234 | 0.21097 | 0.25475 | 0.24263 | 0.17419 | 0.09136 | 0.18841 | 0.21682 | 0.10849 | - | 0.09132 | 0.09019 | 0.06281 |
| (11) Kelantan | 0.28166 | 0.23336 | 0.27263 | 0.28132 | 0.20700 | 0.07503 | 0.19000 | 0.22174 | 0.16855 | 0.09149 | - | 0.04955 | 0.05510 |
| (12) Terengganu | 0.26931 | 0.20123 | 0.28132 | 0.27381 | 0.17677 | 0.12500 | 0.21172 | 0.23059 | 0.11248 | 0.09022 | 0.04962 | - | 0.00803 |
| (13) Pahang | 0.26103 | 0.23608 | 0.25736 | 0.29437 | 0.16614 | 0.11979 | 0.18006 | 0.24261 | 0.09152 | 0.06302 | 0.05523 | 0.00777 | - |

*Above the diagonal: nucleotide subdivision (*Nst*).
Below the diagonal: population subdivision (*Fst*).

**Table 4.11:** Measures of gene flow (number of migrants, *Nm*) and geographical distance (KM) among populations of *R. rattus*, based on nDNA sequences.

| **Populations** | Perlis ($N = 9$) | Kedah ($N = 12$) | Penang mainland ($N = 7$) | Penang island ($N = 6$) | Perak ($N = 8$) | Selangor ($N = 9$) | Kuala Lumpur ($N = 15$) | Negeri Sembilan ($N = 12$) | Melaka ($N = 10$) | Johor ($N = 12$) | Kelantan ($N = 8$) | Tereng-ganu ($N = 10$) | Pahang ($N = 12$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) Perlis | - | 40 | 115 | 125 | 227 | 400 | 400 | 454 | 522 | 678 | 230 | 348 | 453 |
| (2) Kedah | 6.58 | - | 82 | 92 | 187 | 360 | 360 | 414 | 482 | 639 | 209 | 319 | 416 |
| (3) Penang mainland | 2.28 | 5.85 | - | 10 | 125 | 291 | 292 | 347 | 415 | 580 | 227 | 311 | 377 |
| (4) Penang island | 2.57 | 2.07 | 2.48 | - | 135 | 300 | 301 | 357 | 425 | 590 | 237 | 321 | 387 |
| (5) Perak | 9.46 | 28.89 | 407.50 | 6.86 | - | 176 | 173 | 213 | 295 | 456 | 210 | 238 | 261 |
| (6) Selangor | 2.76 | 2.56 | 1.83 | 1.73 | 3.43 | - | 22 | 60 | 127 | 307 | 347 | 307 | 216 |
| (7) Kuala Lumpur | 1.38 | 1.29 | 1.35 | 2.06 | 1.66 | 2.11 | - | 54 | 123 | 297 | 333 | 288 | 194 |
| (8) Negeri Sembilan | 1.53 | 2.72 | 2.38 | 1.25 | 3.22 | 2.25 | 1.50 | - | 68 | 247 | 377 | 317 | 195 |
| (9) Melaka | 2.78 | 2.28 | 1.66 | 1.67 | 2.88 | 4.72 | 2.43 | 1.48 | - | 187 | 434 | 360 | 215 |
| (10) Johor | 1.85 | 1.87 | 1.46 | 1.56 | 2.37 | 4.97 | 2.15 | 1.81 | 4.11 | - | 542 | 434 | 264 |
| (11) Kelantan | 1.28 | 1.64 | 1.33 | 1.28 | 1.94 | 6.16 | 2.13 | 1.75 | 2.47 | 4.96 | - | 131 | 282 |
| (12) Terengganu | 1.36 | 1.98 | 1.28 | 1.33 | 2.33 | 3.50 | 1.86 | 1.67 | 3.95 | 5.04 | 9.58 | - | 170 |
| (13) Pahang | 1.42 | 1.62 | 1.44 | 1.20 | 2.51 | 3.67 | 2.28 | 1.56 | 4.96 | 7.43 | 8.55 | 64.82 | - |

*Above the diagonal: geographical distance (KM) among populations of *R. rattus.*
Below the diagonal: gene flow (number of migrants, *Nm*) among populations of *R. rattus*.

Peninsular Malaysia (refer Table 4.12). Here, populations of *R. rattus* in Peninsular Malaysia demonstrated a higher significant genetic differentiation, based on mtDNA sequences. Mantel test disclosed that there was relatively strong positive insignificant relationship between mtDNA nucleotide divergence and geographic distance among populations of *R. rattus* in Peninsular Malaysia, with correlation coefficient, $r = 0.5565$, $P = 0.061$.

Analysis of AMOVA based on nDNA sequences of IRBP is shown in Table 4.13. The findings signified a statistically significant population subdivision of *R. rattus* in Peninsular Malaysia, with $F_{ST}$ value of 0.19. The variation among populations attributed to 18.57% and the variation within populations attributed to 81.43%. However, the whole populations of *R. rattus* in Peninsular Malaysia demonstrated a lower significant genetic differentiation based on nDNA sequences, as compared to mtDNA sequences. The model of isolation-by-distance (IBD) was also denied through nDNA sequences, following Mantel test which revealed positive insignificant correlation between nucleotide divergence and geographic distance among populations of *R. rattus* in Peninsular Malaysia, with correlation coefficient, $r = 0.2127$, $P = 0.093$.

Statistic of mtDNA (COI, Cyt b, D-loop) sequence variations of *R. rattus* populations is shown in Table 4.14. Populations originating from Penang mainland and Melaka exhibited the highest haplotype diversity (*Hd* = 1.0000), with Penang island population having considerably the lowest haplotype diversity (*Hd* = 0.73333); this coincides with the nucleotide diversity ($\pi$) of the haplotype, where population of Penang island having the lowest value of 0.00102. Population of Kuala Lumpur exhibited the highest $\pi$ of 0.00532. Overall, the entire populations of *R. rattus* in Peninsular Malaysia disclosed high level of genetic diversity, supported with $\pi = 0.00709$ and *Hd* = 0.99237.

**Table 4.12:** Analysis of molecular variance of *R. rattus* populations, based on mtDNA sequences**.**

| **Source of variation** | **d.f.** | **Sum of squares** | **Variance components** | **Percentage of variation** | ***p* value** |
|---|---|---|---|---|---|
| Among populations | 12 | 634.744 | 4.91187 Va | 55.02 | $p < 0.001$ |
| Within populations | 117 | 469.856 | 4.01587 Vb | 44.98 | $p < 0.001$ |
| Total | 129 | 1104.600 | 8.92774 | | |
| Fixation Indices, $F_{ST} = 0.55018$ | | | | | |

*significant ($p < 0.05$).

**Table 4.13:** Analysis of molecular variance of *R. rattus* populations, based on nDNA sequences**.**

| **Source of variation** | **d.f.** | **Sum of squares** | **Variance components** | **Percentage of variation** | ***p* value** |
|---|---|---|---|---|---|
| Among populations | 12 | 176.516 | 1.02600 Va | 18.57 | $p < 0.001$ |
| Within populations | 117 | 526.461 | 4.49967 Vb | 81.43 | $p < 0.001$ |
| Total | 129 | 702.977 | 5.52566 | | |
| Fixation Indices, $F_{ST} = 0.18568$ | | | | | |

*significant ($p < 0.05$).

**Table 4.14:** Summary statistic of mitochondrial DNA sequence variations in 13 populations of *R. rattus*.

| **Populations** | **No. of individuals (*N*)** | **No. of haplotypes (*H*)** | **No. of segregating sites (*S*)** | **Haplotype diversity (*Hd*)** | **Nucleotide diversity ($\pi$)** | **Average number of nucleotide differences (*k*)** | **Tajima's statistic (*D*)** | **Fu's statistic (*Fs*)** | **Fu & Li's statistic (*D**)** | **Fu & Li's statistic (*F**)** |
|---|---|---|---|---|---|---|---|---|---|---|
| Perlis | 9 | 6 | 17 | 0.83333 | 0.00285 | 6.88889 | 0.49340 | 0.708 | 1.32923 | 1.26260 |
| Kedah | 12 | 9 | 23 | 0.95455 | 0.00476 | 11.50000 | 2.26064* | 0.032 | 1.37976* | 1.83315* |
| Penang mainland | 7 | 7 | 11 | 1.00000 | 0.00185 | 4.47619 | -0.01639 | -3.286 | -0.22716 | -0.19743 |
| Penang island | 6 | 3 | 5 | 0.73333 | 0.00102 | 2.46667 | 0.70767 | 1.420 | 0.93883 | 0.94741 |
| Perak | 8 | 4 | 5 | 0.78571 | 0.00104 | 2.50000 | 1.34421 | 0.613 | 0.74709 | 0.97549 |
| Selangor | 9 | 8 | 24 | 0.97222 | 0.00458 | 11.05556 | 1.25129 | -0.834 | 1.46934* | 1.68608* |
| Kuala Lumpur | 15 | 12 | 31 | 0.94286 | 0.00532 | 12.83810 | 1.46120 | -1.222 | 1.30221 | 1.54988 |
| Negeri Sembilan | 12 | 8 | 25 | 0.92424 | 0.00342 | 8.25758 | -0.01125 | 0.249 | 0.18018 | 0.14817 |
| Melaka | 10 | 10 | 30 | 1.00000 | 0.00466 | 11.24444 | 0.28973 | -3.153 | 0.64769 | 0.63015 |
| Johor | 12 | 9 | 8 | 0.93939 | 0.00153 | 3.69697 | 1.57955 | -3.226 | 0.90472 | 1.22377 |
| Kelantan | 8 | 5 | 15 | 0.89286 | 0.00328 | 7.92857 | 1.89576 | 1.826 | 1.08634 | 1.41655 |
| Terengganu | 10 | 8 | 22 | 0.95556 | 0.00337 | 8.13333 | 0.21682 | -0.773 | 1.38569* | 1.23445 |
| Pahang | 12 | 4 | 12 | 0.80303 | 0.00270 | 6.53030 | 2.70283* | 4.723 | 1.58445* | 2.03328* |
| All populations | 130 | 91 | 45 | 0.99237 | 0.00709 | 17.12558 | 3.15316* | -34.256 | 2.19815* | 3.11791* |

*$P < 0.10$ (for *D*, *D**, and *F**). Significance was determined using coalescent simulations in DNASP v5.

The neutrality tests of *R. rattus* based on mtDNA sequences was examined by employing Tajima's test of neutrality, *D*, Fu & Li's *D** and *F**, and Fu's *Fs* statistics. Pahang populations exhibited the highest significant value of *D*, *D**, and *F**, and *Fs*, with the value of 2.70283, 1.58445, 2.03328, and 4.723, respectively. Population of Penang mainland exhibited the insignificant lowest value of *D*, *D**, and *F**, and *Fs*, with the value of -0.01639, -0.22716, -0.19743, and -3.286, respectively. Altogether, entire populations of *R. rattus* from Peninsular Malaysia demonstrated significant value of *D*, *D**, and *F**, and *Fs*, with the value of 3.15316, 2.19815, 3.11791, and -34.256, respectively.

The historic demographic parameters of *R. rattus* populations based on mtDNA sequences is shown in Table 4.15. Small differences were observed between population before expansion ($\theta_0$) and population after expansion ($\theta_1$), which were 3.681 and 54.619, respectively. The relative time since population expansion ($\tau$) value noted for populations of *R. rattus* in Peninsular Malaysia was 17.496. The mismatch distribution analysis of pairwise nucleotide differences among mtDNA sequences for the whole populations of *R. rattus* in Peninsular Malaysia (Figure 4.11) has signified unimodal distribution characteristic (*SSD* = 0.00196; *P* = 0.72). This illustrated that the observed distribution is in agreement with the expected population expansion model. The insignificant *P*-values of the *SSD* stipulated that the peak fit the expected distributions under sudden expansion model. Non-significant raggedness values (*P* = 1.00) were shown in entire populations of *R. rattus*. The outcomes of neutrality tests coincided with the outcomes of mismatch distribution analysis, and the findings accepted the null hypothesis of recent population expansion of *R. rattus*, based on mtDNA sequences.

**Table 4.15:** Historic demographic parameters of *R. rattus* based on mtDNA (COI, Cyt b, D-loop) gene.

| **Populations** | **Relative time since population expansion ($\tau$)** | **Population before expansion ($\theta_0$)** | **Population after expansion ($\theta_1$)** | **Sum of Squared deviation (*SSD*)** | **Harpending's raggedness index** |
|---|---|---|---|---|---|
| Perlis | 15.732 | 0.000 | 9.555 | 0.06770 ($P = 0.33$) | 0.17593 ($P = 0.18$) |
| Kedah | 19.512 | 0.005 | 25.795 | 0.03130 ($P = 0.31$) | 0.07415 ($P = 0.28$) |
| Penang mainland | 5.766 | 0.005 | 21.328 | 0.01791 ($P = 0.80$) | 0.04989 ($P = 0.92$) |
| Penang island | 5.250 | 0.002 | 4.690 | 0.12744 ($P = 0.12$) | 0.29333 ($P = 0.34$) |
| Perak | 4.109 | 0.002 | 14.482 | 0.09801 ($P = 0.07$) | 0.25510 ($P = 0.17$) |
| Selangor | 14.412 | 0.000 | 40.719 | 0.05091 ($P = 0.18$) | 0.16281 ($P = 0.13$) |
| Kuala Lumpur | 17.719 | 0.000 | 33.459 | 0.01470 ($P = 0.60$) | 0.02685 ($P = 0.52$) |
| Negeri Sembilan | 1.018 | 10.169 | 18.186 | 0.04228 ($P = 0.39$) | 0.07186 ($P = 0.42$) |
| Melaka | 10.664 | 3.514 | 30.791 | 0.00948 ($P = 0.97$) | 0.01383 ($P = 0.99$) |
| Johor | 2.850 | 1.705 | 31.523 | 0.03217 ($P = 0.16$) | 0.11111 ($P = 0.16$) |
| Kelantan | 14.461 | 0.000 | 22.173 | 0.09451 ($P = 0.03$) | 0.12117 ($P = 0.41$) |
| Terengganu | 0.992 | 8.694 | 99999.000 | 0.03665 ($P = 0.42$) | 0.07012 ($P = 0.35$) |
| Pahang | 13.000 | 0.000 | 21.119 | 0.16929 ($P = 0.01$) | 0.21878 ($P = 0.07$) |
| All populations | 17.496 | 3.681 | 54.619 | 0.00196 ($P = 0.72$) | 0.00109 ($P = 1.00$) |

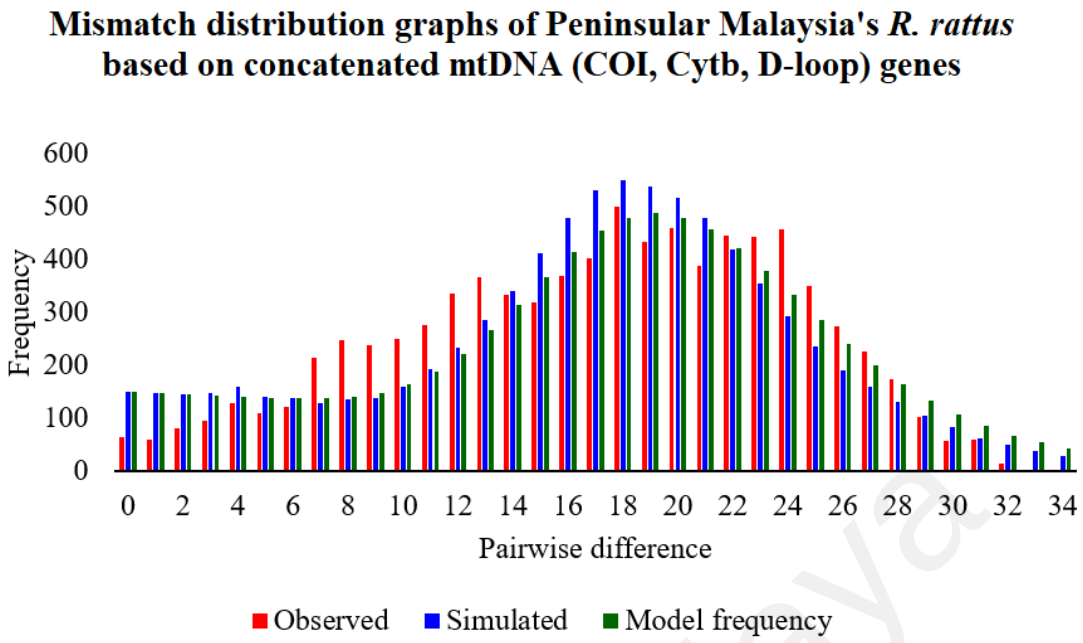


**Figure 4.11:** Mismatch distribution of pairwise nucleotide differences of mtDNA sequences of *R. rattus* with unimodal distribution.

Statistic of nDNA (IRBP) sequence variations of *R. rattus* populations is shown in Table 4.16. Populations originating from Kedah, Penang mainland and Perak exhibited the highest haplotype diversity (*Hd* = 1.0000), with Melaka population having considerably the lowest haplotype diversity (*Hd* = 0.84444). Population of Kelantan showed the lowest value of $\pi$, at 0.00565, while population of Selangor exhibited the highest $\pi$ value of 0.00847. Of nDNA sequences, overall populations of *R. rattus* in Peninsular Malaysia demonstrated high level of genetic diversity, with $\pi$ of 0.00888 and *Hd* of 0.99380.

The neutrality tests of *R. rattus* based on nDNA sequences demonstrated Johor populations exhibited the highest insignificant value of *D*, *D**, and *F**, and *Fs*, with the value of 1.36108, 1.37976, 1.56546, and 2.098, respectively. Population of Kelantan exhibited the insignificant lowest value of *D*, *D**, and *F**, and *Fs*, with the value of -0.01012, -0.02869, -0.01383, and -5.589, respectively. Overall, nDNA sequences

**Table 4.16:** Summary statistic of nuclear DNA sequence variations in 13 populations of *R. rattus*.

| **Populations** | **No. of individuals (*N*)** | **No. of haplotypes (*H*)** | **No. of segregating sites (*S*)** | **Haplotype diversity (*Hd*)** | **Nucleotide diversity ($\pi$)** | **Average number of nucleotide differences (*k*)** | **Tajima's statistic (*D*)** | **Fu's statistic (*Fs*)** | **Fu & Li's statistic (*D**)** | **Fu & Li's statistic (*F**)** |
|---|---|---|---|---|---|---|---|---|---|---|
| Perlis | 9 | 6 | 20 | 0.91667 | 0.00697 | 8.55556 | 0.79981 | 1.201 | 0.80994 | 0.90482 |
| Kedah | 12 | 12 | 24 | 1.00000 | 0.00693 | 8.50000 | 0.30908 | -1.185 | 0.30307 | 0.34690 |
| Penang mainland | 7 | 7 | 21 | 1.00000 | 0.00761 | 9.33333 | 0.50020 | -1.687 | 0.48851 | 0.54158 |
| Penang island | 6 | 5 | 20 | 0.93333 | 0.00831 | 10.20000 | 1.02772 | 0.851 | 0.88917 | 1.00105 |
| Perak | 8 | 8 | 23 | 1.00000 | 0.00821 | 10.07143 | 0.70976 | -2.160 | 0.49500 | 0.60681 |
| Selangor | 9 | 8 | 24 | 0.97222 | 0.00847 | 10.38889 | 0.87638 | -0.962 | 0.94648 | 1.04215 |
| Kuala Lumpur | 15 | 8 | 27 | 0.87619 | 0.00762 | 9.35238 | 0.52824 | 1.690 | 0.52167 | 0.60246 |
| Negeri Sembilan | 12 | 7 | 25 | 0.90909 | 0.00722 | 8.86364 | 0.31492 | 1.501 | 0.87919 | 0.83315 |
| Melaka | 10 | 6 | 25 | 0.84444 | 0.00804 | 9.86667 | 0.55461 | 1.835 | 0.61442 | 0.67626 |
| Johor | 12 | 7 | 23 | 0.87879 | 0.00811 | 9.95455 | 1.36108 | 2.098 | 1.37976* | 1.56546 |
| Kelantan | 8 | 7 | 18 | 0.96429 | 0.00565 | 6.92857 | -0.01012 | -5.589 | -0.02869 | -0.01383 |
| Terengganu | 10 | 9 | 20 | 0.97778 | 0.00605 | 7.42222 | 0.23439 | -2.390 | -0.01314 | 0.04144 |
| Pahang | 12 | 11 | 20 | 0.98485 | 0.00654 | 8.03030 | 0.93353 | -3.597 | 0.90693 | 1.04021 |
| All populations | 130 | 95 | 32 | 0.99380 | 0.00888 | 10.89887 | 2.53837* | -95.634 | 2.00920* | 2.66741* |

*$P < 0.10$ (for *D*, *D**, and *F**). Significance was determined using coalescent simulations in DNASP v5.

demonstrated the whole populations of *R. rattus* in Peninsular Malaysia portrayed significant value of *D*, *D**, and *F**, and *Fs*, with the value of 2.53837, 2.00920, 2.66741, and -95.634, respectively.

The historic demographic parameters of *R. rattus* populations based on nDNA sequences is shown in Table 4.17. Small differences were noticed between $\theta_0$ and $\theta_1$, from 0.000 to 41.758, respectively. The tau ($\tau$) value observed for populations of *R. rattus* in Peninsular Malaysia was 11.266. The mismatch distribution analysis of pairwise nucleotide differences among nDNA sequences for the whole populations of *R. rattus* in Peninsular Malaysia (Figure 4.12) also demonstrated unimodal curve (*SSD* = 0.00183; *P* = 0.10). This showed that the observed distribution is in parallel with the expected population expansion model. The insignificant *P*-values of the *SSD* stipulated that the curve match the expected distributions under sudden expansion model. The entire populations of *R. rattus* from Peninsular Malaysia also exhibited a non-significant raggedness value (*P* = 0.10). The findings of neutrality tests were in agreement with the findings of mismatch distribution analysis, and the outcomes also accepted the null hypothesis of recent population expansion of *R. rattus*, based on nDNA sequences.

**Table 4.17:** Historic demographic parameters of *R. rattus* based on nDNA (IRBP) gene.

| **Populations** | **Relative time since population expansion ($\tau$)** | **Population before expansion ($\theta_0$)** | **Population after expansion ($\theta_1$)** | **Sum of Squared deviation (*SSD*)** | **Harpending's raggedness index** |
|---|---|---|---|---|---|
| Perlis | 10.031 | 0.000 | 2332.500 | 0.03998 ($P = 0.15$) | 0.10725 ($P = 0.14$) |
| Kedah | 8.691 | 0.025 | 313.750 | 0.00971 ($P = 0.65$) | 0.03352 ($P = 0.52$) |
| Penang mainland | 4.953 | 0.000 | 99999.000 | 0.09319 ($P = 0.15$) | 0.05442 ($P = 0.82$) |
| Penang island | 16.029 | 0.000 | 43.420 | 0.08122 ($P = 0.09$) | 0.25333 ($P = 0.29$) |
| Perak | 10.293 | 0.004 | 99999.000 | 0.02516 ($P = 0.48$) | 0.07908 ($P = 0.30$) |
| Selangor | 10.943 | 0.000 | 99999.000 | 0.01970 ($P = 0.52$) | 0.06404 ($P = 0.35$) |
| Kuala Lumpur | 12.375 | 0.000 | 34.414 | 0.03915 ($P = 0.06$) | 0.09297 ($P = 0.02$) |
| Negeri Sembilan | 4.145 | 9.492 | 30.596 | 0.05362 ($P = 0.07$) | 0.08678 ($P = 0.26$) |
| Melaka | 11.729 | 0.000 | 36.237 | 0.08104 ($P = 0.03$) | 0.16247 ($P = 0.04$) |
| Johor | 11.418 | 0.000 | 224.688 | 0.06651 ($P = 0.01$) | 0.13728 ($P = 0.01$) |
| Kelantan | 7.711 | 0.086 | 55.547 | 0.02554 ($P = 0.60$) | 0.07781 ($P = 0.51$) |
| Terengganu | 7.566 | 0.000 | 287.812 | 0.01611 ($P = 0.51$) | 0.04198 ($P = 0.46$) |
| Pahang | 8.660 | 0.000 | 99999.000 | 0.01091 ($P = 0.47$) | 0.02365 ($P = 0.65$) |
| All populations | 11.266 | 0.000 | 41.758 | 0.00183 ($P = 0.10$) | 0.00611 ($P = 0.10$) |

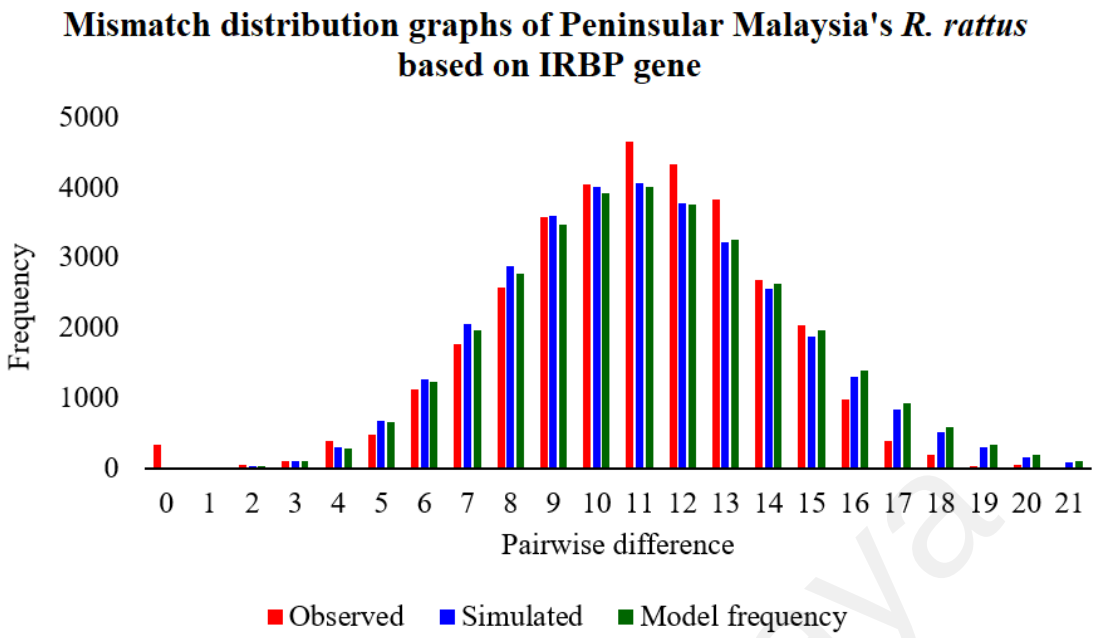


**Figure 4.12:** Mismatch distribution of pairwise nucleotide differences of nDNA sequences of *R. rattus* with unimodal distribution.

## 4.3 Morphometric Analysis of Skull Characteristics of *Rattus rattus*

A total of 130 skulls of *Rattus rattus* which is proportional to individuals of age groups varying from C2 to C4, were employed in the statistical analysis. MANOVA and ANOVA were performed for craniodental measurements of *R. rattus* versus age and sex which displayed that craniodental measurements were significantly correlated with age ($P < 0.0001$) and sex ($P < 0.0001$). This shows the presence of sexual dimorphism. Consequently, the data were analysed separately for males and females. The sex ratio was quite equilibrated with 67 males and 63 females. Table 4.18 and Table 4.19 shows the 14 craniodental measurements of the male and female *R. rattus*, respectively. Generally, males portrayed higher craniodental measurements compared to females. Table 4.20 shows the difference in craniodental measurements for males and females. All variables showed significant difference between males and females except ML.

**Table 4.18:** The 14 craniodental measurements (in mm) of the 67 samples of male *R. rattus* collected in Peninsular Malaysia. Sample size (*N*), mean, minimum (min), maximum (max), and standard deviation (*SD*) value are given for each character.

| Variables | Male (*N* = 67) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | C2 (*N* = 15) | | | | C3 (*N* = 16) | | | | C4 (*N* = 36) | | | |
| | Mean | Min | Max | *SD* | Mean | Min | Max | *SD* | Mean | Min | Max | *SD* |
| **ONL** | 42.25 | 41.23 | 43.81 | 0.89 | 42.32 | 40.03 | 46.03 | 1.46 | 47.76 | 46.39 | 48.86 | 0.65 |
| **BBC** | 15.92 | 14.55 | 18.35 | 0.98 | 16.12 | 11.93 | 18.20 | 1.45 | 19.12 | 16.47 | 20.00 | 0.71 |
| **LD** | 12.20 | 8.95 | 14.06 | 1.35 | 11.43 | 10.18 | 13.74 | 1.00 | 16.86 | 13.27 | 18.39 | 1.18 |
| **LIF** | 7.21 | 6.25 | 8.19 | 0.56 | 7.07 | 5.26 | 8.28 | 0.78 | 9.62 | 8.69 | 10.81 | 0.71 |
| **BM1** | 1.89 | 1.56 | 2.60 | 0.24 | 1.88 | 1.46 | 2.22 | 0.22 | 2.81 | 1.70 | 3.29 | 0.36 |
| **LBP** | 8.07 | 6.26 | 10.39 | 1.11 | 7.17 | 6.24 | 8.19 | 0.57 | 9.90 | 9.09 | 10.27 | 0.30 |
| **LB** | 6.78 | 4.61 | 7.65 | 0.76 | 6.94 | 5.87 | 7.94 | 0.49 | 8.24 | 7.86 | 8.92 | 0.29 |
| **BMF** | 3.22 | 2.55 | 4.11 | 0.45 | 2.96 | 2.31 | 3.84 | 0.45 | 4.77 | 3.71 | 5.02 | 0.25 |
| **BBP** | 4.10 | 3.36 | 4.93 | 0.44 | 4.14 | 3.25 | 5.00 | 0.45 | 6.12 | 5.01 | 6.89 | 0.50 |
| **BZP** | 4.02 | 3.11 | 5.71 | 0.64 | 4.47 | 3.66 | 5.40 | 0.51 | 5.74 | 5.06 | 6.02 | 0.19 |
| **CLM1-3** | 6.81 | 5.65 | 7.44 | 0.52 | 6.73 | 5.80 | 7.45 | 0.53 | 8.25 | 6.81 | 8.91 | 0.45 |
| **HBC** | 10.83 | 9.61 | 11.63 | 0.57 | 12.00 | 11.41 | 12.76 | 0.45 | 12.47 | 12.08 | 12.89 | 0.26 |
| **ML** | 18.91 | 16.50 | 23.88 | 2.22 | 19.47 | 16.02 | 23.42 | 1.97 | 23.53 | 22.60 | 24.39 | 0.57 |
| **M1-M3** | 6.06 | 5.15 | 7.31 | 0.57 | 6.33 | 5.58 | 6.88 | 0.55 | 7.25 | 6.02 | 7.90 | 0.44 |

**Table 4.19:** The 14 craniodental measurements (in mm) of the 63 samples of female *R. rattus* collected in Peninsular Malaysia. Sample size (*N*), mean, minimum (min), maximum (max), and standard deviation (*SD*) value are given for each character.

| Variables | Female (*N* = 63) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | C2 (*N* = 24) | | | | C3 (*N* = 14) | | | | C4 (*N* = 25) | | | |
| | Mean | Min | Max | *SD* | Mean | Min | Max | *SD* | Mean | Min | Max | *SD* |
| **ONL** | 41.63 | 40.01 | 44.88 | 1.15 | 41.38 | 40.10 | 43.20 | 1.11 | 45.74 | 42.71 | 46.80 | 0.99 |
| **BBC** | 15.34 | 14.28 | 16.99 | 0.64 | 16.69 | 15.64 | 17.77 | 0.54 | 17.65 | 17.04 | 18.36 | 0.40 |
| **LD** | 13.37 | 12.26 | 15.17 | 0.64 | 11.51 | 9.85 | 13.31 | 0.89 | 14.99 | 14.64 | 15.38 | 0.19 |
| **LIF** | 7.45 | 6.39 | 8.25 | 0.51 | 7.43 | 6.27 | 8.56 | 0.55 | 8.07 | 7.50 | 8.86 | 0.41 |
| **BM1** | 1.79 | 1.49 | 2.20 | 0.18 | 1.82 | 1.46 | 2.19 | 0.22 | 2.10 | 1.85 | 2.39 | 0.16 |
| **LBP** | 8.19 | 6.99 | 9.26 | 0.51 | 7.11 | 5.71 | 8.78 | 0.93 | 9.23 | 8.27 | 9.90 | 0.31 |
| **LB** | 6.78 | 5.92 | 7.50 | 0.37 | 7.05 | 6.08 | 7.65 | 0.47 | 7.38 | 7.01 | 7.85 | 0.24 |
| **BMF** | 3.30 | 2.76 | 4.48 | 0.37 | 2.91 | 2.11 | 3.65 | 0.48 | 4.30 | 4.04 | 4.63 | 0.19 |
| **BBP** | 4.19 | 3.52 | 5.21 | 0.42 | 4.39 | 3.75 | 5.10 | 0.37 | 5.20 | 4.86 | 5.57 | 0.17 |
| **BZP** | 4.18 | 3.66 | 4.76 | 0.31 | 4.59 | 4.14 | 5.16 | 0.27 | 5.09 | 4.35 | 5.48 | 0.28 |
| **CLM1-3** | 6.48 | 5.69 | 8.21 | 0.59 | 6.72 | 5.82 | 7.31 | 0.46 | 7.41 | 7.00 | 7.78 | 0.21 |
| **HBC** | 11.09 | 10.11 | 12.46 | 0.64 | 11.66 | 10.70 | 13.52 | 0.68 | 11.81 | 11.26 | 12.07 | 0.23 |
| **ML** | 20.49 | 18.89 | 22.98 | 0.89 | 20.88 | 19.54 | 22.24 | 0.82 | 22.05 | 21.47 | 22.59 | 0.30 |
| **M1-M3** | 6.07 | 5.36 | 6.71 | 0.27 | 6.11 | 5.36 | 6.71 | 0.31 | 6.38 | 6.07 | 6.72 | 0.21 |

**Table 4.20:** t-test for difference in mean of craniodental measurements between male and female *R. rattus*.

| **Craniodental measurements** | ***P*** |
|---|---|
| ONL | *** |
| BBC | *** |
| LD | * |
| LIF | *** |
| BM1 | *** |
| LBP | * |
| LB | *** |
| BMF | ** |
| BBP | ** |
| BZP | ** |
| CLM1-3 | *** |
| HBC | *** |
| ML | 0.3416 (NS) |
| M1-M3 | *** |

*** $P < 0.0001$, ** $P < 0.01$, * $P < 0.05$, NS-not significant.

The PCA plots for males (Figure 4.13) and females (Figure 4.14) showed three groups. The first group comprised the specimens with age C2, the second with age C3, and third with age C4. These plots revealed that age groups C2 and C3 shared similar measurements and significant difference can be seen between these groups with age C4.

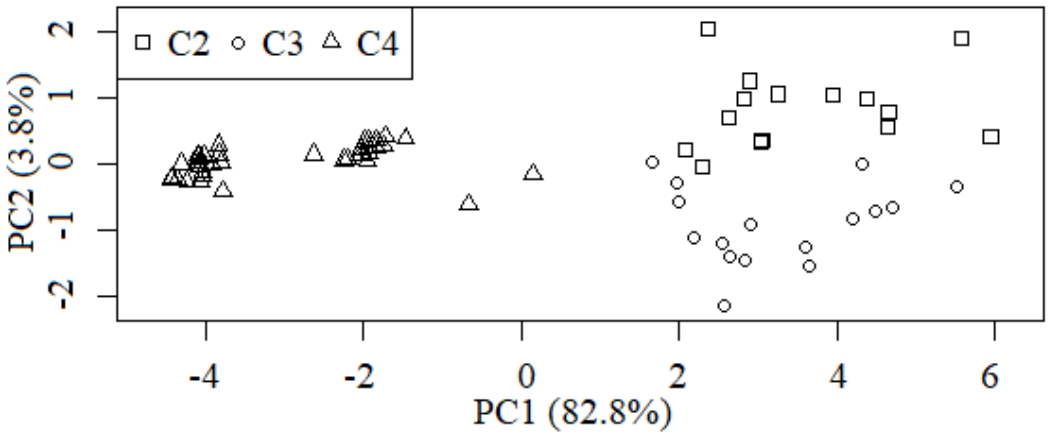


**Figure 4.13:** PC1 and PC2 of the PCA for 14 craniodental measurements of 67 male *R. rattus*.

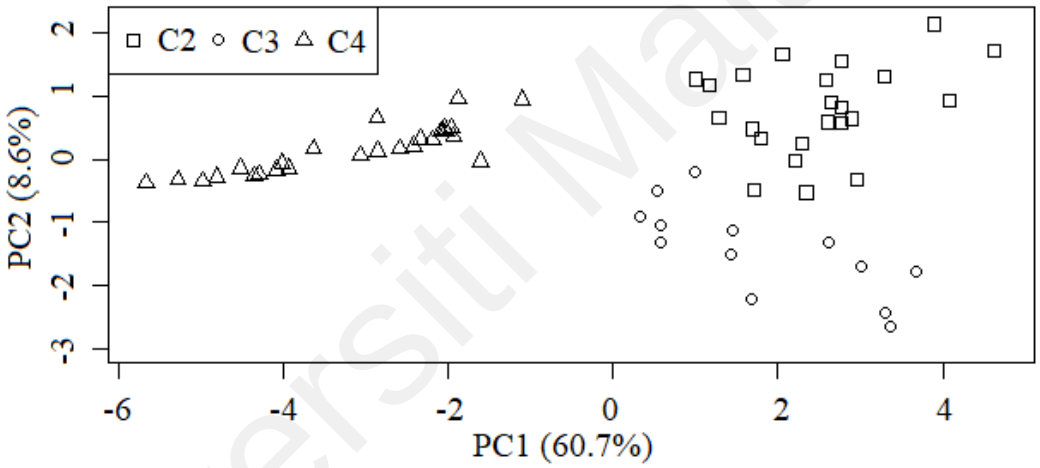


**Figure 4.14:** PC1 and PC2 of the PCA for 14 craniodental measurements of 63 female *R. rattus*.

For males, axes 1 and 2 of the PCA plot explained 86.6% of the total variation where 82.8% of the total variation was described by axis 1 while 3.8% was described by axis 2. Based on the results in Table 4.21, PC1 was negatively correlated with all the variables, PC2 showed strongest negative correlation with HBC, and PC3 showed strongest positive correlation with ML and M1-M3. The linear discriminant analysis (LDA) was performed on the basis of the results of the PCA which showed three age groups (Figure 4.15) with C2 and C3 sharing similar characteristics while C4 was significantly different. The plot

of the linear discriminants, LD1 and LD2 displayed separations between the specimens from age groups C2, C3, and C4. LD1 displayed 77.6% of the total variation while LD2 accounted 22.4%.

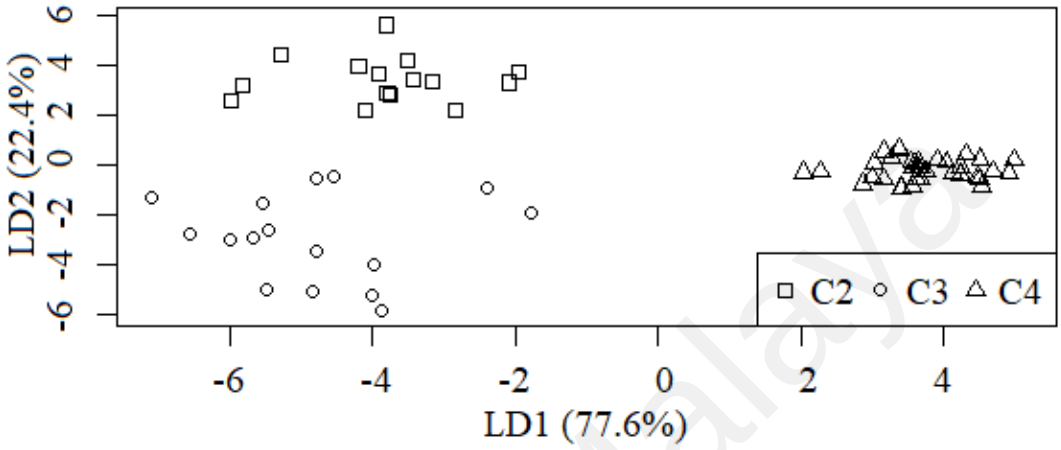


**Figure 4.15:** LDA plot of the 14 craniodental measurements of male *R. rattus*.

For females, the first two axes of the PCA plot accounted to 69.3% of the total variation where axis 1 and axis 2 explained 60.7% and 8.6% of the total variations, respectively. From Table 4.21, PC1 displayed negative correlations with all the variables studied with the highest negative correlation with ONL (-0.3102), PC2 and PC3 showed the highest negative correlation with HBC (PC2: -0.5685, PC3: -0.5219). The highest positive correlations involving PC2 were seen with LD (0.4374) and LBP (0.4419). The LDA for the female *R. rattus* also showed three age groups by displaying three clusters (Figure 4.16). Axis LD1 accounted to 75.5% of the total variation and LD2, 24.5%.

**Table 4.21:** Variable coefficients and percentage variation explained by the first three components (PC1-PC3) of the PCA on 14 standardised craniodental measurements from male and female *R. rattus*.

| Craniodental measurements | Male | | | Female | | |
|---|---|---|---|---|---|---|
| | **PC1** | **PC2** | **PC3** | **PC1** | **PC2** | **PC3** |
| ONL | -0.2805 | 0.0598 | -0.0272 | -0.3102* | 0.0882 | -0.0732 |
| BBC | -0.2723 | 0.0375 | 0.3056 | -0.2967 | -0.2690 | 0.2575 |
| LD | -0.2782 | 0.1496 | 0.0318 | -0.2667 | 0.4374* | -0.3115 |
| LIF | -0.2780 | 0.0991 | -0.1744 | -0.2318 | 0.0537 | 0.0868 |
| BM1 | -0.2696 | 0.1376 | -0.1053 | -0.2745 | 0.0561 | 0.2412 |
| LBP | -0.2656 | 0.3015 | 0.3229 | -0.2473 | 0.4419* | -0.2070 |
| LB | -0.2671 | -0.0173 | -0.0114 | -0.2541 | -0.0425 | 0.4486* |
| BMF | -0.2698 | 0.2625 | -0.1273 | -0.2923 | 0.2145 | -0.2832 |
| BBP | -0.2807 | 0.0199 | -0.0990 | -0.2984 | -0.0493 | -0.0866 |
| BZP | -0.2687 | -0.0499 | -0.1300 | -0.2911 | -0.3404 | -0.1421 |
| CLM1-3 | -0.2781 | 0.1345 | -0.0841 | -0.2883 | 0.0239 | 0.2728 |
| HBC | -0.2230 | -0.7816* | -0.3453 | -0.1567 | -0.5685* | -0.5219* |
| ML | -0.2541 | -0.2033 | 0.5692* | -0.2658 | -0.0562 | 0.2289 |
| M1-M3 | -0.2501 | -0.3325 | 0.5155* | -0.2288 | -0.1847 | -0.1198 |
| % variance | 82.81 | 3.84 | 2.98 | 60.73 | 8.61 | 7.08 |

Note: Asterisk in the table denote variables that explain the highest percentage of variation on each principal component (PC).

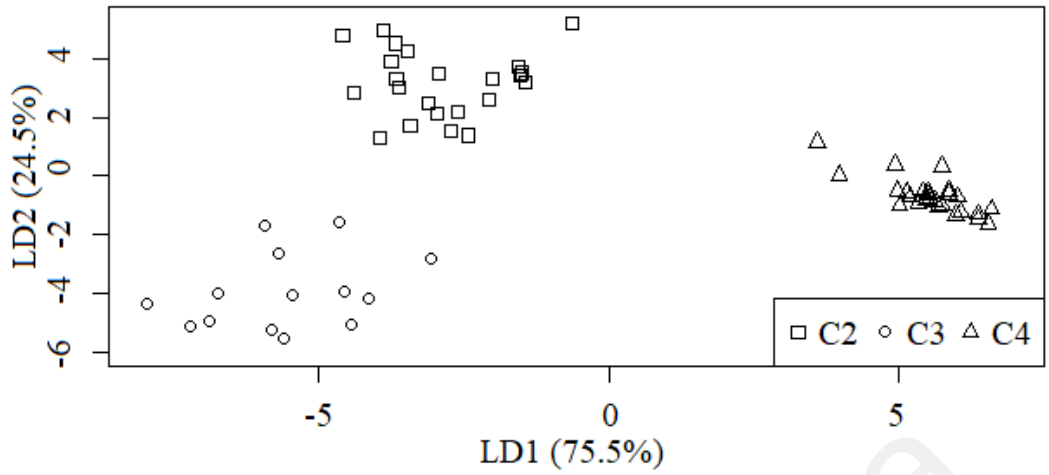


**Figure 4.16:** LDA plot of the 14 craniodental measurements of female *R. rattus*.

Analysis of the 'a posteriori' Scheffé's test were performed to observe if the morphometric variables differed significantly for males and females separately from age groups C2, C3, and C4. Table 4.22 show the results of the 'a posteriori' Scheffé's test comparing 14 craniodental measurements individually between age groups for males and females. It can be seen that age groups C2 and C3 have some craniodental measurement which do not differ significantly. However, the differences became more prominent when C3 and C4 were compared for both genders. This is because not much difference takes place in the growth of *R. rattus* between the C2 and C3 subadult stages. The difference becomes more obvious when *R. rattus* attains C4 adult stage.

**Table 4.22:** Results of the 'a posteriori' Scheffé's test to evaluate the significant differences of craniodental measurements between ages C2, C3, and C4 for male and female *R. rattus*.

| Craniodental measurements | Male | | | Female | | |
|---|---|---|---|---|---|---|
| | C2 and C3 | C2 and C4 | C3 and C4 | C2 and C3 | C2 and C4 | C3 and C4 |
| ONL | 0.9804 (NS) | *** | *** | 0.9425 (NS) | *** | *** |
| BBC | 0.8427 (NS) | *** | *** | *** | *** | *** |
| LD | 0.2004 (NS) | *** | *** | *** | *** | *** |
| LIF | 0.8425 (NS) | *** | *** | 0.9981 (NS) | *** | *** |
| BM1 | 0.9949 (NS) | *** | *** | 0.8773 (NS) | *** | *** |
| LBP | ** | *** | *** | *** | *** | *** |
| LB | 0.6695 (NS) | *** | *** | 0.0512 (NS) | *** | * |
| BMF | 0.1438 (NS) | *** | *** | ** | *** | *** |
| BBP | 0.9721 (NS) | *** | *** | 0.0637 (NS) | *** | *** |
| BZP | * | *** | *** | *** | *** | *** |
| CLM1-3 | 0.8863 (NS) | *** | *** | 0.1442 (NS) | *** | *** |
| HBC | *** | *** | ** | ** | *** | 0.7864 (NS) |
| ML | 0.5627 (NS) | *** | *** | 0.0739 (NS) | *** | *** |
| M1-M3 | 0.3240 (NS) | *** | *** | 0.7546 (NS) | *** | ** |

*** $P < 0.0001$, ** $P < 0.01$, * $P < 0.05$, NS-not significant.

According to the results of the regression in Table 4.23, none of the geographical factors (latitude, longitude, and precipitation) showed statistical significance for both males and females. Thus, there is lack of geographic evidence related to the skull variation of *R. rattus* in Peninsular Malaysia observed in this study.

**Table 4.23:** Regressions of the different skull size parameters of male and female *R. rattus* in relation to the geographical factors in Peninsular Malaysia.

| **Factors** | **Male** | **Female** |
|---|---|---|
| | **Log data (significance)** | **Log data (significance)** |
| Latitude | 0.0780 (NS) | 0.1115 (NS) |
| Longitude | 0.2542 (NS) | 0.3603 (NS) |
| Precipitation | 0.4784 (NS) | 0.3061 (NS) |

* $P < 0.05$ (significant), NS-not significant.

## 4.4 Prevalence and Genetic Divergence of *Leptospira* spp. in *Rattus rattus*

Throughout this study, it is confirmed that *Rattus rattus* are actively involved in the transmission of leptospirosis. Both *Leptospira*'s 16S rRNA and LipL32 sequences were successfully amplified from the kidney and liver tissues of *R. rattus*, with expected amplicon size of 412 and 240 bp, respectively (Figure 4.17).

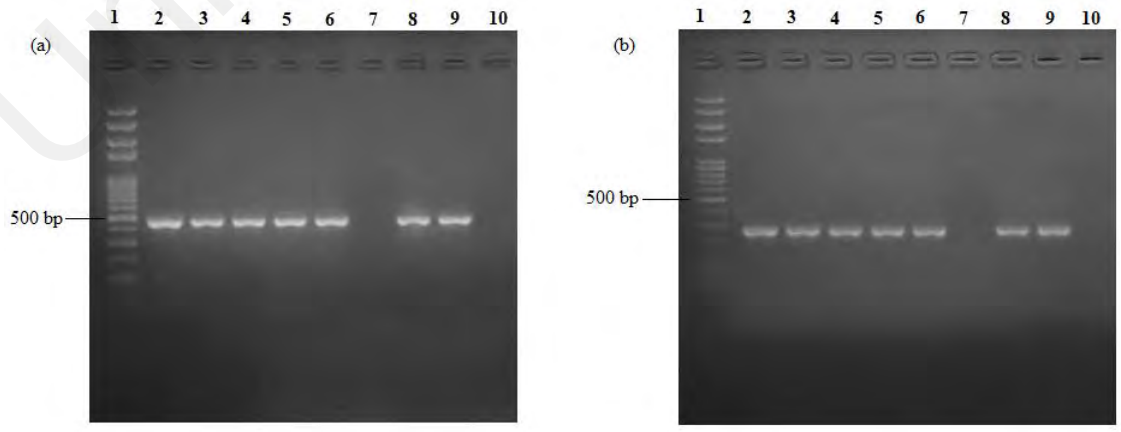


**Figure 4.17:** Detection of leptospiral DNA in kidney and liver of *Rattus rattus* by PCR; (a) 16S rRNA gene, and (b) LipL32 gene. (Lane 1: 100 bp DNA ladder; lanes 2-6: positive samples; lane 7: negative sample; lanes 8-9: positive controls; lane 10: negative control).

Out of the 130 individuals of *R. rattus* studied, a total of 51 (39.2%) individuals were positive for leptospirosis detected by using PCR method (Figure 4.18). The sequences amplified using both primers denoted that all the positive DNA were *Leptospira* spp. and belongs to the pathogenic leptospires (98%-100% maximum identity). The results of nucleotide BLAST showing similarity with 16S rRNA gene of *Leptospira* spp. from GenBank is indicated in Appendix F. Two pathogenic *Leptospira* spp. namely, *L. interrogans* and *L. borgpetersenii* were discovered circulating among *R. rattus*. The most dominant species is *L. borgpetersenii* carried by 29 (56.9%) individuals while another 22 (43.1%) individuals were colonised by *L. interrogans* (Figure 4.19). No intermediate and saprophytic leptospires species were identified in this study. The nucleotide sequences of all the positive controls and samples were deposited into GenBank database with accession numbers of MH997573-MH997626 for 16S rRNA gene and GenBank accession numbers of MK026014-MK026067 for LipL32 gene.

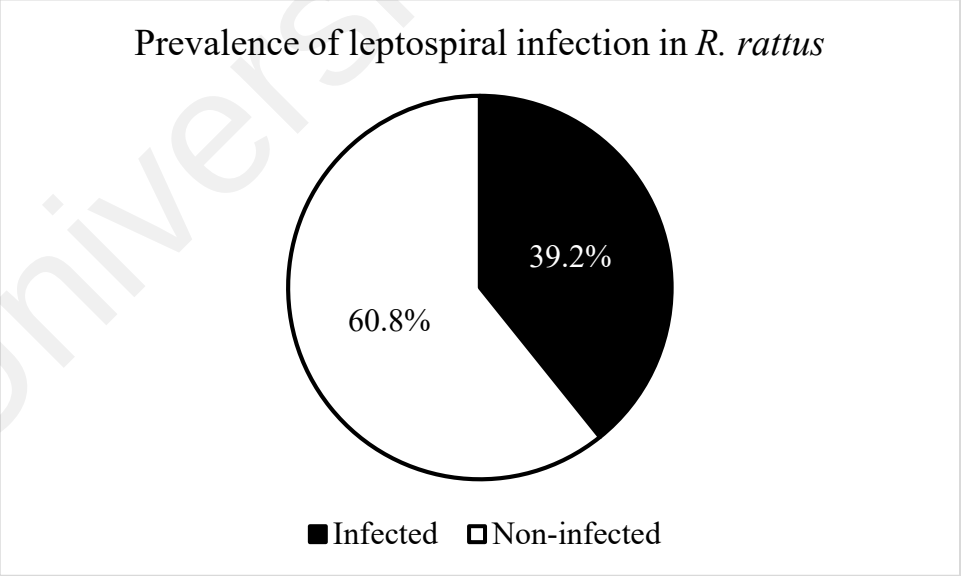


**Figure 4.18:** Prevalence of leptospiral infection in *R. rattus* collected in Peninsular Malaysia from December 2016 to February 2018.

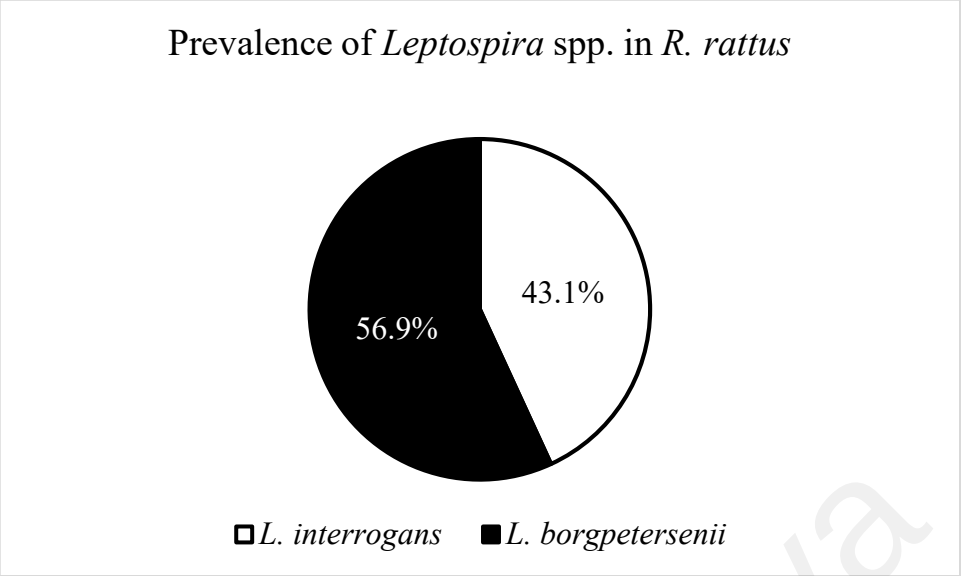


**Figure 4.19:** Prevalence of *Leptospira* spp. circulating among *R. rattus*.

All of the captured *R. rattus* comprised 69 subadult and 61 adult rats. The leptospires infection, according to age group was quite similar with 39.1% and 39.3% occurrence in subadults and adults' rats, respectively. No juvenile rats were trapped in this study. Males *R. rattus* are the most prevalent to leptospires infection compared to females (Figure 4.20). Out of 67 males *R. rattus*, 28 (41.8%) individuals are the vectors for leptospirosis, in which 12 and 16 of them were infected with *L. interrogans* and *L. borgpetersenii*, respectively. From 63 females *R. rattus*, a total of 23 (36.5%) individuals were infected by leptospires, in which ten and 13 individuals harboured *L. interrogans* and *L. borgpetersenii*, respectively.

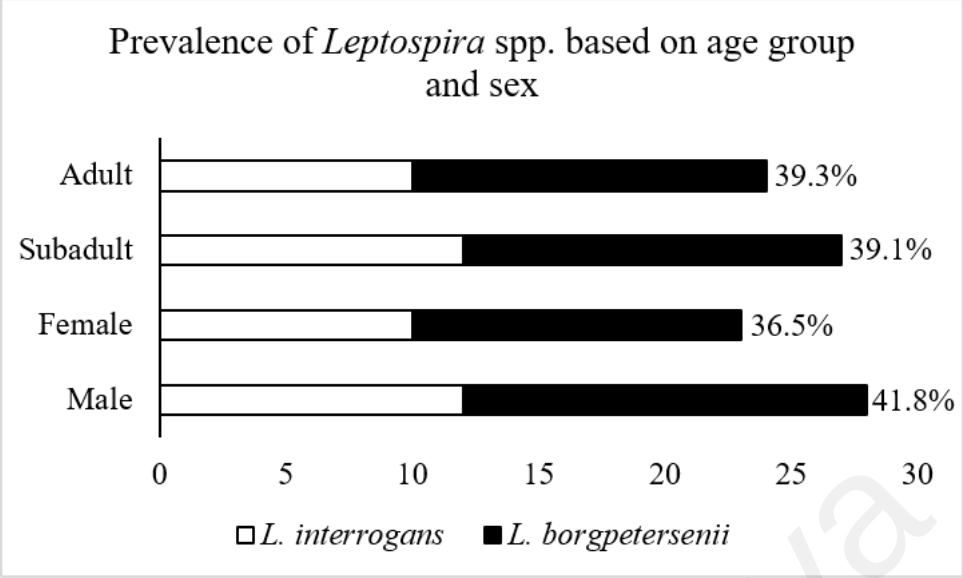


**Figure 4.20:** Prevalence of *Leptospira* spp. based on age group and sex of *R. rattus*.

Based on the 13 cities in Peninsular Malaysia where *R. rattus* were trapped, this study indicates that the most prevalent city with rat's leptospirosis is Penang island (Figure 4.21). Of six caught rats, two individuals each carrying *L. interrogans* and *L. borgpetersenii* were identified. The uncommon city with rat's leptospirosis is represented by Terengganu, in which only one out of ten individuals carrying *L. borgpetersenii* with no evidence of *L. interrogans* infection.

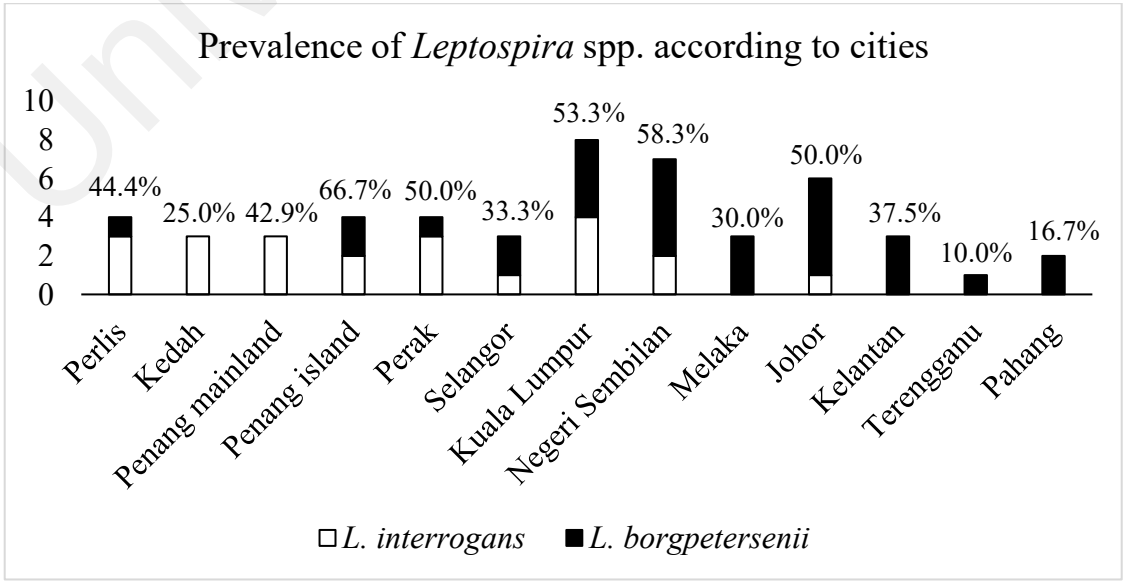


**Figure 4.21:** Prevalence of *Leptospira* spp. according to cities in Peninsular Malaysia.

The sampling areas involved in this study mainly focus on three preferable niches of fresh markets, seaside, and housing areas where populations of *R. rattus* are abundant. The results demonstrated that rats caught from fresh markets are the most prevalent with leptospires infection (Figure 4.22). Of 35 rats caught in fresh markets, 19 (54.3%) of them are carriers of leptospires, in which nine and ten of them carrying *L. interrogans* and *L. borgpetersenii*, respectively. Meanwhile, of total 27 rats caught in seaside niche, 14 (51.9%) of them are infected with leptospirosis, wherein six of them were infected by *L. interrogans* and another eight individuals corresponded to *L. borgpetersenii*. However, with the highest number of 68 rats trapped in the housing areas, only 18 (26.5%) individuals harbouring leptospires. Seven and 11 of them carrying *L. interrogans* and *L. borgpetersenii*, respectively.

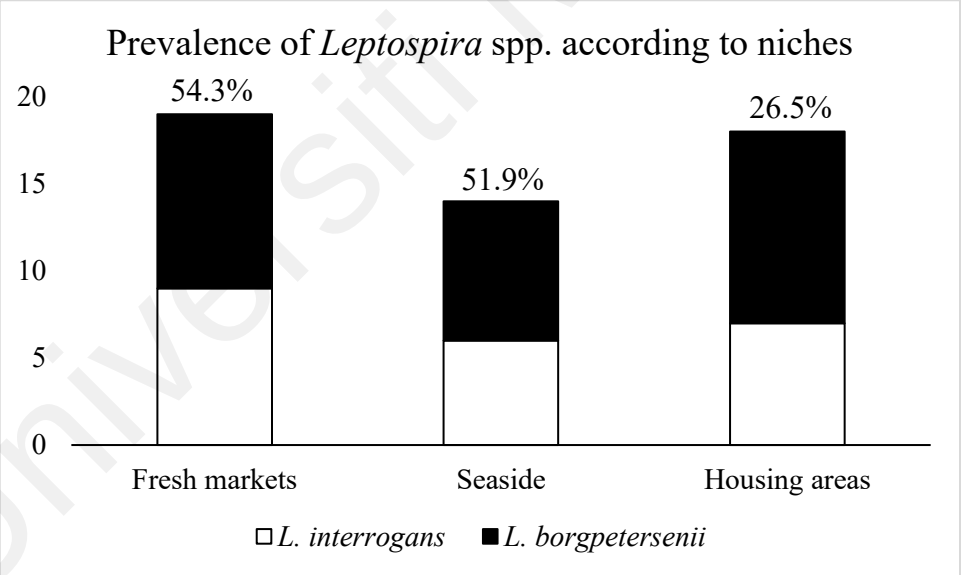


**Figure 4.22:** Prevalence of *Leptospira* spp. according to niches.

Furthermore, this study portrays the occurrence of leptospiral DNA in both kidney and liver tissues of *R. rattus*. Overall, from each kidney and liver tissues of the 130 rats utilised in the screening of infection, 11 (8.5%) liver tissues were colonised by *Leptospira* spp., in which three individuals were infected with *L. interrogans* and another eight individuals are in agreements with *L. borgpetersenii*. A total of 40 (30.8%) individuals of *R. rattus*

were proved leptospirosis positive after the kidney tissues were screened. Nineteen *L. interrogans* and 21 *L. borgpetersenii* species colonised the kidney of the rats (Figure 4.23). Surprisingly, none of the individuals got infected in both kidney and liver tissues simultaneously. The leptospiral DNA presented in either one of the two screened organs of an individual at each time point.

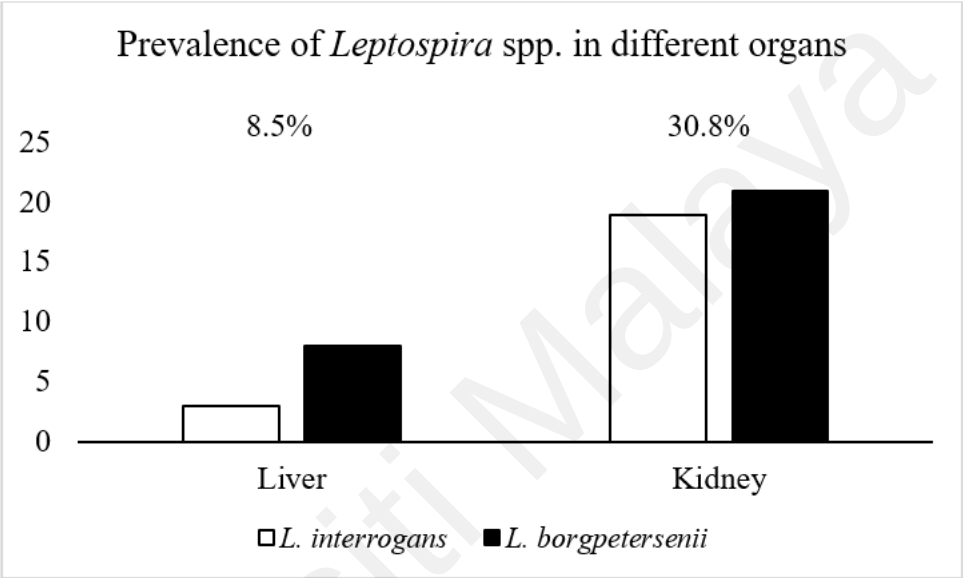


**Figure 4.23:** Prevalence of *Leptospira* spp. in different organs.

Analysis of all nucleotide sequences of *L. interrogans* and *L. borgpetersenii* used in this study shows a total of 412 characters in the 16S rRNA sequence with 400 (97.1%) conserved sites and 12 (2.9%) variable sites. For *L. interrogans* and *L. borgpetersenii* sequences associated to LipL32 gene, a total of 240 characters were found with 225 (93.8%) conserved sites and 15 (6.2%) variable sites. The findings of 2.7% and 6.2%, informative sites in 16S rRNA and LipL32 genes suggested that both genes are very conserved, respectively. The transition-transversion rate analysis was also conducted which shows that the 16S rRNA and LipL32 sequences undergoes higher transition rates as compared to transversion rates, resulting in the Ti/Tv ratio of 3.1 and 2.8, respectively. The nucleotide composition analysis of 16S rRNA nucleotide sequences shows that

Guanine (G) has the highest composition in the sequences with an average of 32.4%, followed by Adenine (A) with an average of 27.3%, Thymine (T) with an average of 20.2%, and Cytosine (C) with an average of 20.1%. For LipL32 nucleotide sequences, Adenine (A) demonstrated the highest average of 29.5%, followed by Thymine (T) with 24.1%, Guanine (G) with 24.0%, and the lowest average of 22.4% by Cytosine (C).

For species identification using phylogenetic analysis, the Kimura 2-Parameter model was used in NJ tree reconstructions tested with a bootstrap value of 1000. From 16S rRNA gene sequences, three well-defined clades of *Leptospira* spp. were successfully resolved (Figure 4.24), referred to the clades of pathogenic, intermediate, and saprophytic species. The leptospires DNA obtained from kidney and liver tissues of *R. rattus* were clustered into the pathogenic clades of *L. interrogans* and *L. borgpetersenii*. The positive control samples were also segregated accordingly. Essentially, the tree revealed monophyly of the genus *Leptospira* with respect to the outgroup, *Leptonema illini*. Phylogenetic comparisons were initiated between the two *Leptospira* spp. using LipL32 gene. Figure 4.25 indicates that the similar tree topology was derived from all the sequences, in which two separated clades of *L. interrogans* and *L. borgpetersenii* were formed.

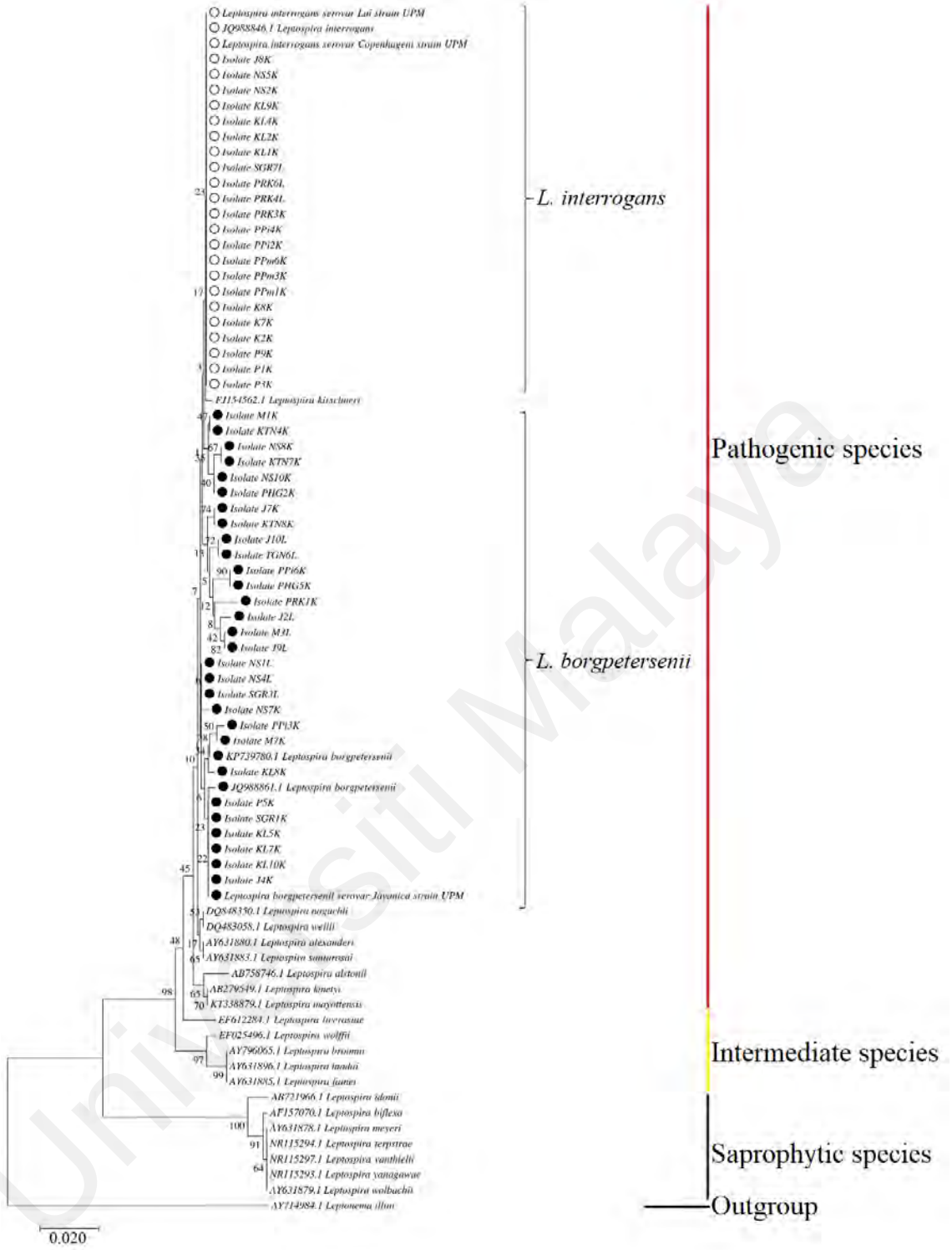


**Figure 4.24:** The neighbor-joining (NJ) phylogenetic tree of *Leptospira* spp. estimated using the Kimura 2-parameter algorithm and 1000 bootstrap replications, based on 16S rRNA gene. Blank circles represent *L. interrogans* while filled circles represent *L. borgpetersenii*. The optimal tree with the sum of branch length = 0.3224 is shown and bootstrap values are indicated on the branches. The bar represents 0.02 nucleotide substitutions per alignment position.

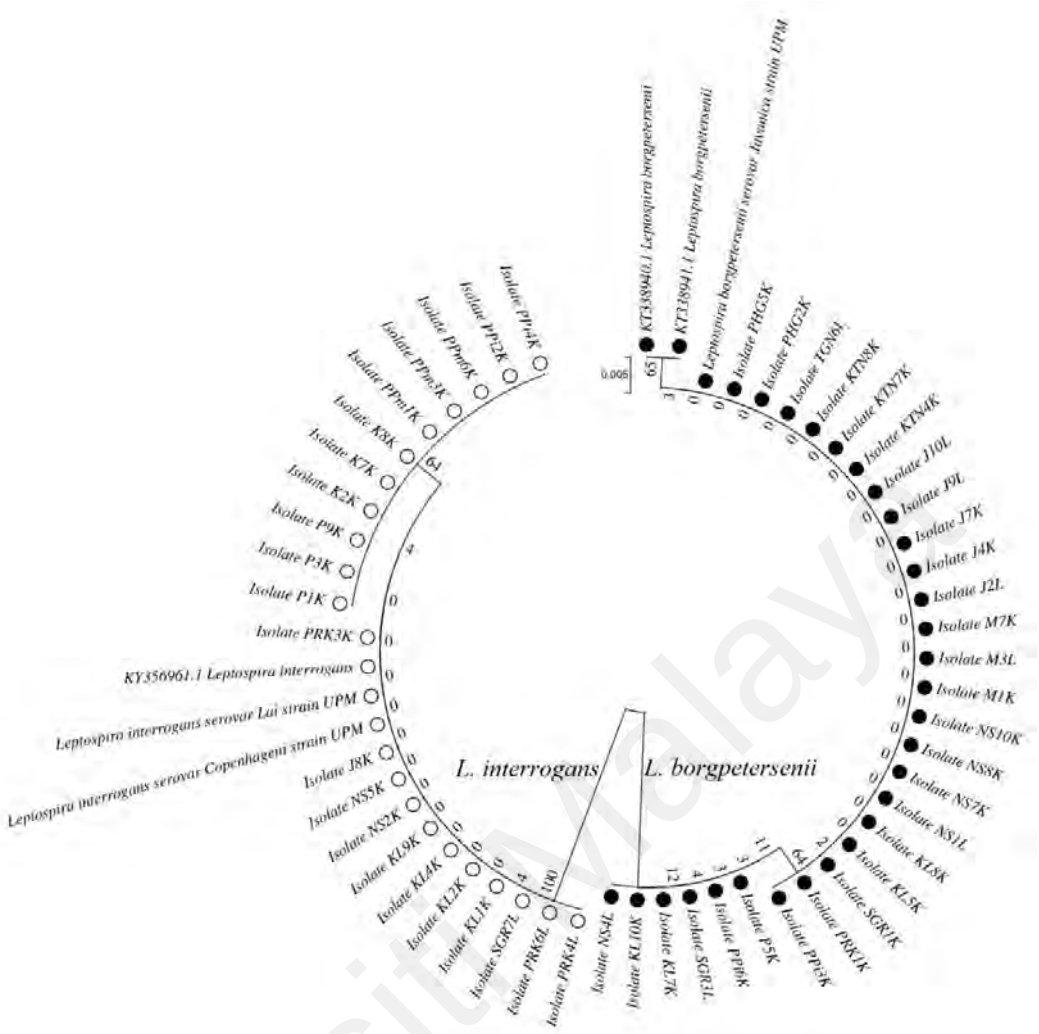


**Figure 4.25:** The neighbor-joining (NJ) phylogenetic tree of *L. interrogans* and *L. borgpetersenii* estimated using the Kimura 2-parameter algorithm and 1000 bootstrap replications, based on LipL32 gene. Blank circles represent *L. interrogans* while filled circles represent *L. borgpetersenii*. The optimal tree with the sum of branch length = 0.0652 is shown and bootstrap values are indicated on the branches. The bar represents 0.005 nucleotide substitutions per alignment position.

The average percentages of pairwise genetic distances among *Leptospira* spp. and outgroup based on 16S rRNA gene calculated using Kimura 2-Parameter algorithm model is tabulated in Table 4.24. From the analysis, pathogenic leptospires of *L. interrogans* and *L. borgpetersenii* observed in this study demonstrated an interspecific genetic distance of 0.6% between them. Interestingly, the genetic distances between pathogenic, intermediate, and saprophytic groups of leptospires revealed the most prominent results. The relationships of the pathogenic group with intermediate and saprophytic groups are in the range of 1.7%-3.7% and 8.8%-10.2%, respectively. Species of genus *Leptospira*

**Table 4.24:** Average percentages of pairwise genetic distances among *Leptospira* spp. and *Leptonema illini* (Outgroup), based on the 16S rRNA gene.

| Taxon | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) *L. interrogans* | - | | | | | | | | | | | | | | | | | | | | | | |
| (2) *L. kirschneri* | 0.2 | - | | | | | | | | | | | | | | | | | | | | | |
| (3) *L. noguchii* | 0.5 | 0.7 | - | | | | | | | | | | | | | | | | | | | | |
| (4) *L. weilii* | 0.5 | 0.7 | 0.0 | - | | | | | | | | | | | | | | | | | | | |
| (5) *L. alexanderi* | 0.5 | 0.7 | 0.2 | 0.2 | - | | | | | | | | | | | | | | | | | | |
| (6) *L. santarosai* | 0.5 | 0.7 | 0.2 | 0.2 | 0.0 | - | | | | | | | | | | | | | | | | | |
| (7) *L. borgpetersenii* | 0.6 | 0.8 | 0.7 | 0.7 | 0.7 | 0.7 | - | | | | | | | | | | | | | | | | |
| (8) *L. kmetyi* | 1.0 | 1.2 | 0.5 | 0.5 | 0.7 | 0.7 | 1.2 | - | | | | | | | | | | | | | | | |
| (9) *L. mayottensis* | 1.0 | 1.2 | 0.5 | 0.5 | 0.7 | 0.7 | 1.2 | 0.0 | - | | | | | | | | | | | | | | |
| (10) *L. alstonii* | 1.2 | 1.4 | 0.7 | 0.7 | 0.9 | 0.9 | 1.4 | 0.2 | 0.2 | - | | | | | | | | | | | | | |
| (11) *L. licerasiae* | 1.7 | 1.7 | 1.7 | 1.7 | 2.0 | 2.0 | 2.3 | 2.0 | 2.0 | 2.5 | - | | | | | | | | | | | | |
| (12) *L. wolffii* | 2.2 | 2.2 | 2.7 | 2.7 | 2.7 | 2.7 | 2.8 | 2.7 | 2.7 | 3.5 | 2.5 | - | | | | | | | | | | | |
| (13) *L. broomii* | 2.7 | 2.5 | 3.0 | 3.0 | 3.0 | 3.0 | 3.1 | 3.0 | 3.0 | 3.7 | 3.2 | 1.0 | - | | | | | | | | | | |
| (14) *L. inadai* | 2.7 | 2.5 | 3.0 | 3.0 | 3.0 | 3.0 | 3.1 | 3.0 | 3.0 | 3.7 | 3.2 | 1.0 | 0.0 | - | | | | | | | | | |
| (15) *L. fainei* | 2.7 | 2.5 | 3.0 | 3.0 | 3.0 | 3.0 | 3.1 | 3.0 | 3.0 | 3.7 | 3.2 | 1.0 | 0.0 | 0.0 | - | | | | | | | | |
| (16) *L. idonii* | 9.1 | 8.8 | 9.6 | 9.6 | 9.6 | 9.6 | 9.3 | 9.9 | 9.9 | 10.2 | 9.6 | 9.9 | 9.4 | 9.4 | 9.4 | - | | | | | | | |
| (17) *L. biflexa* | 9.1 | 8.8 | 9.6 | 9.6 | 9.6 | 9.6 | 9.2 | 9.9 | 9.9 | 9.6 | 9.9 | 9.9 | 9.9 | 9.9 | 9.9 | 1.5 | - | | | | | | |
| (18) *L. meyeri* | 9.1 | 8.8 | 9.6 | 9.6 | 9.6 | 9.6 | 9.2 | 9.9 | 9.9 | 9.6 | 9.6 | 9.9 | 9.9 | 9.9 | 9.9 | 1.2 | 0.2 | - | | | | | |
| (19) *L. terpstrae* | 9.1 | 8.8 | 9.6 | 9.6 | 9.6 | 9.6 | 9.2 | 9.9 | 9.9 | 9.6 | 9.6 | 9.9 | 9.9 | 9.9 | 9.9 | 1.2 | 0.2 | 0.0 | - | | | | |
| (20) *L. vanthielii* | 9.1 | 8.8 | 9.6 | 9.6 | 9.6 | 9.6 | 9.2 | 9.9 | 9.9 | 9.6 | 9.6 | 9.9 | 9.9 | 9.9 | 9.9 | 1.2 | 0.2 | 0.0 | 0.0 | - | | | |
| (21) *L. yanagawae* | 9.1 | 8.8 | 9.6 | 9.6 | 9.6 | 9.6 | 9.2 | 9.9 | 9.9 | 9.6 | 9.6 | 9.9 | 9.9 | 9.9 | 9.9 | 1.2 | 0.2 | 0.0 | 0.0 | 0.0 | - | | |
| (22) *L. wolbachii* | 9.1 | 8.8 | 9.6 | 9.6 | 9.6 | 9.6 | 9.2 | 9.9 | 9.9 | 9.6 | 9.6 | 9.9 | 9.9 | 9.9 | 9.9 | 1.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | - | |
| (23) *Leptonema illini* | 15.8 | 15.8 | 15.5 | 15.5 | 15.4 | 15.4 | 15.8 | 15.5 | 15.5 | 15.2 | 15.8 | 15.8 | 16.1 | 16.1 | 16.1 | 18.0 | 17.4 | 17.4 | 17.4 | 17.4 | 17.4 | 17.4 | - |

Pathogenic species

Intermediate species

Saprophytic species

exhibited the highest value of a pairwise distance with *Leptonema illini* between 15.2%-18.0%, which were highly expected as they were originated from two different genera. From the calculation of pairwise distances, it can be concluded that the genetic distances of 16S rRNA sequences can explain the relationship between the species and genus level. Additionally, the pairwise genetic distance between *L. interrogans* and *L. borgpetersenii* corresponded to LipL32 gene shows greater nucleotide divergence with the value of 5.8%. The result signified that 16S rRNA gene had a lower rate of base substitution and was more conserved as compared to LipL32 gene.

Single-nucleotide polymorphisms (SNP) analysis conducted on *L. interrogans* and *L. borgpetersenii* sequences based on the 16S rRNA and LipL32 genes revealed 12 and 15 segregating sites, respectively (Appendix G). Throughout the analysed sequences of the two species, 18 haplotypes were defined for 16S rRNA gene and five haplotypes were derived from LipL32 gene. Minimum-spanning network (MSN) was generated with the haplotype data obtained to illustrate the relationships of the two *Leptospira* spp. For 16S rRNA gene, *L. interrogans* solely colonised in a single haplotype (H_1) and the remaining 17 haplotypes were conquered by *L. borgpetersenii* (H_2-H_18) (see Figure 4.26). Among the five haplotypes detected using LipL32 gene, two haplotypes were observed for *L. interrogans* (H_1-H_2), as well as three from *L. borgpetersenii* (H_3-H_5) (see Figure 4.27). The network analysis of both genes revealed that the haplotypes are unique to each species, as all of them lacked a common haplotype.

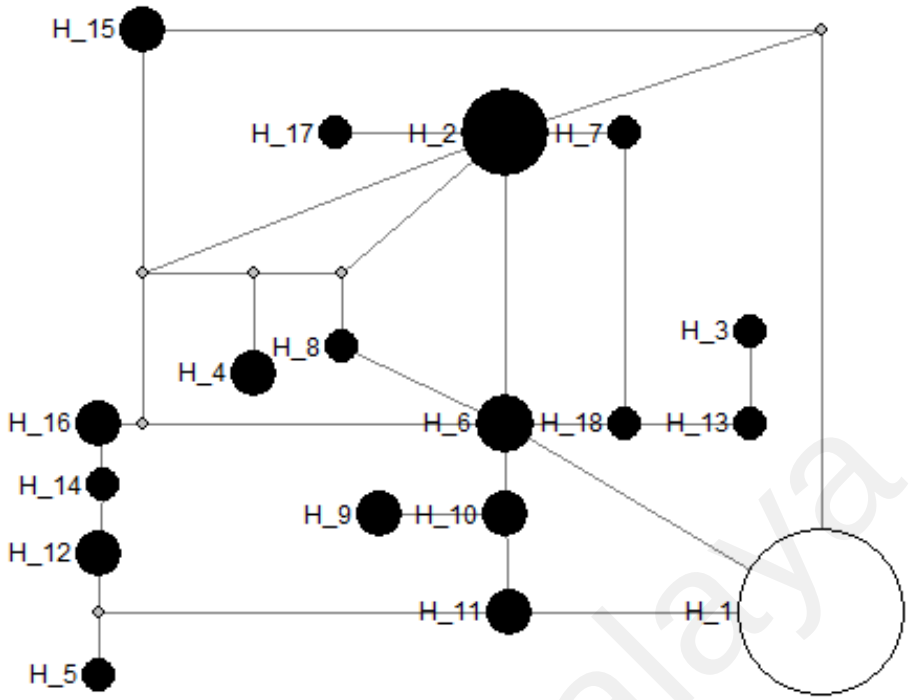


**Figure 4.26:** Minimum-spanning network of *L. interrogans* and *L. borgpetersenii* haplogroup defining 18 haplotypes, based on 16S rRNA gene. Haplotype 1 (blank circle) – *L. interrogans*; Haplotype 2-18 (filled circles) – *L. borgpetersenii*. The circle size of *Leptospira* spp. haplotypes is proportional to the frequency of the haplotype.

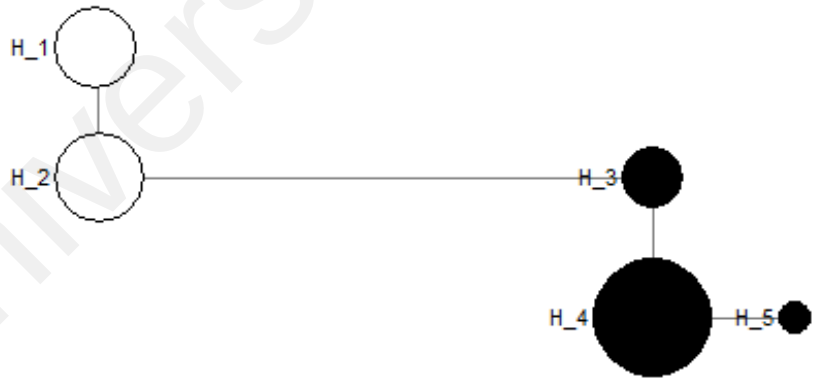


**Figure 4.27:** Minimum-spanning network of *L. interrogans* and *L. borgpetersenii* haplogroup defining five haplotypes, based on LipL32 gene. Haplotype 1 and 2 (blank circles) – *L. interrogans*; Haplotype 3-5 (filled circles) – *L. borgpetersenii*. The circle size of *Leptospira* spp. haplotypes is proportional to the frequency of the haplotype.

The estimation of genetic diversity of *L. interrogans* and *L. borgpetersenii*, based on 16S rRNA and LipL32 genes was tabulated in Table 4.25. Both genes revealed higher nucleotide and haplotypes diversities in *L. borgpetersenii* (16S rRNA: $\pi$ = 0.00768, *Hd*

= 0.93750; LipL32: $\pi$ = 0.00214, *Hd* = 0.51333) in contrast to *L. interrogans* (16S rRNA: $\pi$ = 0.00000, *Hd* = 0.00000; LipL32: $\pi$ = 0.00181, *Hd* = 0.41129). Besides, the statistics of genetic variability between the two *Leptospira* spp. were analysed as shown in Table 4.26. Based on the 16S rRNA gene, the observed nucleotide diversity ($\pi$) is very low, which is 0.5%. The net nucleotide divergence (*Da*) between the species was also low with the value of 0.2%. However, $\pi$ and *Da* values calculated between the two species using LipL32 gene were found to exhibit 2.9% and 5.4%, respectively.

**Table 4.25:** Estimates of genetic diversity of *L. interrogans* and *L. borgpetersenii*, based on 16S rRNA and LipL32 genes.

| **Variables** | **16S rRNA** | | **LipL32** | |
|---|---|---|---|---|
| | ***L. interrogans*** | ***L. borgpetersenii*** | ***L. interrogans*** | ***L. borgpetersenii*** |
| No. of individuals (*N*) | 25 | 32 | 25 | 32 |
| No. of haplotypes | 1 | 17 | 2 | 3 |
| Haplotype diversity (*Hd*) | 0.00000 | 0.93750 | 0.41129 | 0.51333 |
| No. of polymorphic sites | 0 | 12 | 1 | 2 |
| Nucleotide diversity ($\pi$) | 0.00000 | 0.00768 | 0.00181 | 0.00214 |

**Table 4.26:** Summary statistics of genetic variability among the two leptospiral species (*L. interrogans* and *L. borgpetersenii*), based on 16S rRNA and LipL32 genes.

| **Variables** | **16S rRNA** | **LipL32** |
|---|---|---|
| No. of individuals (*N*) | 57 | 57 |
| Nucleotide diversity ($\pi$) | 0.005 | 0.029 |
| Net nucleotide divergence (*Da*) | 0.002 | 0.054 |
| Haplotype diversity (*Hd*) | 0.793 | 0.726 |
| Average number of nucleotide differences (*k*) | 2.262 | 6.906 |
| Nucleotide subdivision (*Nst*) | 0.385 | 0.966 |
| Estimate of population subdivision (*Fst*) | 0.386 | 0.964 |
| Number of migrants per generation (*Nm*) | 0.40 | 0.01 |

Genetic differentiation (*Fst*, *Nst*, and *Nm*) values were studied to further clarify the relationships between the two species (refer Table 4.26). Through 16S rRNA gene, the differentiation of both species portrays the *Fst* and *Nst* values of 0.386 and 0.385,

respectively. Based on LipL32 gene, the division of *L. interrogans* from *L. borgpetersenii* demonstrated a higher *Fst* value of 0.964. *Nst* analysis results were totally parallel with *Fst*, with *L. interrogans*-*L. borgpetersenii* showing higher *Nst* value of 0.966. It is known that the *Fst* value higher than 0.25 secured a high genetic differentiation due to genetic structure, while *Nst* value of 1 shows a complete population division. The analysis of gene flow (*Nm*) between the two species discovered the number of migrants in 0.4 in 16S rRNA and 0.01 for LipL32 genes, proposing that the gene flow between these two species was cut-off over time. Apparently, the $Nm < 1$ indicates the genetic divergence over time whilst the $Nm > 1$ stipulates the retaining of the gene flow. Hence, the *Nm* value will be contrary to both *Fst* and *Nst*. Presently, *Nm* results verified both *Fst* and *Nst*. The AMOVA analysis utilising both genes disclosed significant genetic divergence between the two species (Table 4.27 and Table 4.28). For 16S rRNA gene, the disparity assigned between species was 35.51% while variation within species attributed to 64.49% of the total variance ($F_{ST} = 0.35505$; $P < 0.001$). In furtherance, LipL32 revealed the divergence existed between species was 96.47% while variation within species ascribed to 3.53% of the total variance ($F_{ST} = 0.96473$; $P < 0.001$).

**Table 4.27:** Results of analysis of molecular variance (AMOVA) testing genetic structure of *L. interrogans* and *L. borgpetersenii*, based on 16S rRNA gene.

| Source of variation | Variance | Percentage of variation | *F*-statistic | *P*-values |
|---|---|---|---|---|
| Among species | 0.48795 | 35.51 | $F_{ST} = 0.35505$ | $P = 0.00000$ |
| Within species | 0.88636 | 64.49 | | |

**Table 4.28:** Results of analysis of molecular variance (AMOVA) testing genetic structure of *L. interrogans* and *L. borgpetersenii*, based on LipL32 gene.

| Source of variation | Variance | Percentage of variation | *F*-statistic | *P*-values |
|---|---|---|---|---|
| Among species | 6.42047 | 96.47 | $F_{ST} = 0.96473$ | $P = 0.00000$ |
| Within species | 0.23473 | 3.53 | | |

## 4.5 Transcriptome Study of *Rattus rattus* Infected with *Leptospira* spp.

For the transcriptome project, the main intention was to observe the gene expression level on both kidney and liver of *Rattus rattus* infected with leptospirosis. Thus, four samples with a good RIN number were sequenced in total, using BGISEQ-500 platform (Beijing Genome Institute, Shenzhen, China). The four samples included healthy kidney, infected kidney, healthy liver, and infected liver. Figure 4.28 shows the result of RNA quality of the four samples which was inspected using Agilent 2100 Bioanalyzer (Agilent Technologies, USA).

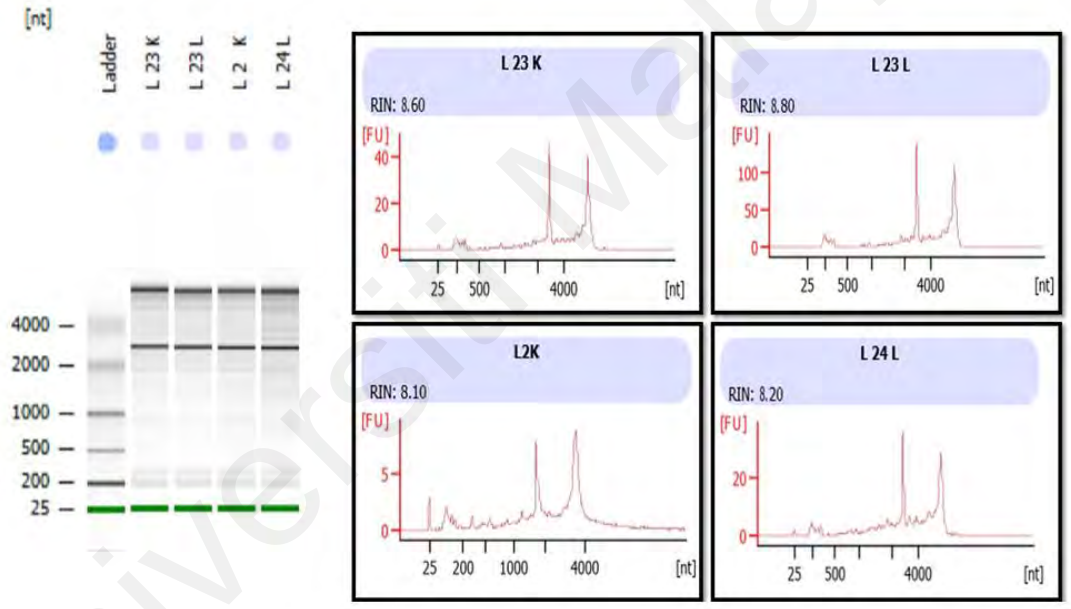


**Figure 4.28:** Result of RNA quality of *R. rattus* samples inspected using Agilent 2100 Bioanalyzer (Agilent Technologies, USA). (L23K: Healthy kidney; L23L: Healthy liver; L2K: Infected kidney; L24L: Infected liver).

### 4.5.1 Sequencing reads filtering

The sequencing reads which containing low-quality, adaptor-polluted, and high content of unknown base (N) reads, were processed and removed before downstream analyses. The original data performance is shown in Figure 4.29. The clean reads quality metrics are shown as Table 4.29. For the clean reads ratio, infected kidney shows the

higher percentage of 95.93%. The lowest clean reads ratio was shown from a healthy kidney, at 92.11%.

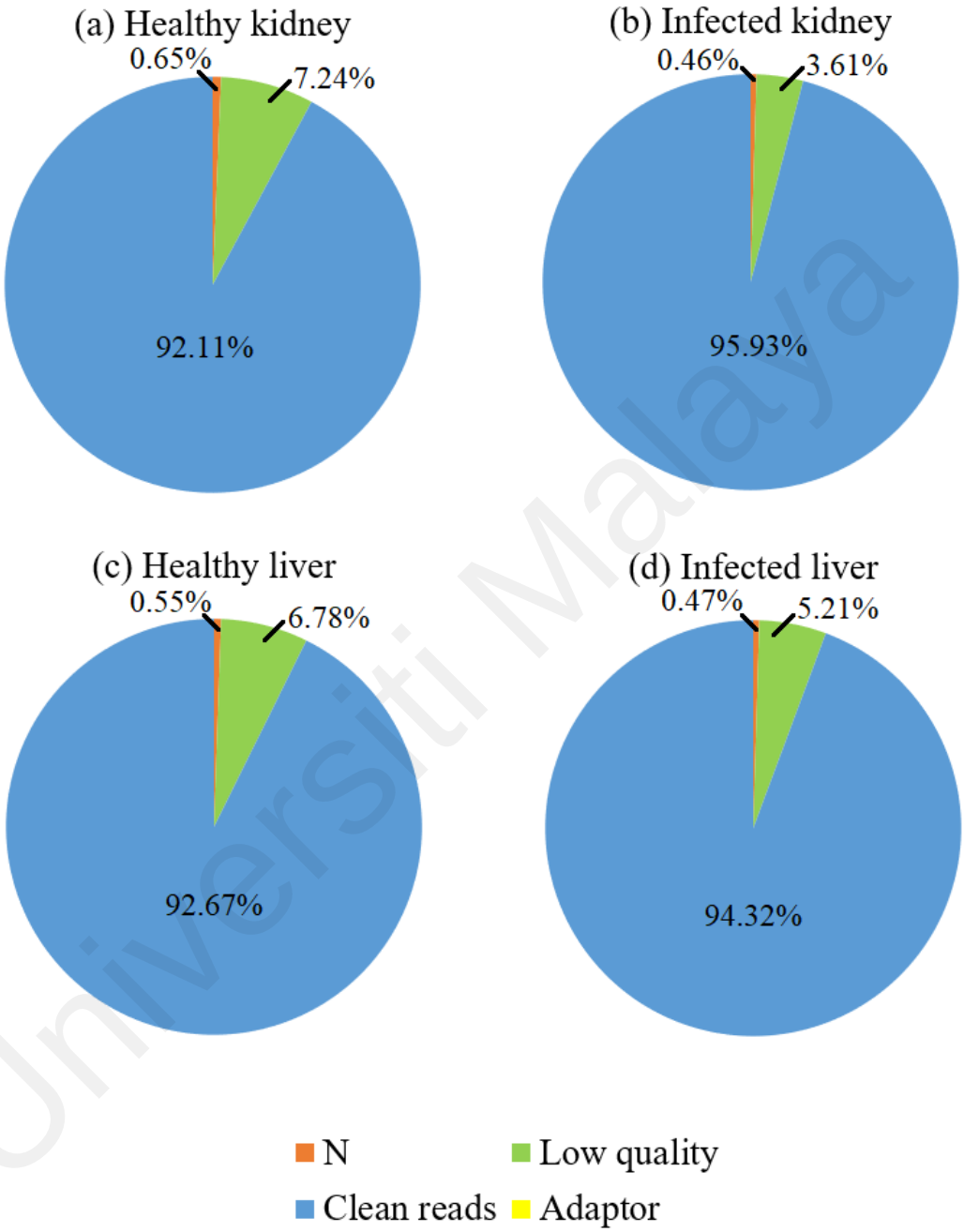


**Figure 4.29:** Raw data filter composition chart. (a) healthy kidney, (b) infected kidney, (c) healthy liver, and (d) infected liver. N: The total amount of reads which contain more than 5% unknown N base; the N reads ratio; Adaptor: The total amount of reads which contain adaptors; the adaptor ratio; Low quality: More than 20% of bases in the total read have quality score lower than 15; low quality reads ratio; Clean reads: Reads filtered with N reads, reads have adaptors and low quality reads; clean reads ratio.

**Table 4.29:** Clean reads quality metrics.

| Sample | Total Raw Reads (Mb) | Total Clean Reads (Mb) | Total Clean Bases (Gb) | Clean Reads Q20 (%) | Clean Reads Q30 (%) | Clean Reads Ratio (%) |
|---|---|---|---|---|---|---|
| Healthy kidney | 84.67 | 77.98 | 7.80 | 97.39 | 90.94 | 92.11 |
| Healthy liver | 84.67 | 78.46 | 7.85 | 97.46 | 91.04 | 92.67 |
| Infected kidney | 83.20 | 79.82 | 7.98 | 98.20 | 91.00 | 95.93 |
| Infected liver | 84.67 | 79.86 | 7.99 | 97.59 | 91.42 | 94.32 |

Total Raw Reads (Mb): The reads amount before filtering, Unit: Mb; Total Clean Reads (Mb): The reads amount after filtering, Unit: Mb; Total Clean Bases (Gb): The total base amount after filtering, Unit: Gb; Clean Reads Q20 (%): The Q20 value for the clean reads; Clean Reads Q30 (%): The Q30 value for the clean reads; Clean Reads Ratio (%): The ratio of the amount of clean reads.

### 4.5.2 Genome mapping

The filtered clean reads were mapped to a reference genome of *Rattus norvegicus* using HISAT. On average, 63.19% reads were mapped. The uniformity of the mapping result for each sample proposes that the samples are comparable. The mapping details are shown as Table 4.30. The highest and lowest uniquely mapping ratio of 39.82% and 33.44% were shown by infected liver and infected kidney, respectively.

**Table 4.30:** Summary of genome mapping.

| Sample | Total Clean Reads | Total Mapping Ratio | Uniquely Mapping Ratio |
|---|---|---|---|
| Healthy kidney | 77984062 | 57.45% | 34.28% |
| Healthy liver | 78462914 | 70.19% | 39.53% |
| Infected kidney | 79819242 | 55.42% | 33.44% |
| Infected liver | 79855252 | 69.69% | 39.82% |

Total Clean Reads: The amount of clean reads; Total Mapping Ratio: The percentage of mapped reads; Uniquely Mapping Ratio: The percentage of reads that map to only one location of reference.

### 4.5.3 Novel transcripts prediction

StringTie was used to reconstruct transcripts, after genome mapping. Using genome annotation information, the novel transcripts were identified by using Cuffcompare and the coding ability of those new transcripts were predicted using CPC. Altogether, 18713 genes were identified in which 14477 of them are known genes and 4455 of them are novel genes. In total, 20530 novel transcripts were identified, in which 10022 of them are previously unknown splicing event for known genes, 4455 of them are novel coding transcripts without any known features, and the remaining 6053 are long non-coding RNA. The detailed information is shown as Table 4.31.

**Table 4.31:** Summary of novel transcripts.

| Total Novel Transcript | Coding Transcript | Non-coding Transcript | Novel Isoform | Novel Gene |
|---|---|---|---|---|
| 20530 | 14477 | 6053 | 10022 | 4455 |

Total Novel Transcript: The amount of predicted novel transcripts; Coding Transcript: The amount of predicted coding transcripts; Non-coding Transcript: The amount of predicted non-coding transcripts; Novel Isoform: The amount of predicted coding transcripts that previously unknown splicing event for a known gene; Novel Gene: The amount of predicted coding transcripts that previously unknown.

### 4.5.4 SNP and INDEL detection

Genome Analysis Toolkit (GATK) was used to find SNP and INDEL variant for each sample, after genome mapping. The SNP summary is shown as Table 4.32 and Figure 4.30. The highest total of SNP variant was shown by an infected kidney and the lowest total of SNP variant was shown by a healthy liver.

**Table 4.32:** SNP variant type summary.

| Sample | A-G | C-T | Transi-tion | A-C | A-T | C-G | G-T | Trans-version | Total |
|---|---|---|---|---|---|---|---|---|---|
| Healthy kidney | 266464 | 264289 | 530753 | 54679 | 43145 | 53419 | 54383 | 205626 | 736379 |
| Healthy liver | 255658 | 254165 | 509823 | 52878 | 42502 | 50663 | 52807 | 198850 | 708673 |
| Infected kidney | 346569 | 344499 | 691068 | 72146 | 57935 | 70115 | 72378 | 272574 | 963642 |
| Infected liver | 280139 | 278355 | 558494 | 58357 | 46826 | 55863 | 58381 | 219427 | 777921 |

A-G: The amount of A-G variant type; C-T: The amount of C-T variant type; Transition: The amount of A-G and C-T variant type; A-C: The amount of A-C variant type; A-T: The amount of A-T variant type; C-G: The amount of C-G variant type; G-T: The amount of G-T variant type; Transversion: The amount of A-C, A-T, C-G, and G-T variant type; Total: The amount of all variant type.

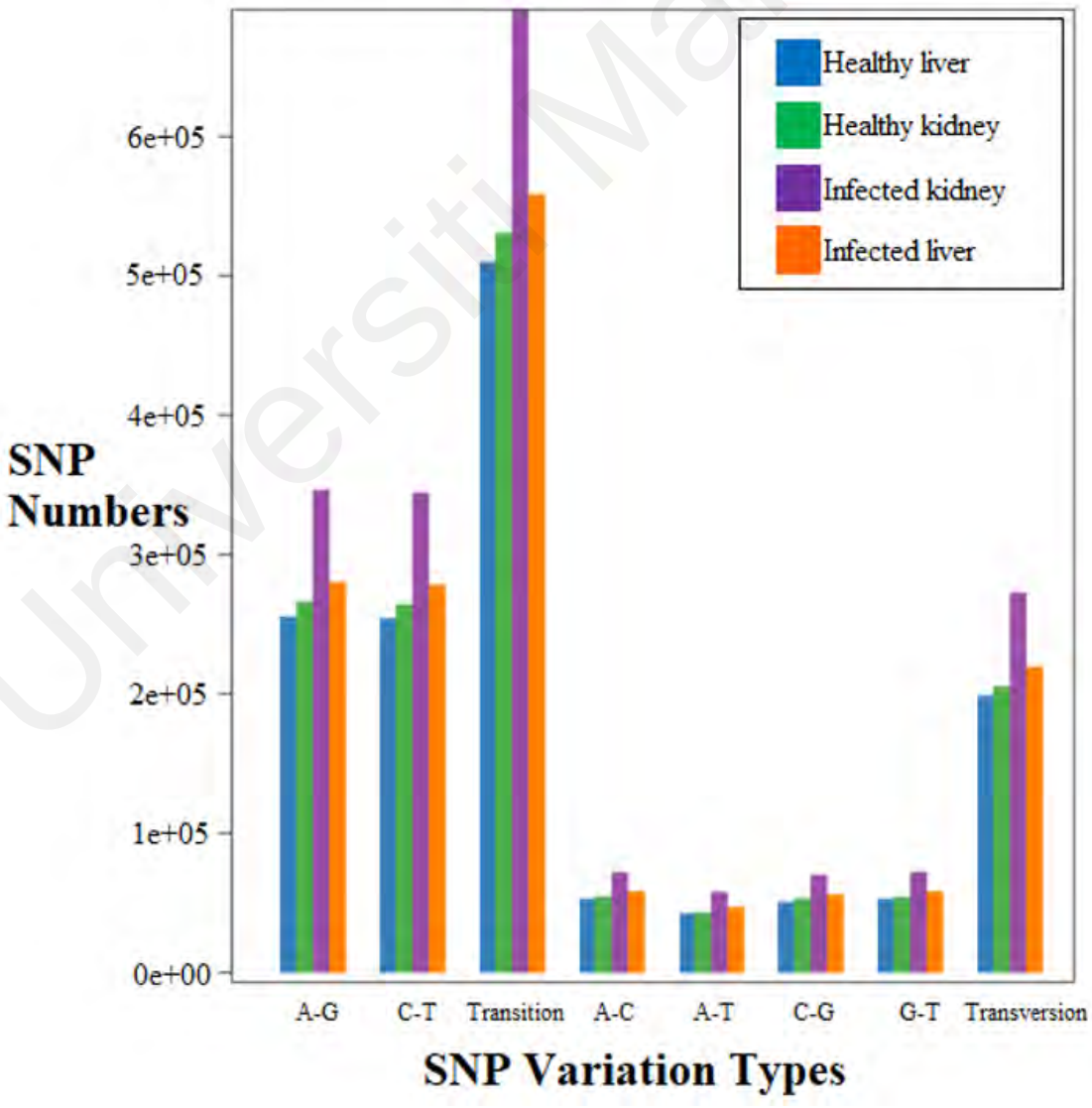


**Figure 4.30:** SNP variant type distribution.

Then, the location of SNP was identified, shown as Figure 4.31. For the healthy kidney and liver, the percentage of SNP were higher in the exon, while infected kidney and liver illustrated the higher percentage of SNP in the intron.

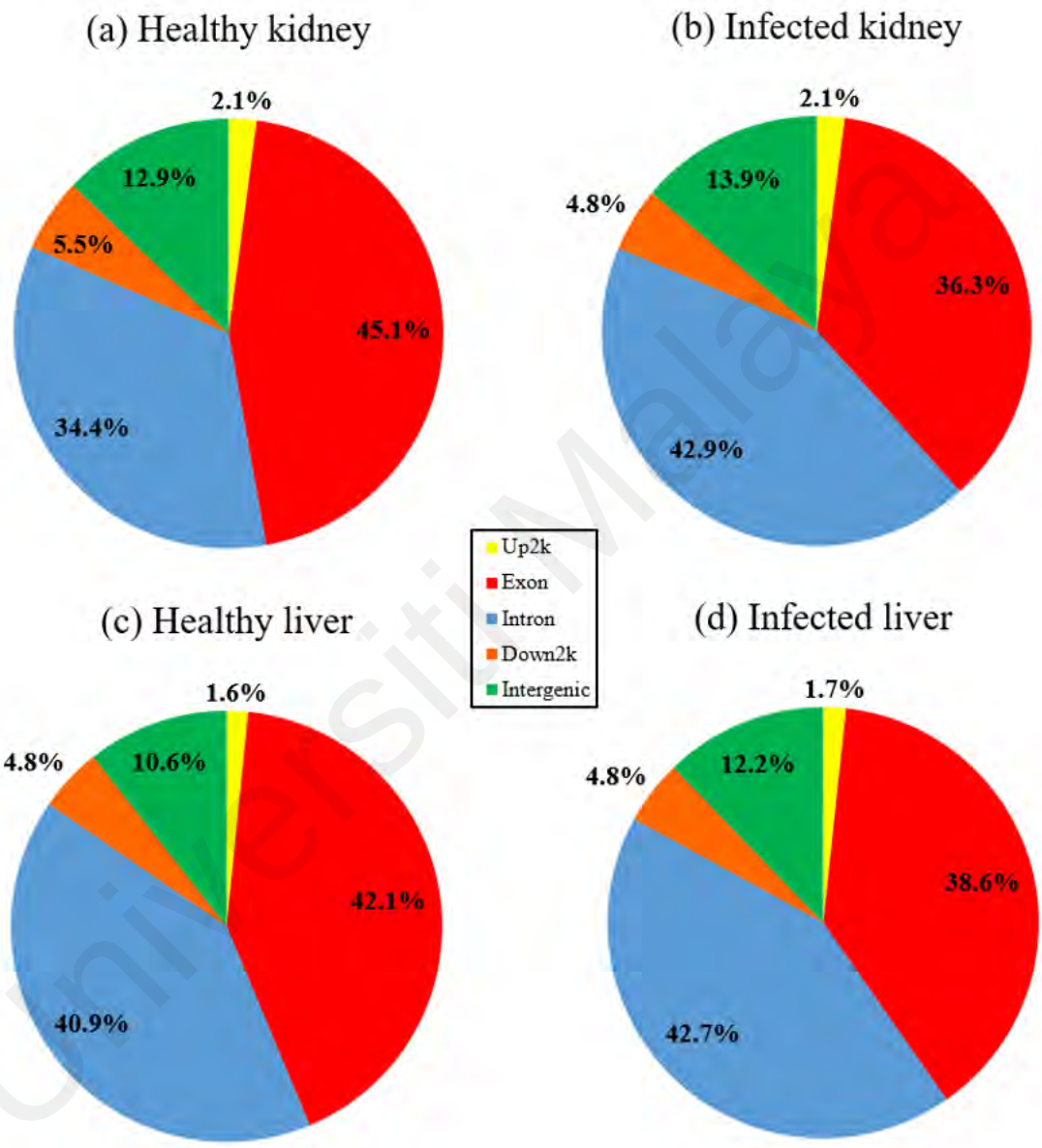


**Figure 4.31:** Distribution of SNP location. (a) healthy kidney, (b) infected kidney, (c) healthy liver, and (d) infected liver. (Up2k means upstream 2000 bp area of a gene. Down2k means downstream 2000 bp area of a gene).

The location of INDEL were shown as Figure 4.32. For a healthy kidney, the percentage of INDEL were higher in the exon, while infected kidney, healthy liver, and infected liver demonstrated the higher percentage of INDEL in the intron.

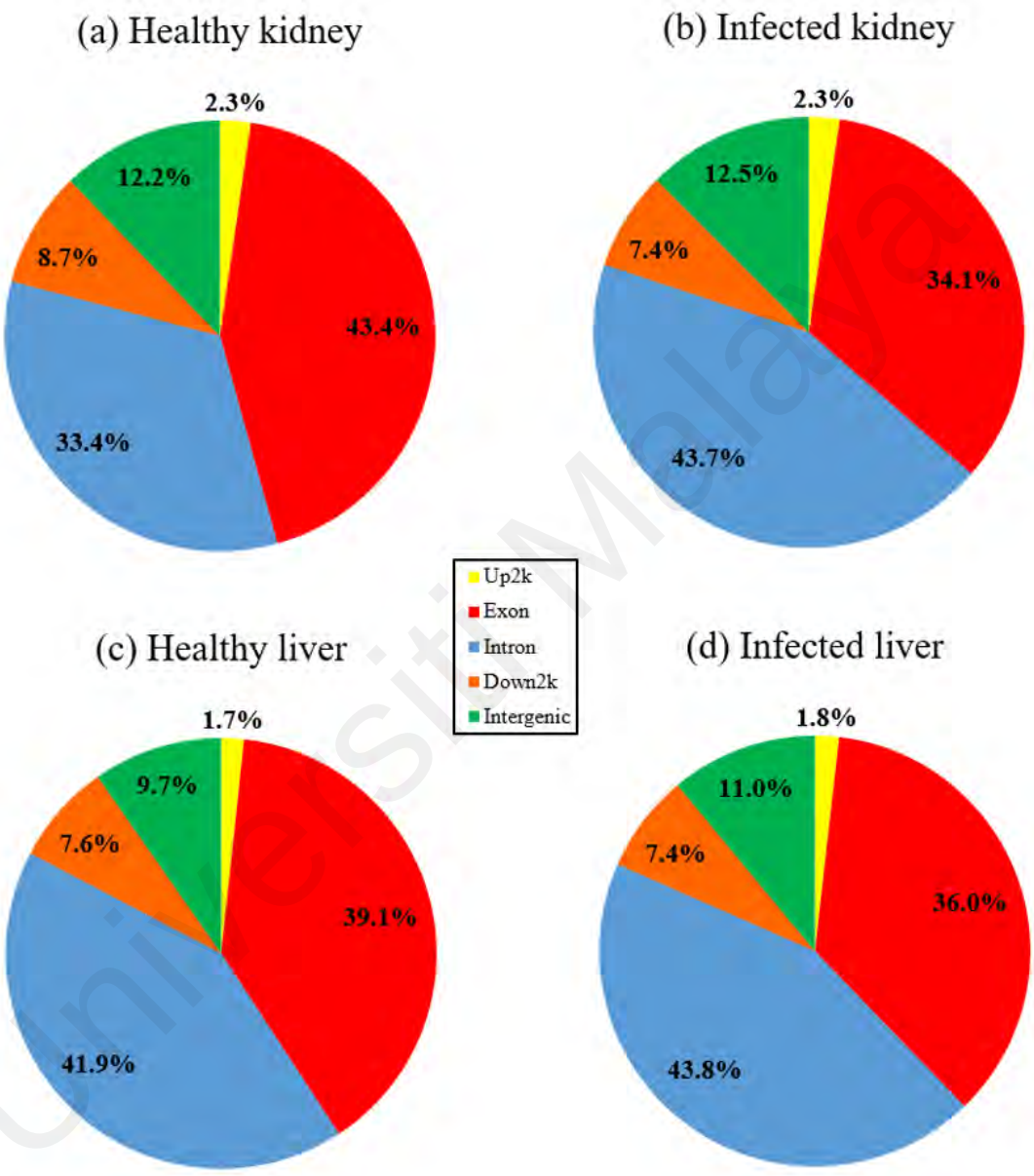


**Figure 4.32:** Distribution of INDEL location. (a) healthy kidney, (b) infected kidney, (c) healthy liver, and (d) infected liver. (Up2k means upstream 2000 bp area of a gene. Down2k means downstream 2000 bp area of a gene).

### 4.5.5 Differentially splicing gene detection

After genome mapping, rMATS were used to detect differentially splicing gene (DSG) between samples. DSGs are regulated by alternative splicing (AS), which allows the

production of a variety of different isoforms from one gene only. Changes in relative abundance of isoforms, regardless of the expression change, signify a splicing-related mechanism. Five types of AS events were discovered, including Skipped Exon (SE), Alternative 5' Splicing Site (A5SS), Alternative 3' Splicing Site (A3SS), Mutually exclusive exons (MXE), and Retained Intron (RI). The summary of gene splicing is shown in Figure 4.33. The highest numbers of alternative splicing events of all the five types were shown by an infected kidney and the lowest were represented by a healthy liver.

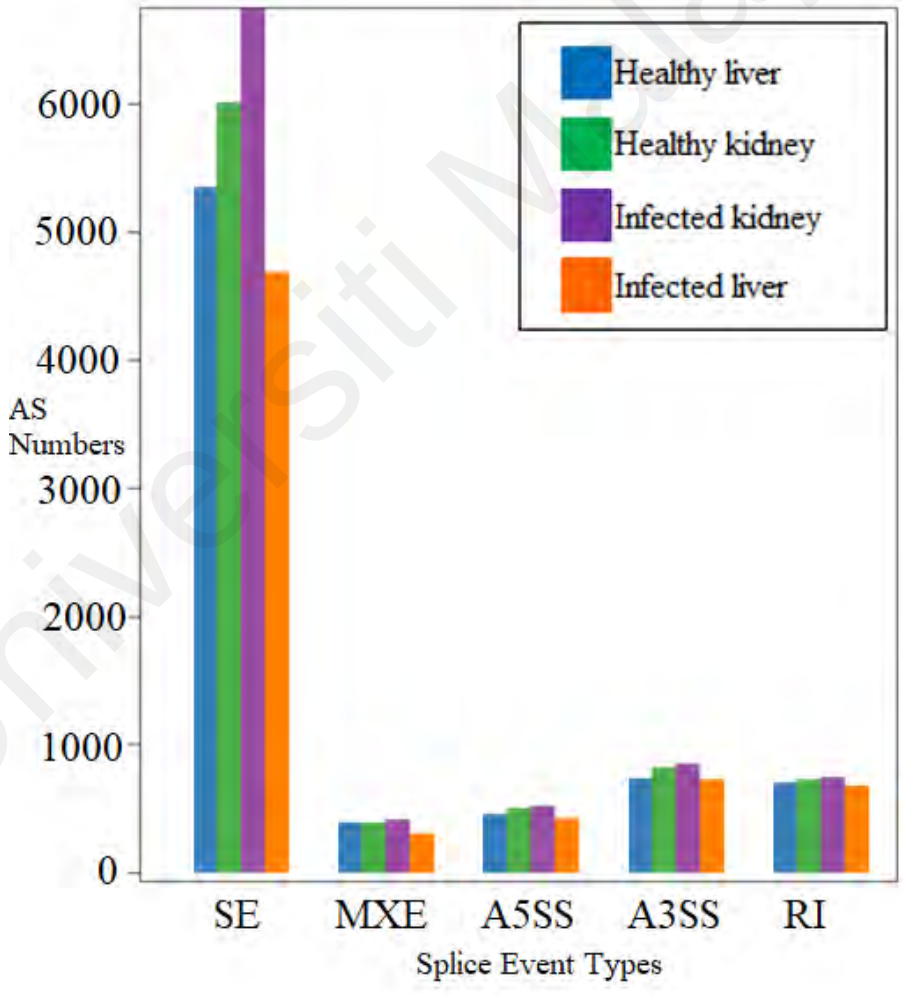


**Figure 4.33:** Statistic of splicing.

### 4.5.6 Gene expression analysis

#### 4.5.6.1 Gene mapping and expression

After novel transcript detection, the novel coding transcripts were merged with reference transcripts to get complete reference. The clean reads were mapped into it using Bowtie2, then the gene expression level was calculated for each sample with RSEM. The gene mapping ratio is shown as Table 4.33. The average gene mapping rate is 59.74%. The highest and lowest uniquely mapping ratio of 58.98% and 41.49% were shown by infected liver and infected kidney, respectively.

**Table 4.33:** Summary of gene mapping ratio.

| Sample | Total Clean Reads | Total Mapping Ratio | Uniquely Mapping Ratio |
|---|---|---|---|
| Healthy kidney | 77984062 | 53.63 | 45.91 |
| Healthy liver | 78462914 | 68.72 | 58.05 |
| Infected kidney | 79819242 | 48.62 | 41.49 |
| Infected liver | 79855252 | 67.97 | 58.98 |

Total Clean Reads: The amount of Clean reads; Total Mapping Ratio: The percentage of mapped reads (%); Uniquely Mapping Ratio: The percentage of uniquely mapped reads (%).

The number of genes and transcripts of each sample is shown as Table 4.34. The highest and lowest total gene number, known gene number, novel gene number, total transcript number, known transcript number, and novel transcript number were signified by an infected kidney and a healthy liver, respectively.

**Table 4.34:** Genes and transcripts statistics.

| Sample | Total Gene Number | Known Gene Number | Novel Gene Number | Total Transcript Number | Known Transcript Number | Novel Transcript Number |
|---|---|---|---|---|---|---|
| Healthy kidney | 17260 | 13442 | 3818 | 25581 | 13135 | 12446 |
| Healthy liver | 16446 | 12792 | 3654 | 24247 | 12346 | 11901 |
| Infected kidney | 17570 | 13745 | 3825 | 26029 | 13449 | 12580 |
| Infected liver | 16535 | 12863 | 3672 | 24274 | 12371 | 11903 |

Total Gene Number: The amount of all genes; Known Gene Number: The amount of known genes; Novel Gene Number: The amount of novel genes; Total Transcript Number: The amount of all transcripts; Known Transcript Number: The amount of known transcripts; Novel Transcript Number: The amount of novel transcripts.

### 4.5.6.2 Correlation between samples

Pearson correlation between all samples was calculated, shown as Figure 4.34. The correlation between healthy kidney and infected kidney was at 0.904. The correlation between healthy liver and infected liver was at 0.885.

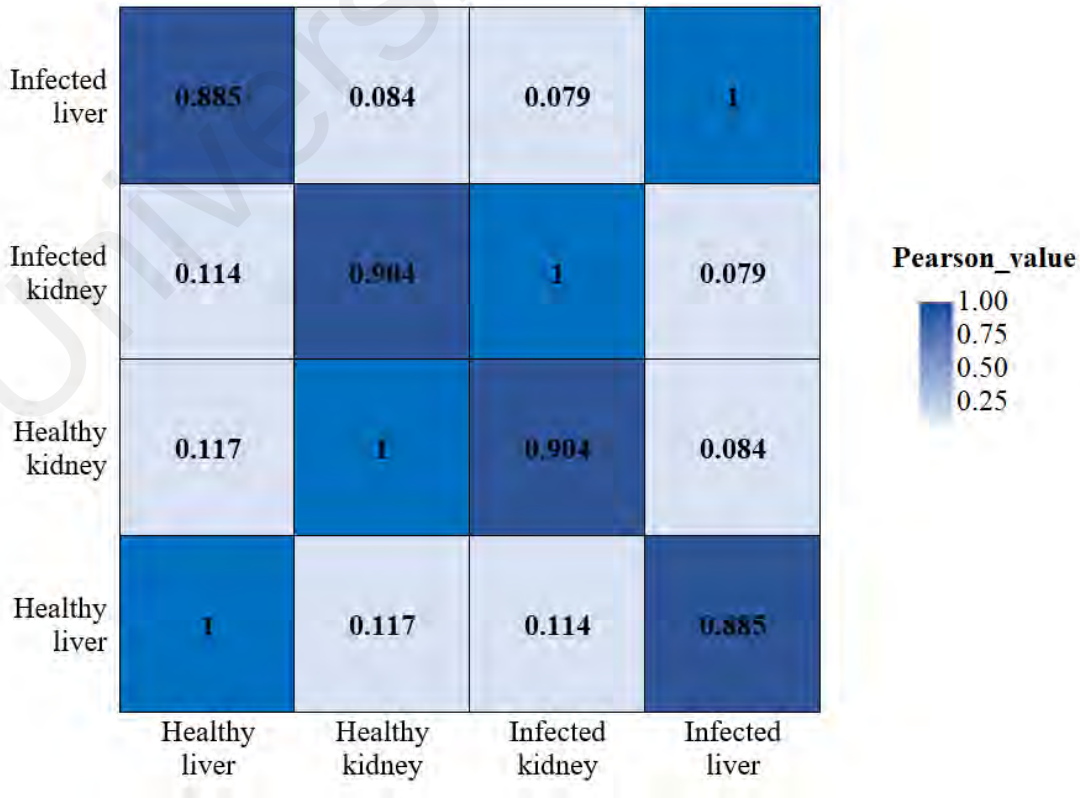


**Figure 4.34:** Heatmap of Pearson correlation between samples.

The hierarchical clustering between all samples is also performed, shown as Figure 4.35. The dendrogram illustrated the clustering pattern of healthy and infected kidney in one clade, as well as healthy and infected liver in another clade. Here, the algorithm successively pairs together samples showing the highest degree of similarity.

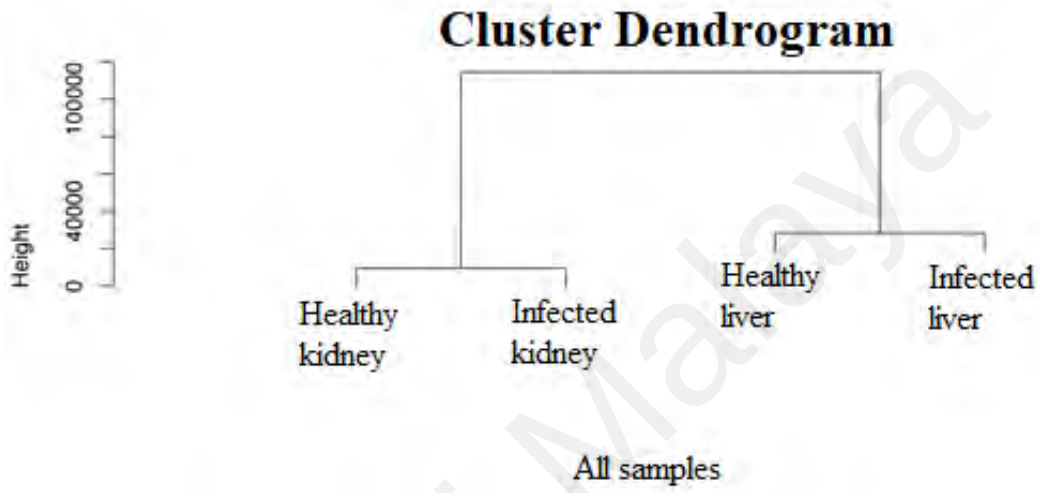


**Figure 4.35:** Hierarchical clustering between samples.

#### 4.5.6.3 The distribution of gene expression

To show the gene amount under different FPKM value, the gene amount was calculated under three different FPKM ranges (FPKM ≤ 1, FPKM 1-10, FPKM ≥ 10), shown as Figure 4.36. An infected kidney represented the highest gene number of 7476 for FPKM ≥ 10, while an infected liver represented the highest gene number of 7287 and 4829 for FPKM 1-10 and FPKM ≤ 1, respectively.

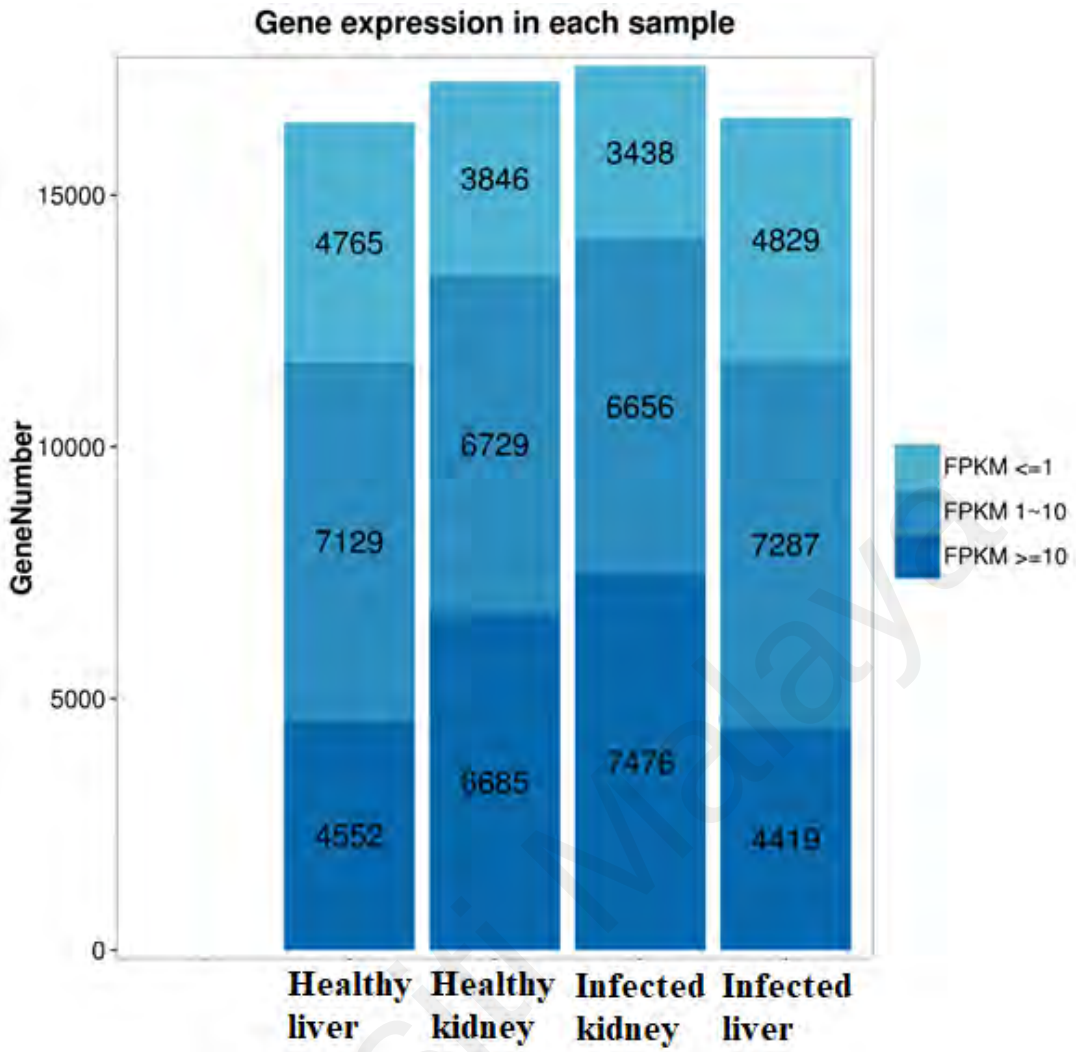


**Figure 4.36:** Gene expression distribution.

To show the amount of novel genes and known genes, the gene amount was calculated (shown as Figure 4.37) and the expression ratio was analysed (shown as Figure 4.38). An infected kidney highlighted the highest number of novel and known genes, while a healthy liver showed the lowest number of novel and known genes. Both novel and known genes were highly expressed in all four samples of healthy and infected kidney, as well as healthy and infected liver.

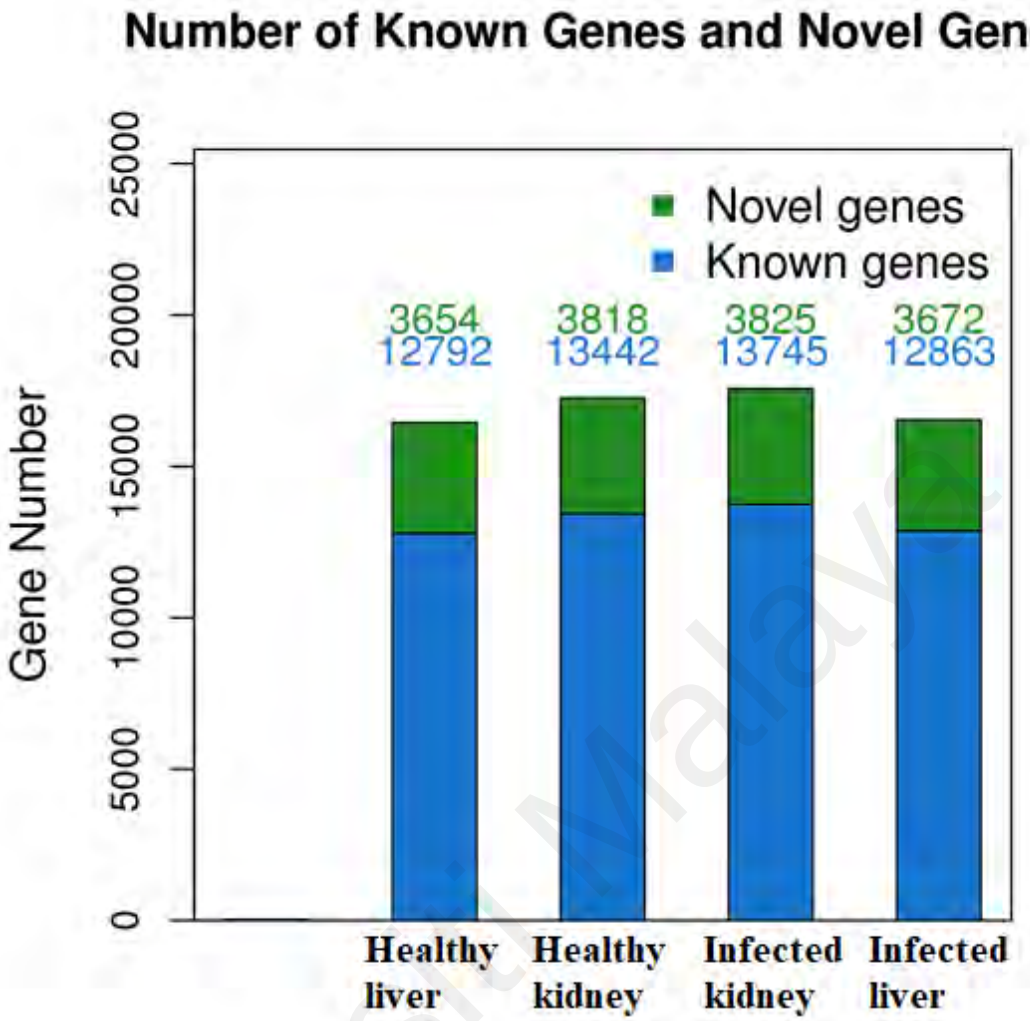


**Figure 4.37:** Statistics of novel genes and known genes.

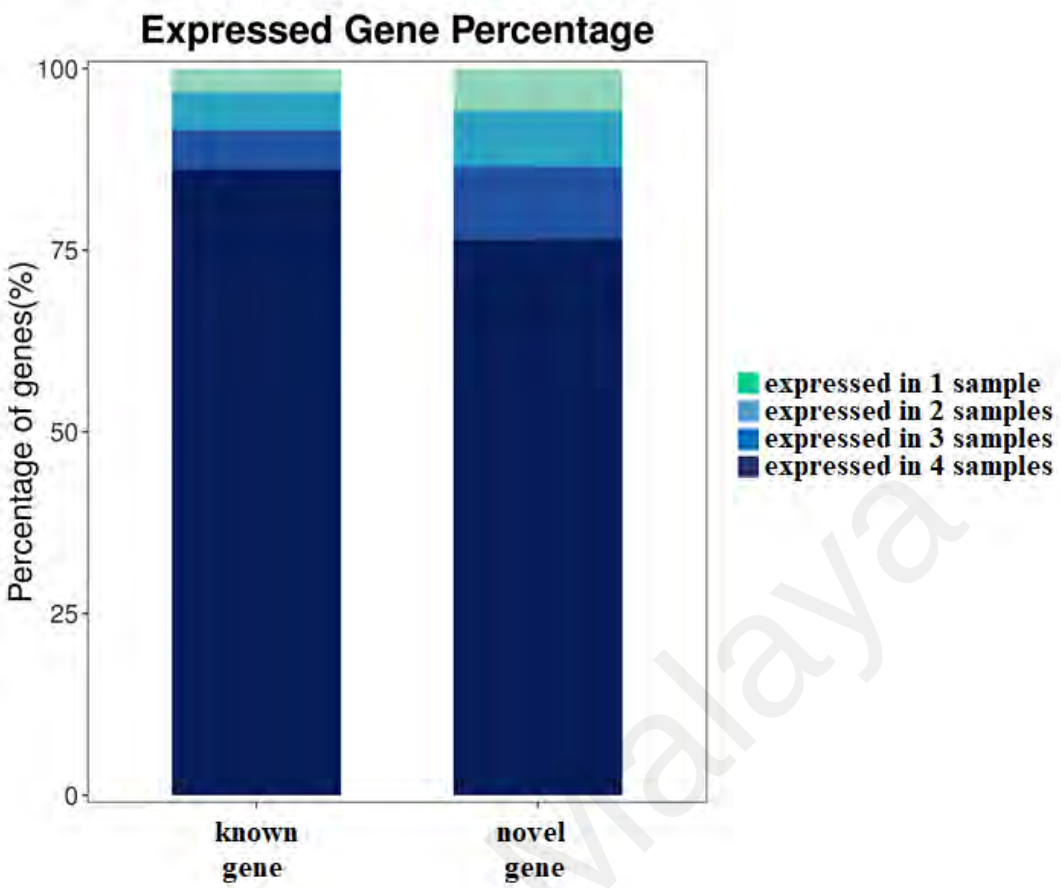


**Figure 4.38:** Expression distribution of novel and known genes.

### 4.5.7 Differentially expressed gene detection

Based on the gene expression level, the Differentially Expression Genes (DEGs) were identified between groups. The PossionDis algorithms was used to detect the DEGs. Summary of DEGs is shown in Figure 4.39. The Scatter plot was used to show the distributions of DEGs in Figure 4.40. Based on significant differences in the expression of relative transcript abundance between the healthy kidney-infected kidney genes, 3365 host genes were significantly upregulated and 928 genes were significantly downregulated by infection with leptospirosis. In contrast, cohort of healthy liver-infected liver revealed the highest number of downregulated genes of 1006 and the lowest number of upregulated genes of 885 during infection.

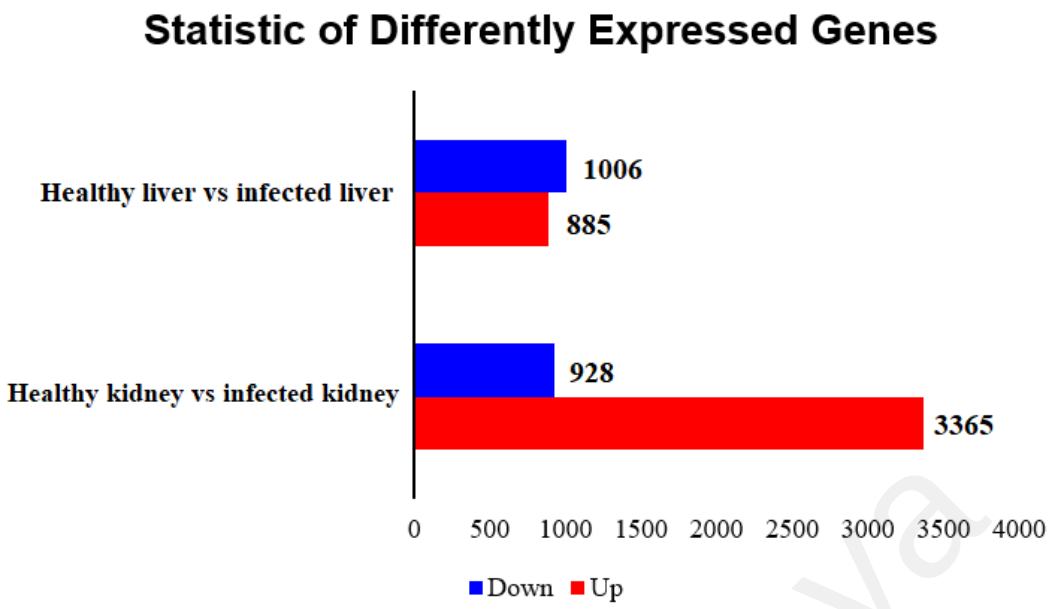


**Figure 4.39:** Summary of DEGs.

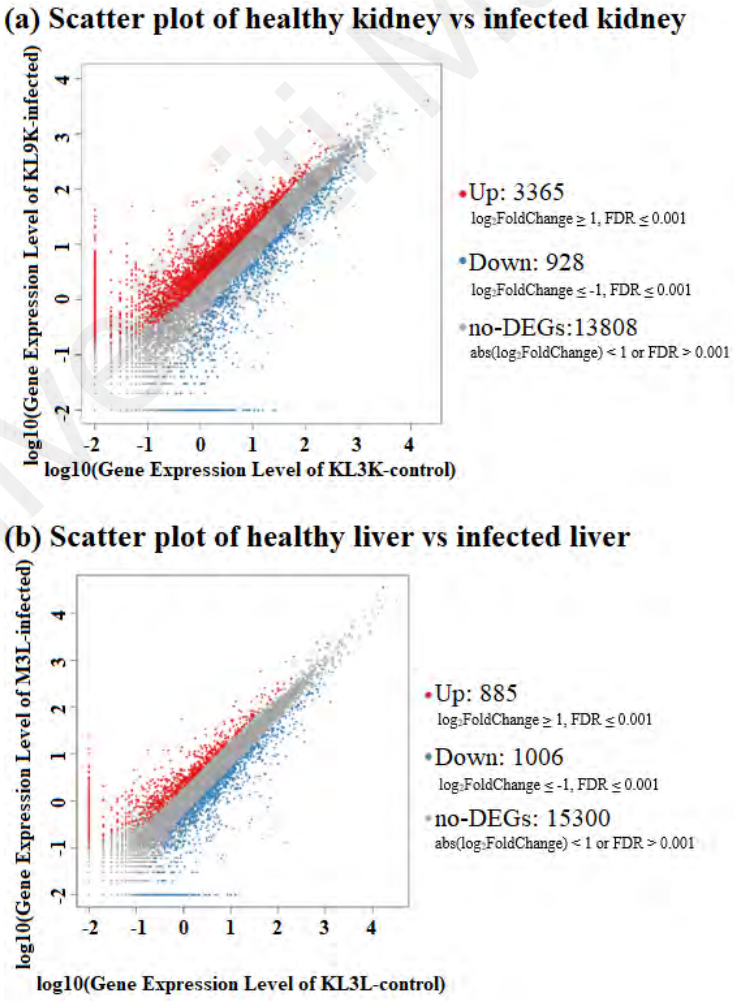


**Figure 4.40:** Scatter plot of DEGs. (a) healthy kidney vs infected kidney, (b) healthy liver vs infected liver.

### 4.5.8 Hierarchical clustering analysis of DEGs

Hierarchical clustering for DEGs were performed to further demonstrate the relationships between DEGs with various expression patterns, shown as Figure 4.41. For union DEGs, both cohorts of healthy-infected kidney and healthy-infected liver signified the low values of the log2 transformed fold change. However, the intersection DEGs of both cohorts signified the high values of the log2 transformed fold change.

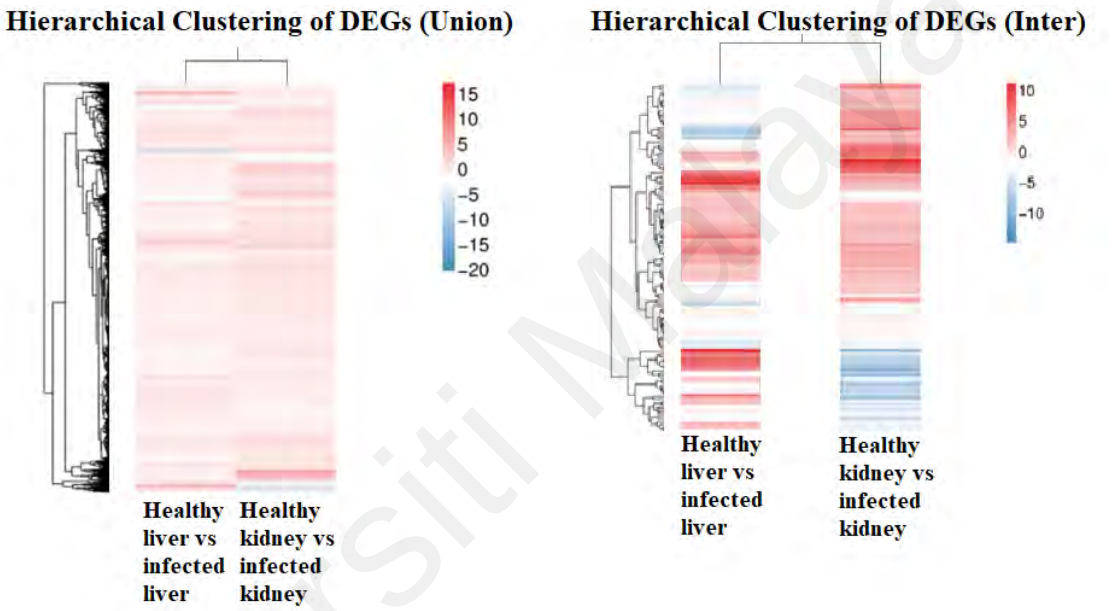


**Figure 4.41:** Heatmap of hierarchical clustering of DEGs. Colouring indicates the log2 transformed fold change (high: red, low: blue).

### 4.5.9 Gene ontology analysis of DEGs

With DEGs, Gene Ontology (GO) classification and functional enrichment were performed. GO has three ontologies: biological process, cellular component, and molecular biological function. The GO classification results are shown as Figure 4.42. For cohort of healthy-infected kidney, overall, 21234 genes (49.72%) were mapped to biological process, 16456 genes (38.53%) were mapped to cellular component, and 5018 genes (11.75%) were mapped to molecular biological function. With regard to the gene ontology (GO) assignment to biological processes, most were involved in cellular process

(12.6%), biological regulation (9.8%), and regulation of biological process (9.2%). GO assignments for the cellular component genes were associated with cell (17.2%), parts of cell (17.2%), and cell organelle (14.5%). Additionally, most of the GO assignments of the molecular function genes were associated with binding (47.9%) or catalytic activity (22.8%), with a smaller proportion associated with molecular function regulator (6.8%).

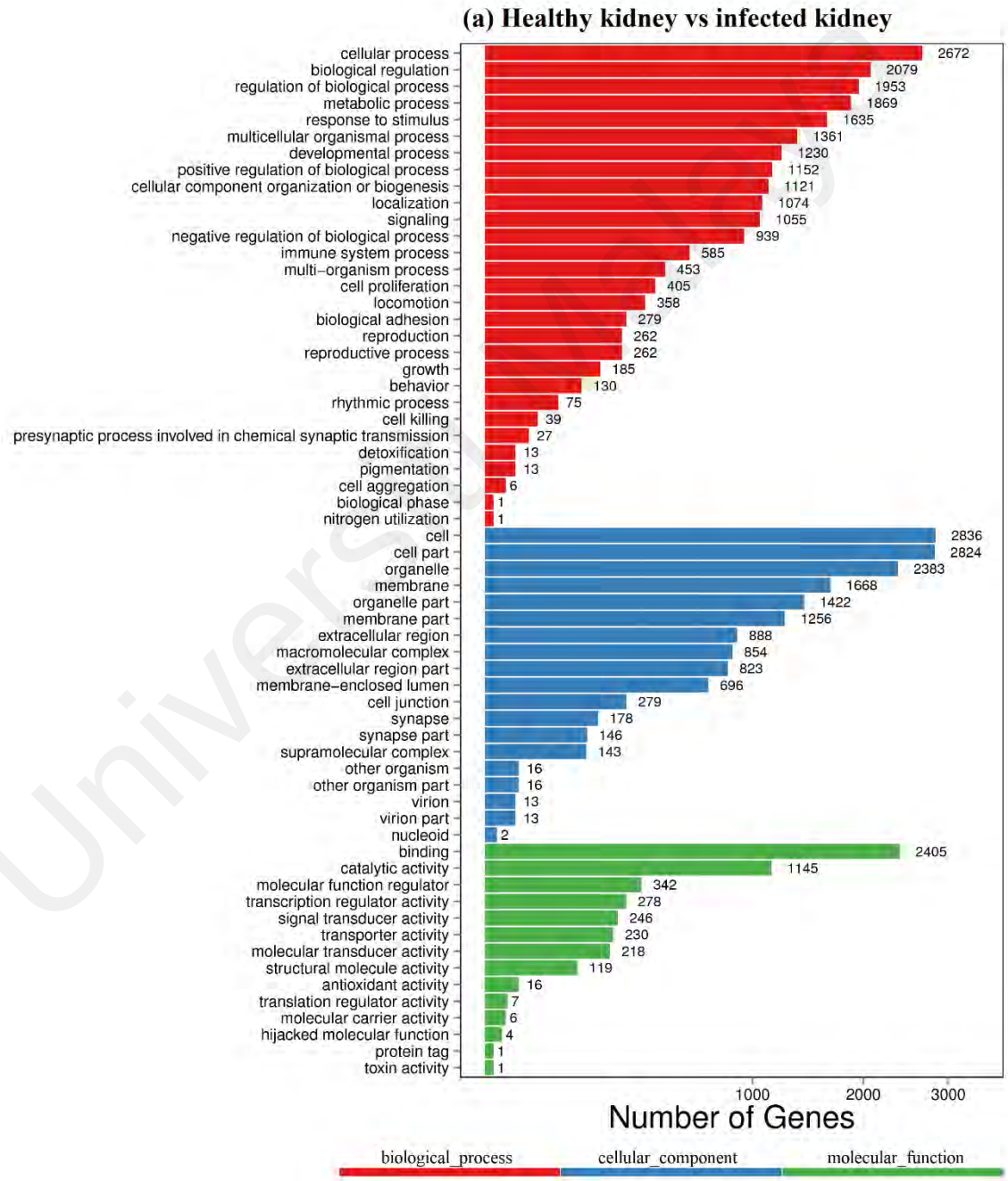


**Figure 4.42:** GO classification of DEGs. (a) healthy kidney vs infected kidney, (b) healthy liver vs infected liver.

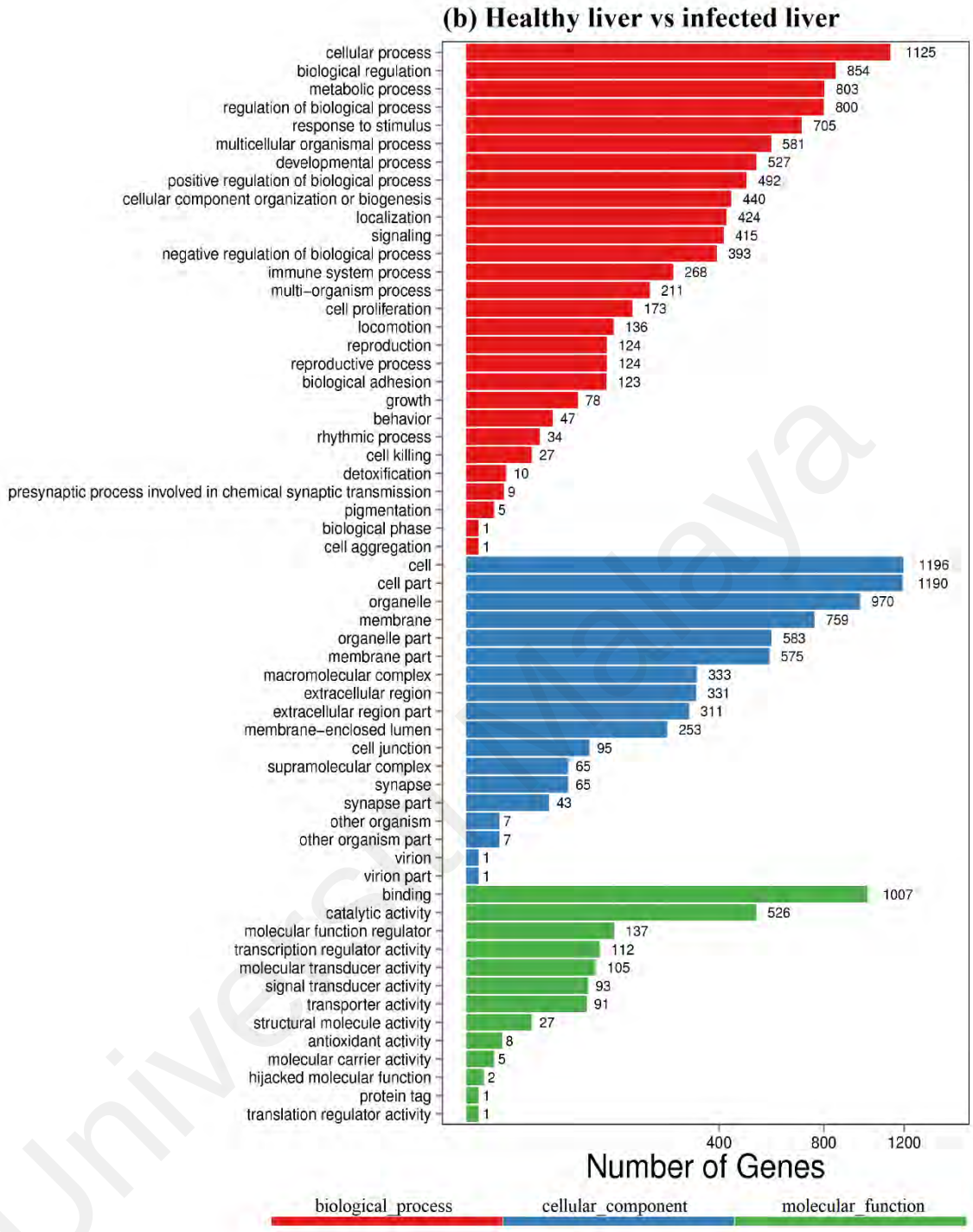


**Figure 4.42:** continued.

For cohort of healthy-infected liver, in total, 8930 genes (50.08%) were mapped to biological process, 6785 genes (38.05%) were mapped to cellular component, and 2115 genes (11.86%) were mapped to molecular biological function. With regard to the gene ontology (GO) assignment to biological processes, most were involved in cellular process

(12.6%), biological regulation (9.6%), and metabolic process (9.0%). GO assignments for the cellular component genes were associated with cell (17.6%), parts of cell (17.5%), and cell organelle (14.3%). Additionally, most of the GO assignments of the molecular function genes were associated with binding (47.6%) or catalytic activity (24.9%), with a smaller proportion associated with molecular function regulator (6.5%).

The GO classifications of upregulated and downregulated genes are shown as Figure 4.43. For cohort of healthy-infected kidney, the percentage of significantly upregulated genes were higher compared to the genes which were significantly downregulated by infection with leptospirosis. In contrast, cohort of healthy-infected liver revealed the contradicted results, in which the percentage of significantly downregulated genes were higher compared to the genes which were significantly upregulated by the infection.

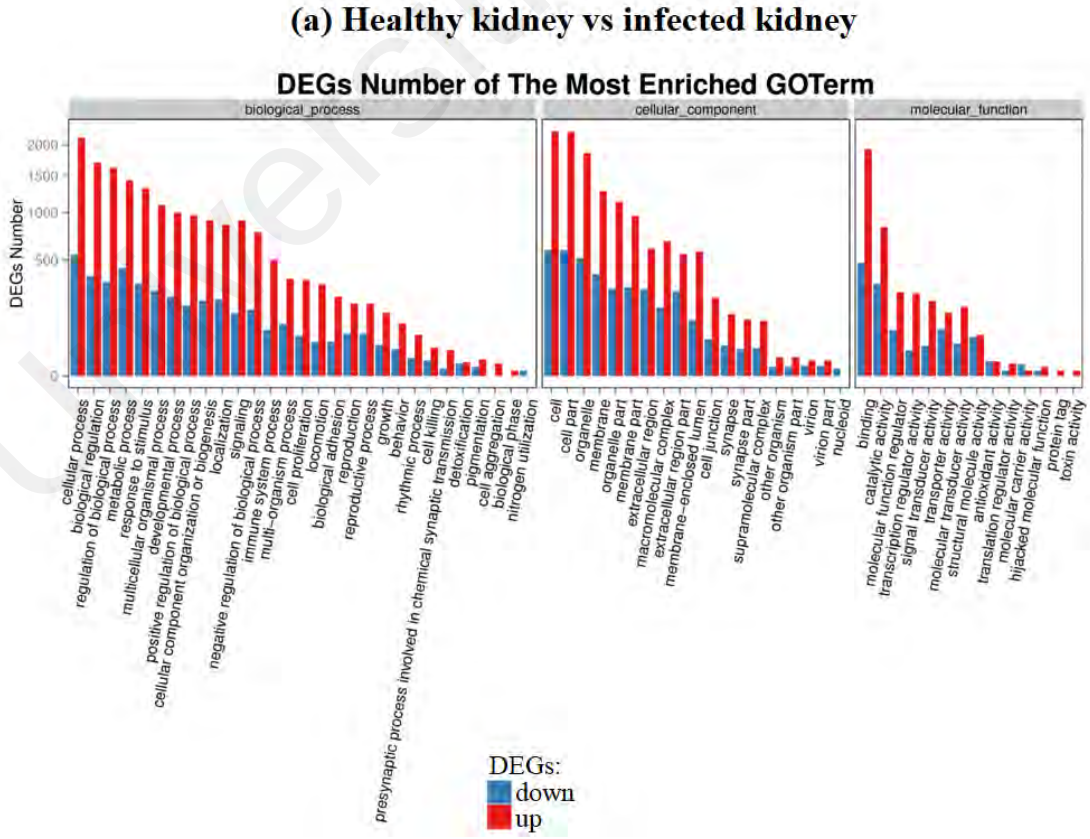


**Figure 4.43:** GO classification of upregulated and downregulated genes. (a) healthy kidney vs infected kidney, (b) healthy liver vs infected liver.

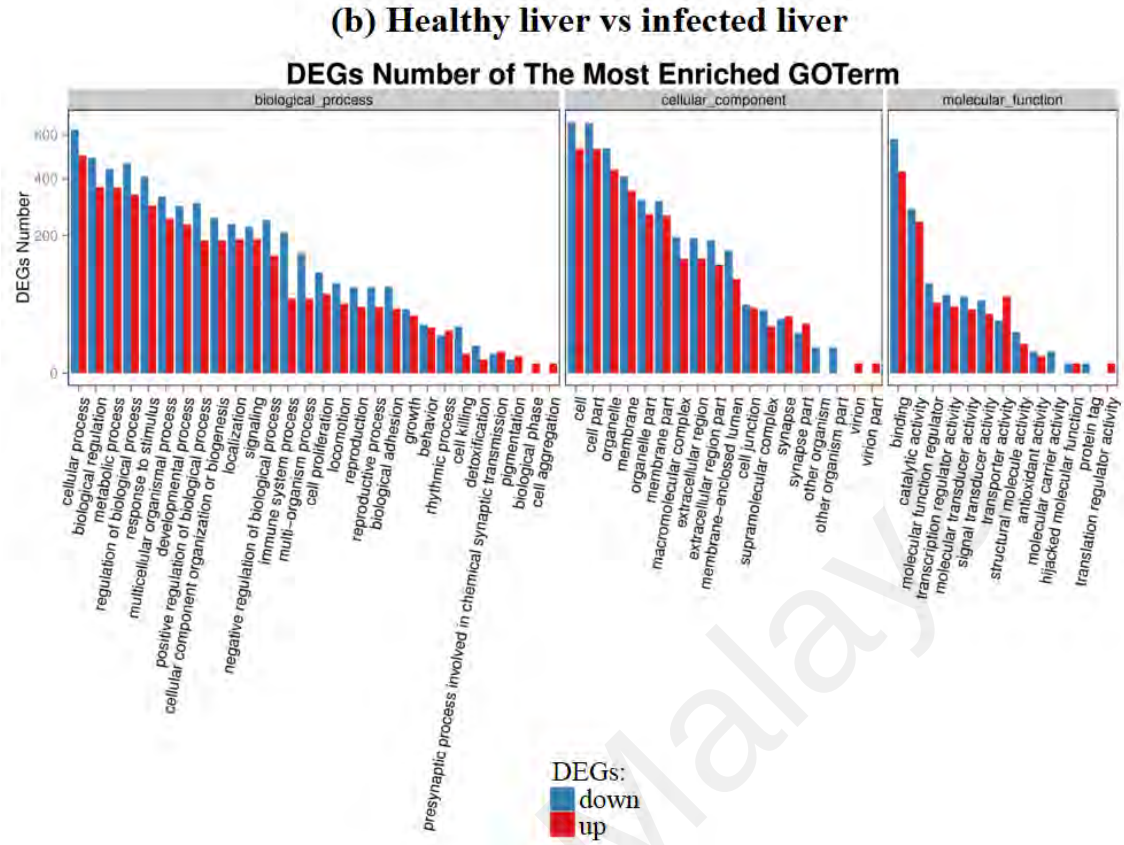


**Figure 4.43:** continued.

### 4.5.10 Pathway analysis of DEGs

With DEGs, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway classification and functional enrichment were performed. The pathway classification results are shown as Figure 4.44. All the differentially expressed genes related to leptospiral infection of the host gene were characterised by mapping them against the referential canonical pathways in the KEGG database. For cohort of healthy-infected kidney, the most abundant categories were associated with signal transduction (659 genes), immune system (508 genes), and cancers overview (450 genes). In furtherance, the most abundant categories for cohort of healthy-infected liver were related with signal transduction (279 genes), immune system (236 genes), and viral infectious diseases (212 genes).

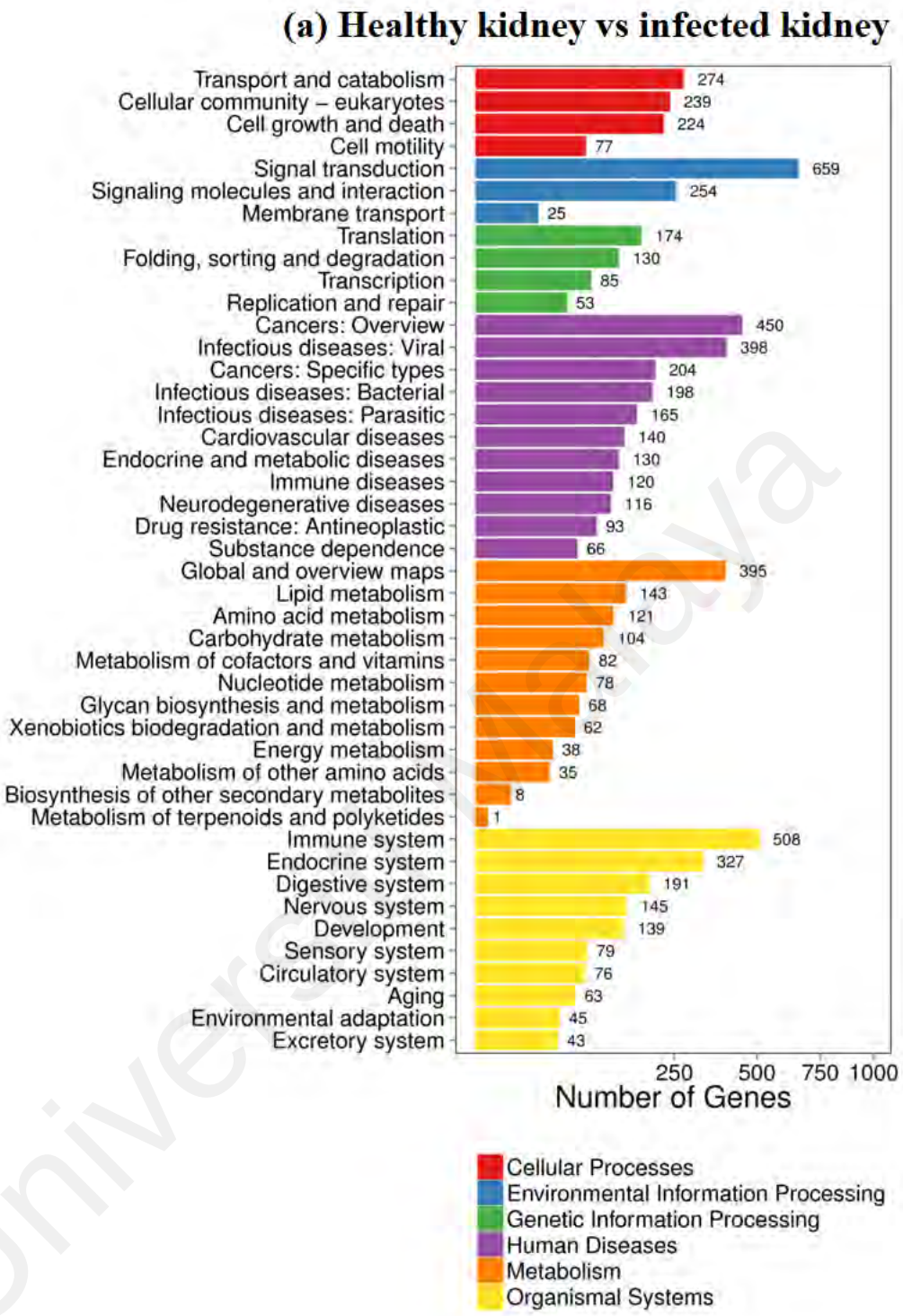


**Figure 4.44:** Pathway classification of DEGs. (a) healthy kidney vs infected kidney, (b) healthy liver vs infected liver.

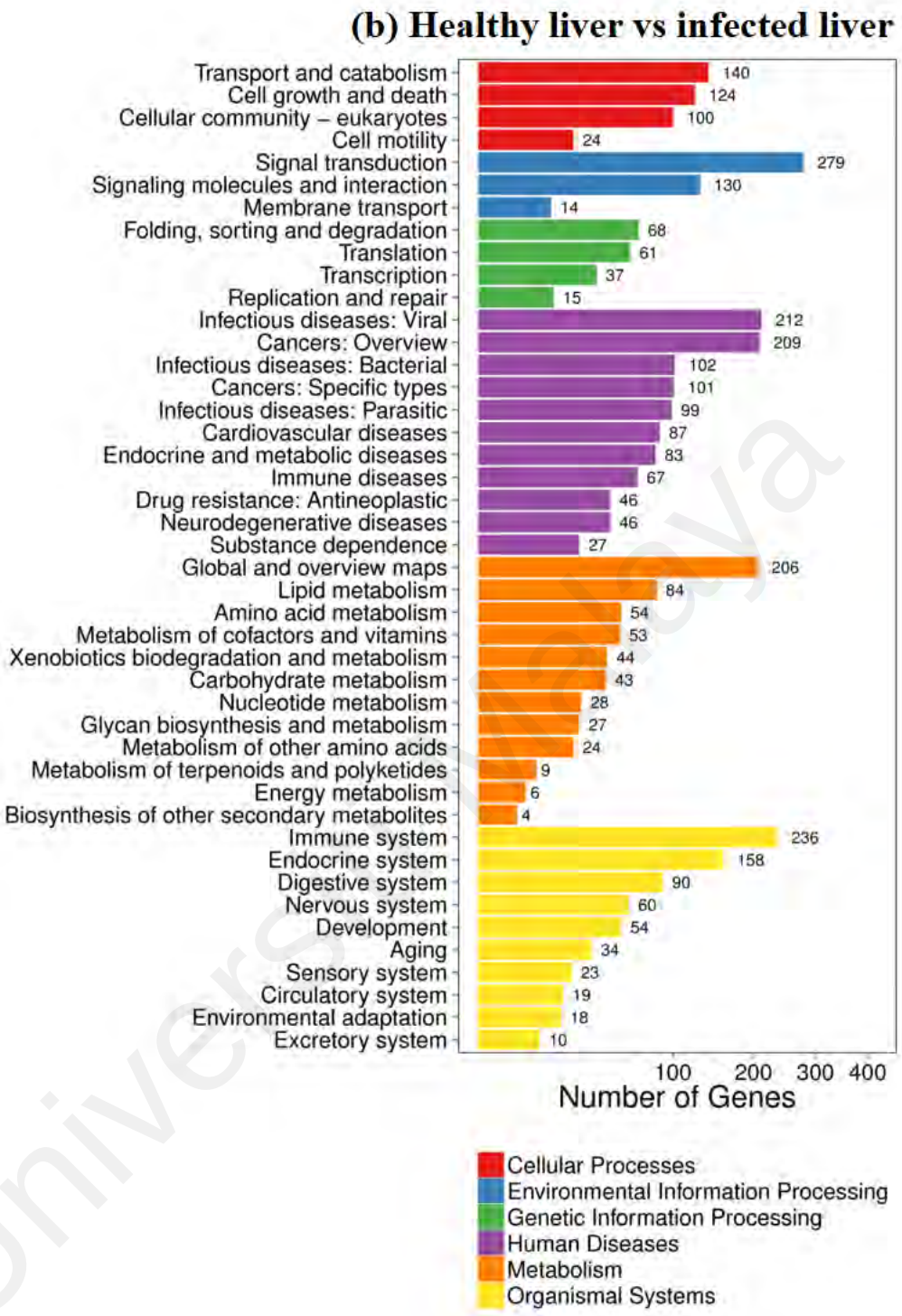


**Figure 4.44:** continued.

The pathway functional enrichment results are shown as Figure 4.45. The results highlighted the top intersection pathway functional enrichment of DEGs for both cohorts of healthy-infected kidney and liver were mostly related to Herpes simplex infection and Epstein-Barr virus infection. However, the pathway functional enrichment results for up

and downregulation genes illustrated the distinct patterns of genes regulation (Figure 4.46), in which the percentage of significantly upregulated genes were higher in cohort of healthy-infected kidney compared to significantly downregulated genes. On the contrary, the percentage of significantly downregulated genes were higher in cohort of healthy-infected liver compared to significantly upregulated genes.

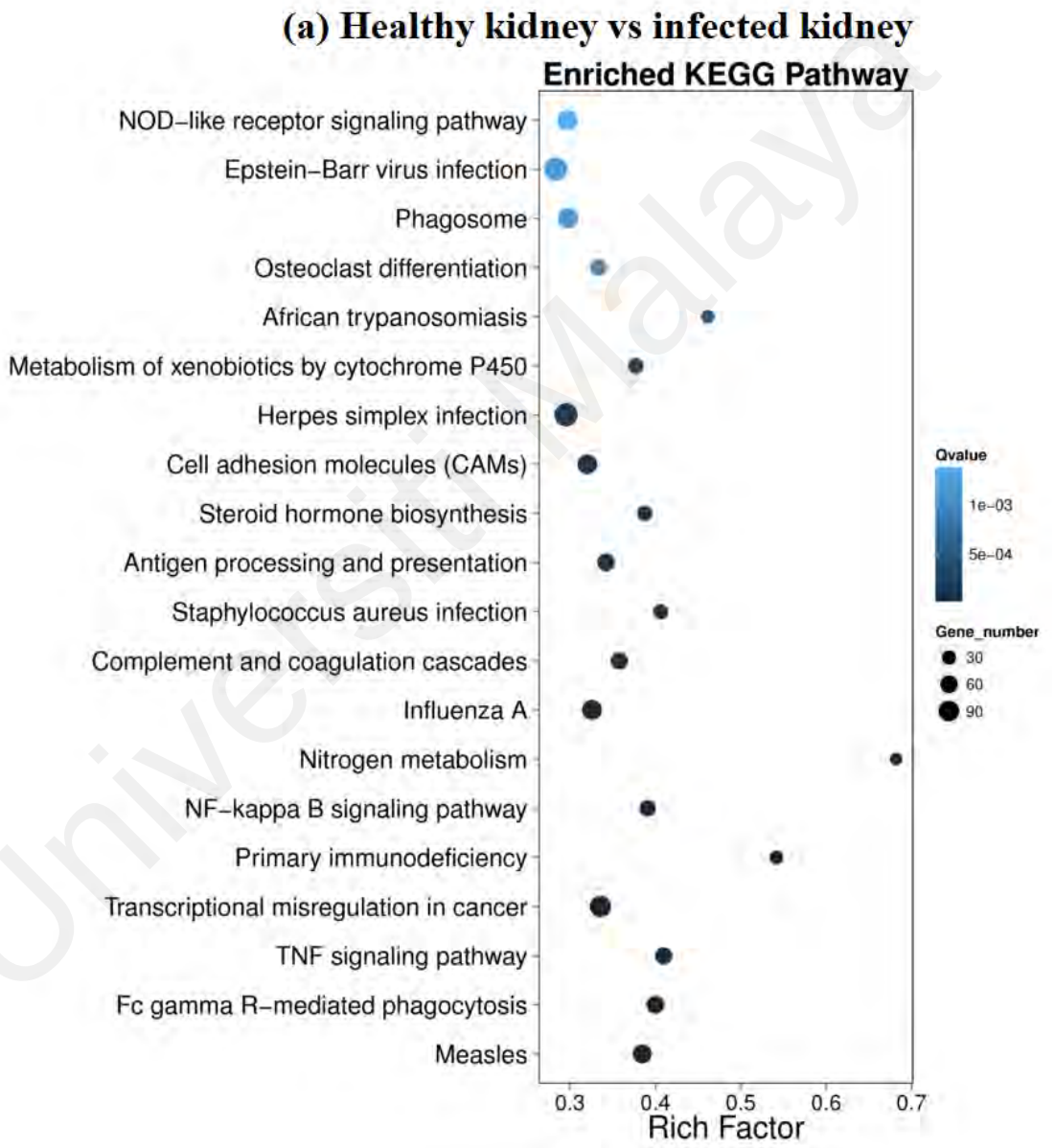


**Figure 4.45:** Pathway functional enrichment of DEGs. (a) healthy kidney vs infected kidney, (b) healthy liver vs infected liver.

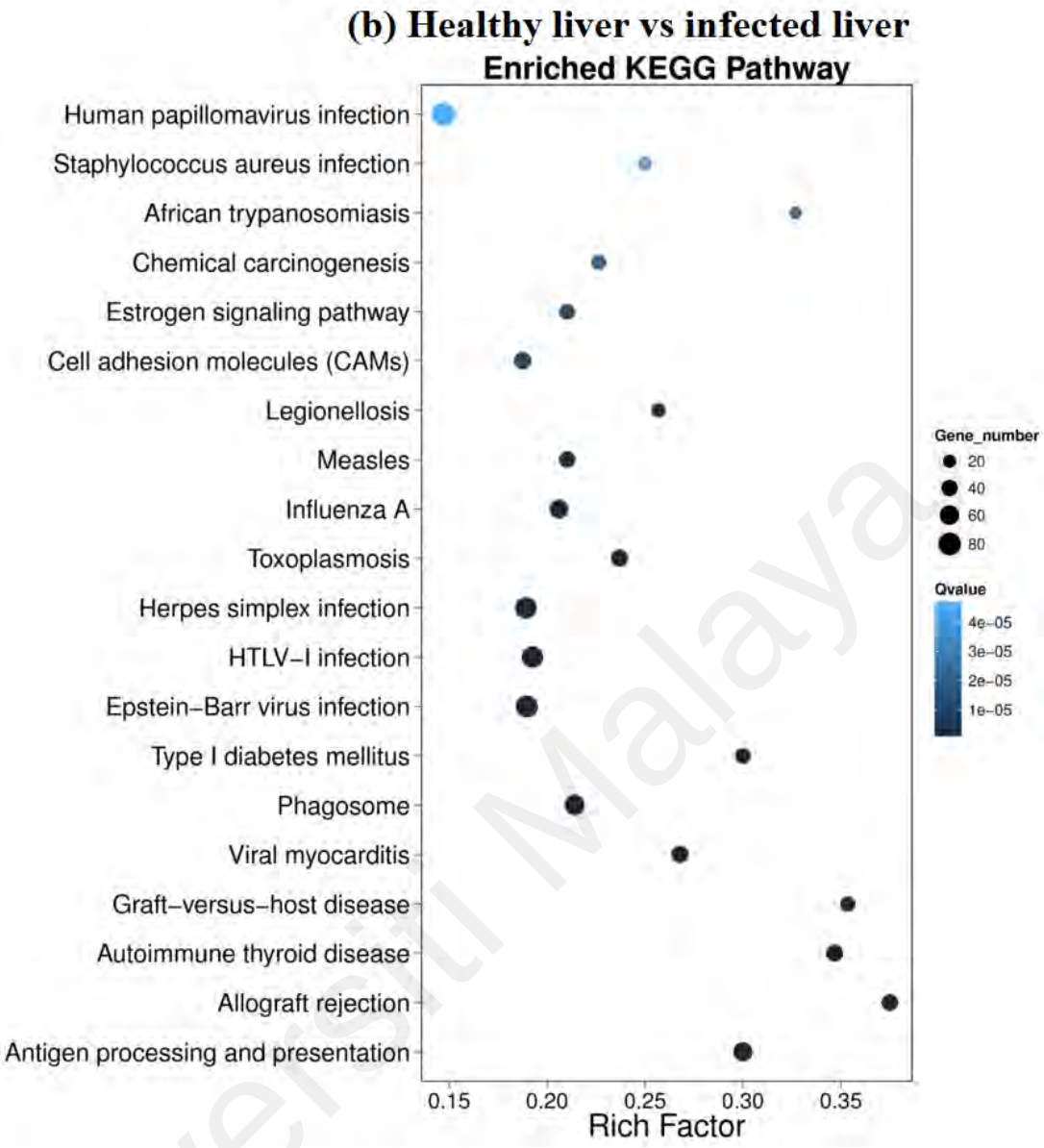


**Figure 4.45:** continued.

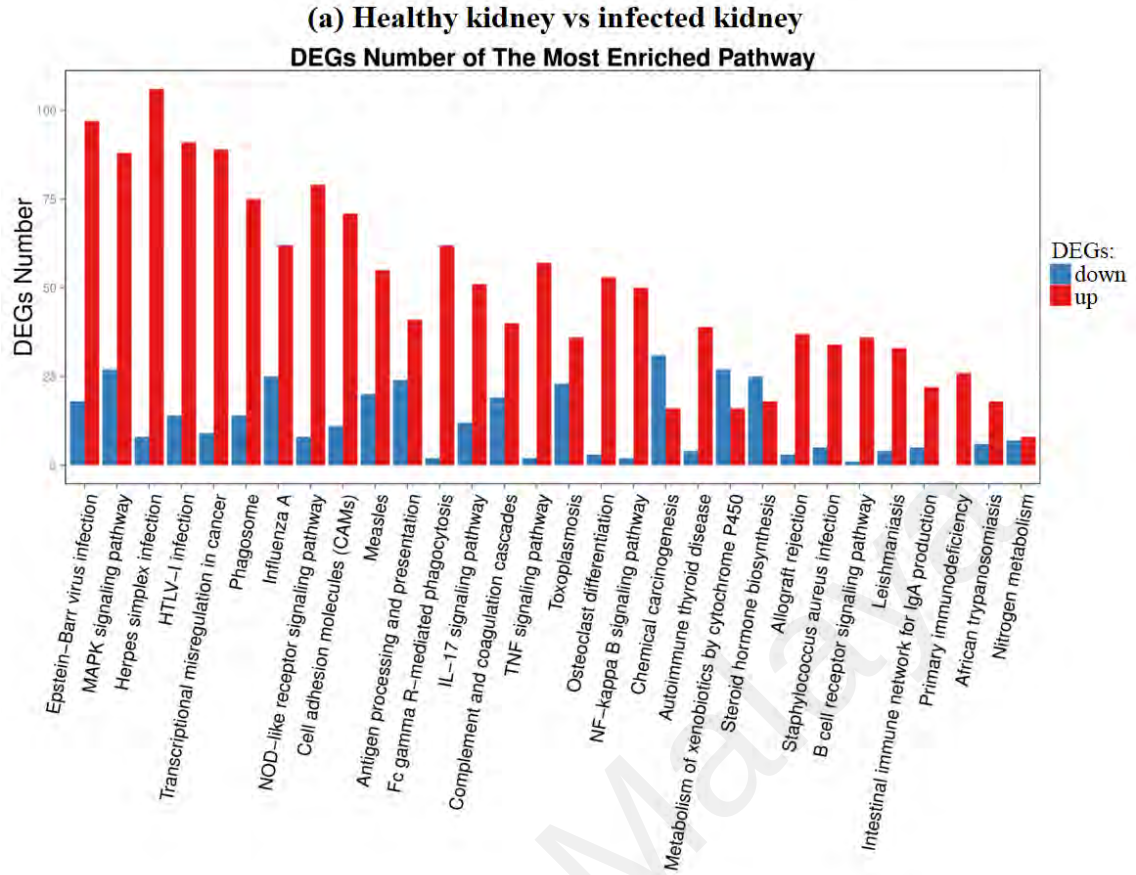


**Figure 4.46:** Pathway functional enrichment result for up/down regulation genes. (a) healthy kidney vs infected kidney, (b) healthy liver vs infected liver.

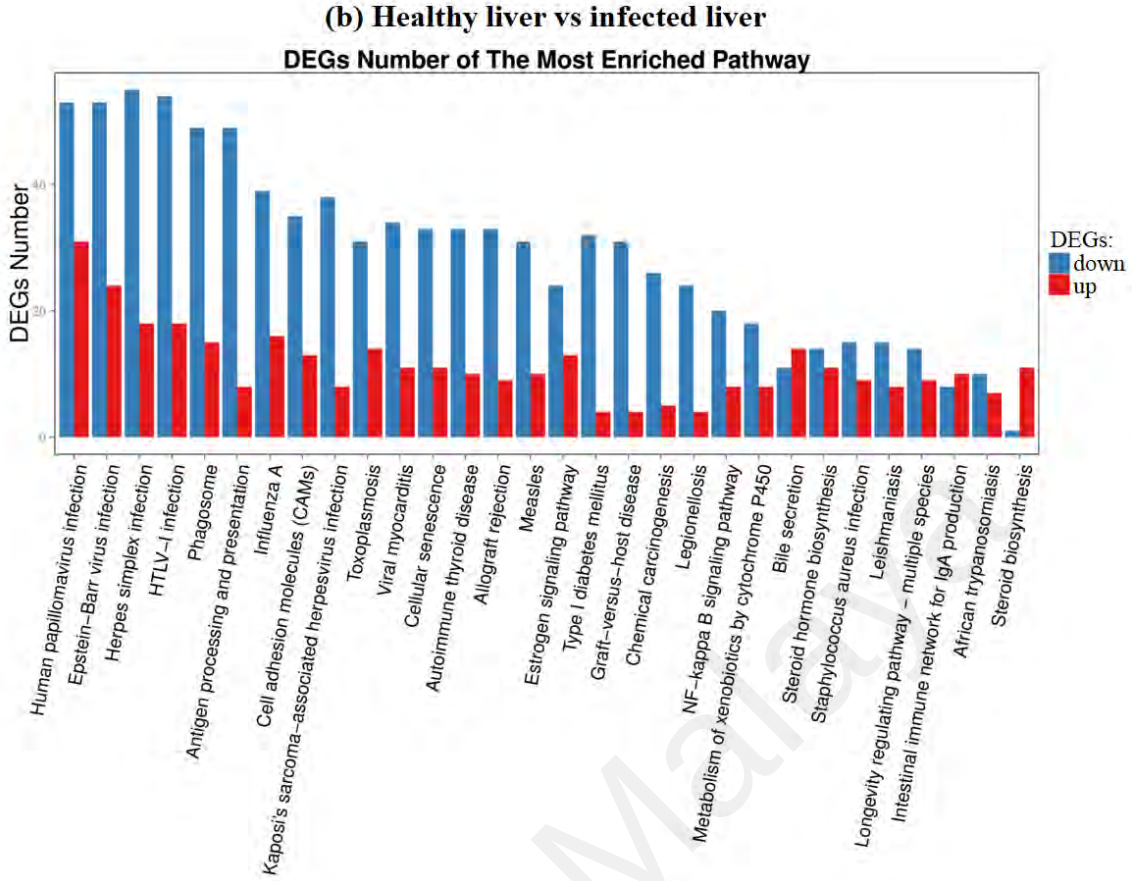


**Figure 4.46:** continued.

### 4.5.11 Quantitative real-time PCR (qRT-PCR) analysis

Several genes were found to be differentially expressed in the leptospiral-infected rats compared to the uninfected controls. Validation of the BGISEQ-500 platform sequencing results involved the choosing of six differentially expressed *R. rattus* immune genes; tumor necrosis factor-α (TNF-α), interleukin (IL)-1β, IL-6, IL-10, macrophage inflammatory protein-1α (MIP-1α/C-C-type chemokine ligand 3, CCL3), and gamma interferon (IFN-γ)-inducible protein 10 kDa (IP-10/C-X-C-type chemokine ligand 10, CXCL10) for quantitative RT-PCR analysis (qRT-PCR). These genes were greatly observed and reported in many previous studies on various hosts experimentally infected with *Leptospira* spp. Figure 4.47 shows the gel image from the optimisation of qRT-PCR primers using PCR.

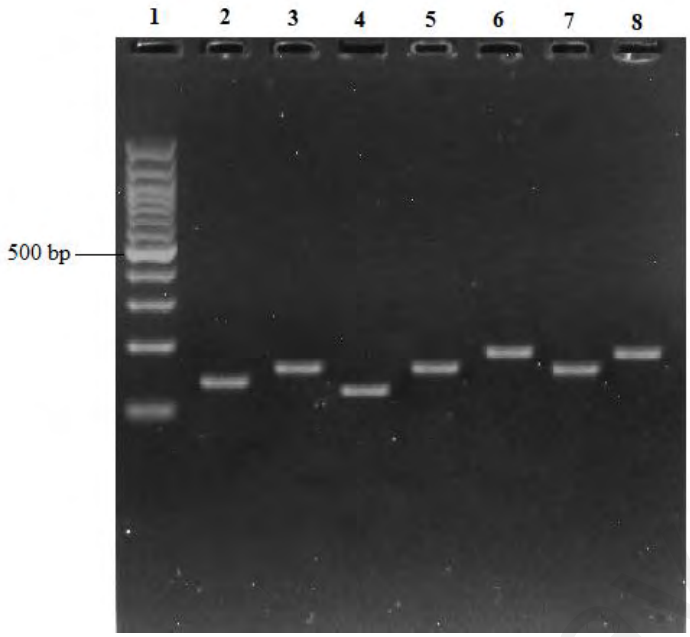


**Figure 4.47:** Optimisation of qRT-PCR primers using PCR. The annealing temperature at 60 °C was ideal for all primers. (Lane 1: 100 bp DNA ladder; lane 2: β-actin; lane 3: IL-1β; lane 4: IL-6; lane 5: IL-10; lane 6: TNF-α; lane 7: MIP-1α; lane 8: IP-10).

The levels of expressed genes were successfully quantified using relative qRT-PCR. The specificity of the qRT-PCR amplification was verified through a melting curve analysis by generating a dissociation curve (Appendix H). During leptospiral infection, all of these six genes were upregulated in the kidney (Figure 4.48), however, only interleukin-6 (IL-6) was upregulated in the liver and the remaining genes were downregulated (Figure 4.49). Overall, the expression level of these six immune genes proved the data obtained from the BGISEQ-500 platform sequencing results in both kidney and liver, which showing similar trends in the up and downregulation of the host genes.

For instance, based on the BGISEQ-500 platform sequencing analysis, TNF-α, IL-1β, IL-6, IL-10, MIP-1α, and IP-10 were upregulated 3.72, 1.74, 2.63, 1.82, 0.58, and 5.38 log2-fold, respectively; and the same elements showed 1.25, 1.85, 4.59, 2.63, 1.14, and 3.47 log2-fold change, respectively in the qRT-PCR analysis, during leptospiral infection in the kidney (Appendix I).

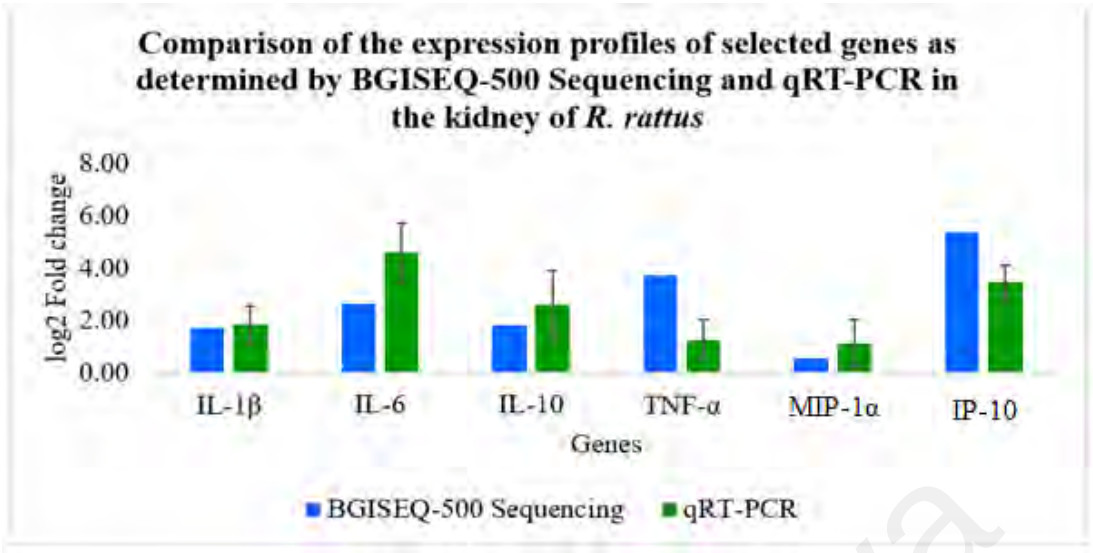


**Figure 4.48:** Comparison of the expression profiles in the kidney of *R. rattus* as determined by BGISEQ-500 sequencing and qRT-PCR.

From the BGISEQ-500 platform sequencing analysis during leptospiral infection in the liver, only IL-6 was upregulated 0.6, while TNF-α, IL-1β, IL-10, MIP-1α, and IP-10 were downregulated 1.67, 0.54, 2.97, 1.18, and 2.78 log2-fold, respectively; and the similar components showed upregulation of 1.08 by IL-6, and downregulation of 18.29, 5.70, 7.31, 18.64, and 4.22 log2-fold change by TNF-α, IL-1β, IL-10, MIP-1α, and IP-10, respectively in the qRT-PCR analysis (Appendix J).

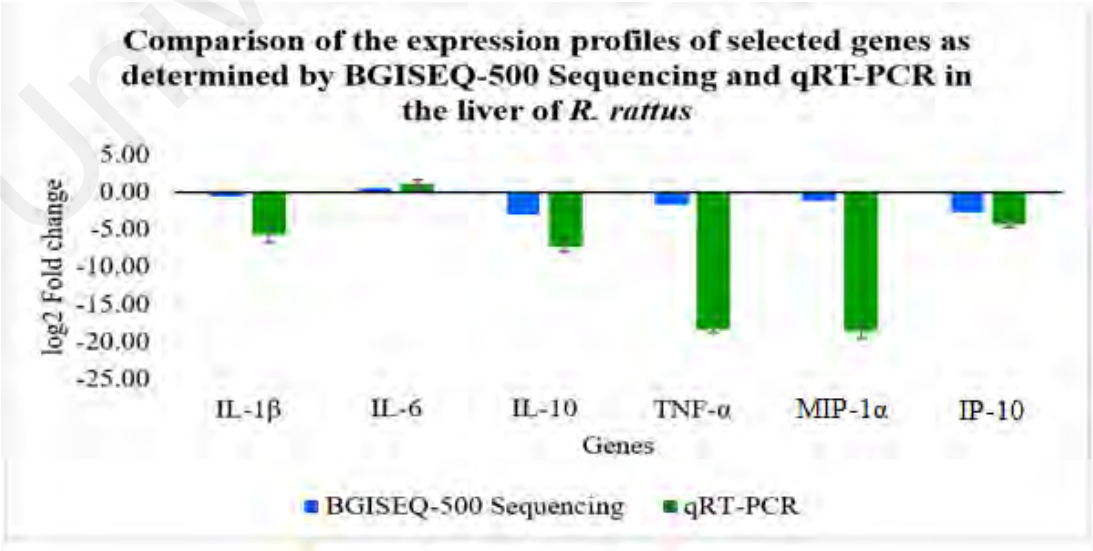


**Figure 4.49:** Comparison of the expression profiles in the liver of *R. rattus* as determined by BGISEQ-500 sequencing and qRT-PCR.

Figure 4.50 illustrated the gene ontology classification of the six annotated genes. All the six genes were corresponded to inflammatory response. Inflammation is the most common during leptospiral infection. Out of six, five of them were related in response to lipopolysaccharide (LPS) and cellular response to LPS. LPS are complex, amphipathic biomolecules found on the major surface component of Gram-negative bacteria, i.e. *Leptospira* spp. which produce LPS.

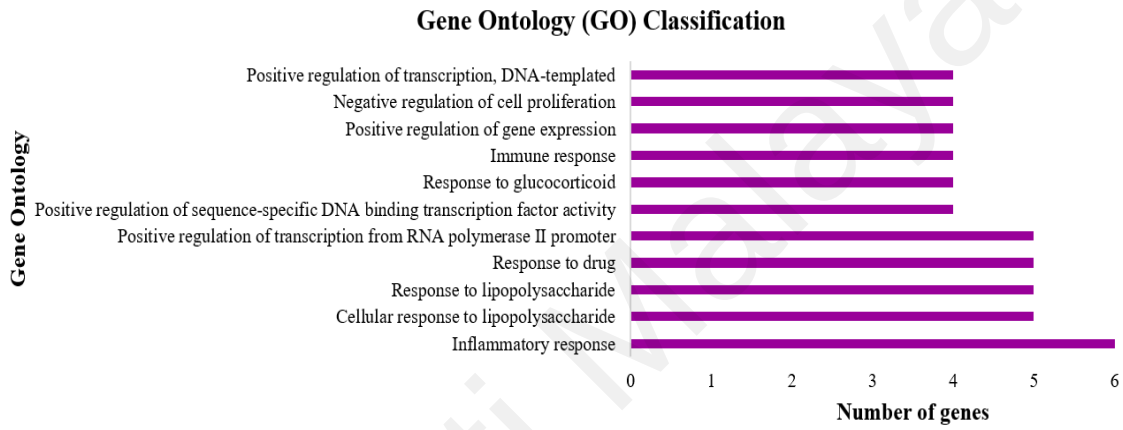


**Figure 4.50:** Gene ontology (GO) classification of the six annotated genes.

The top three KEGG biological pathway classification for the annotated genes (Figure 4.51) demonstrated the pathway of cytokine-cytokine receptor interaction which engaged in innate as well as adaptive inflammatory host defences, Chagas disease which is an inflammatory disease, and toll-like receptor (TLR) signalling pathway. TLR play a crucial role in innate immune responses. They involve in the first line of defence against invading pathogens and serve as an important role in inflammation.

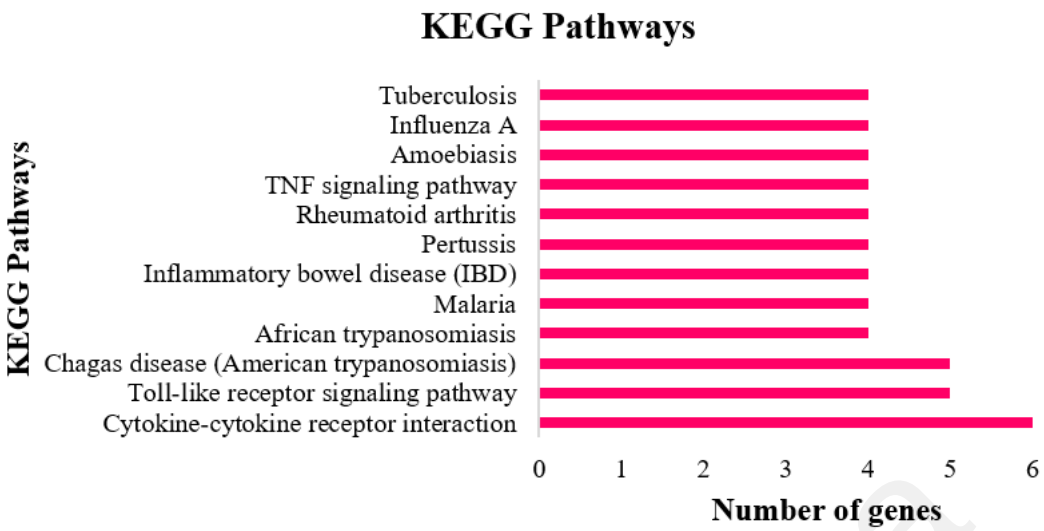


**Figure 4.51:** KEGG pathway classification for annotated genes.

# CHAPTER 5: DISCUSSION

## 5.1 Species Diversity of *Rattus* (Rodentia: Muridae) on the Malay Peninsula and Their Phylogenetic Relationships Inferred from Complete Mitochondrial Cytochrome b Gene Sequences

### 5.1.1 The number and identity of *Rattus* species in Peninsular Malaysia and the *Rattus rattus* complex

The results from this study suggest that there are six species of *Rattus* on the Malay Peninsula: Asian house rat (*R. tanezumi*), Malaysian wood rat (*R. tiomanicus*), Norway/brown rat (*R. norvegicus*), Pacific rat (*R. exulans*), an unnamed black/house rat (*R. rattus cf*. lineage IV sensu Aplin et al., 2011), and the ricefield rat (*R. argentiventer*). Francis (2008) also listed six species in the genus *Rattus* on Peninsular Malaysia, but the list differs in the addition of the Asian house rat, the adjustment of the black rat from *R. rattus* to a different unnamed species in the same complex, and the subtraction of *R. annandalei* (which is present in the Malaysian Peninsula, but is no longer in the genus *Rattus*). The results are in accordance with the suggestion by Aplin et al. (2011) that *Rattus rattus* sensu stricto should be absent in the region.

Mitochondrial lineage terminologies on the *Rattus rattus* complex mentioned by Aplin et al. (2011) suggested that *R. rattus* complex lineage I group (*Rattus rattus* sensu stricto) is the true Oceanian *R. rattus* with chromosome numbers of 2n = 38 (Pagès et al., 2010; Robins et al., 2007; Chinen et al., 2005; Musser & Carleton, 2005; Aplin et al., 2003). Recently, Aplin et al. (2011) proposed the taxonomic usage of *R. rattus* LI for this group. Following them, the terminology is widely used even in this present study. The representatives of this group usually present in India and Europe. Hence, this group was not sampled in Peninsular Malaysia. *Rattus rattus* complex lineage III group also categorised as a part of Asian *R. rattus*, but the chromosome numbers are incompletely

studied. Yosida (1980) examined the sampled Himalayan populations to have diploid chromosome numbers of 42. Aplin et al. (2011) also suggested that the chromosome numbers might be similar to Rrc LII. Nepal and Pakistan are the two countries which rats of this group were found. Thus, it is not captured in Peninsular Malaysia. Samples exist in this group is uncategorised by Aplin et al. (2003). Nonetheless, Musser & Carleton (2005) classified it as a part of *R. tanezumi* and it was further named as *R. rattus* LIII by Aplin et al. (2011). *Rattus losea* with a diploid chromosome number of 42 is ordered under *R. rattus* complex lineage V group (Aplin et al., 2011). However, it is not sampled in the present study and the taxon is not further discussed herein.

***Asian house rat***

For *R. rattus* complex lineage II group, it is classified as a part of Asian *R. rattus* with chromosome numbers of 2n = 42. Previous study by Aplin et al. (2003) suggested that rats belong to this group as a northern *R. tanezumi*, however Pagès et al. (2010), Robins et al. (2007), Chinen et al. (2005), and Musser & Carleton (2005), applied the taxonomic name of *R. tanezumi* without emphasising on the geographic region. Further work by Aplin et al. (2011) proposed the rats in this group as *R. rattus* LII. The present study named the rats belonging to this group as *R. tanezumi*. From the collections in the city area, *tanezumi* rats had a gold brown to reddish brown fur colour blended with black fur on the upper part and a uniform distribution of light-brown hair on the belly.

*Rattus tanezumi* originated in Southeast Asia and distributed to east Asia and Africa with humans and became invasive in many places (Musser & Carleton, 2005). Musser & Carleton (2005) proposed a two-taxon segregation of this species: an ‘endemic southeast Asian taxon’ and a ‘northern and south Asian taxon’. The Asian house rats sampled here in Peninsular Malaysia were very closely related to previously reported sequences of

Asian house rats from Vietnam, Myanmar, China, Japan, Taiwan, and Laos. There is, however, a deep divergence between the haplotypes in this widespread lineage and those reported from Nepal. The type locality is in Japan, so if the species needs to be split, the widespread lineage which includes the animals sampled on the Malay Peninsula, would retain the name.

***An unnamed black rat***

The *R. rattus cf*. lineage IV sensu Aplin et al., 2011 is designated as part of the Asian *R. rattus* complex because of its diploid chromosome numbers of 40. This is due to the 2n = 40 karyotype of *R. kandianus* found in central highlands of Sri Lanka, as described by Yosida (1980). The taxonomic name used for rats in this group is the most controversial. Aplin et al. (2003) and Musser & Carleton (2005) suggested that they belong to *R. tanezumi*, Robins et al. (2007) referred to them as *diardi* clade, R3 was utilised by Pagès et al. (2010), and finally Aplin et al. (2011) used *R. rattus* LIV.

The phylogeny suggests divergence between Malay Peninsula rats at south of the Isthmus of Kra and those north of that biogeographic barrier. Possibly, this finding coincided with the north-south of the Isthmus of Kra that may have influenced a postulated disjunction to represent geographical patterns of population subdivision within *R. rattus* lineage IV. The populations of *R. rattus* which were isolated in the northern side of the Isthmus of Kra became the founders of the Indochinese (Vietnam, Laos, Cambodia) populations, whereas those on the southern side founded the Sundaic Thai Gulf (Peninsular Malaysia, southern Thailand, Philippines) populations. The divergence of the northern-southern side has occurred rapidly around 0.39 Mya. This divergence time corresponds to the Pleistocene epoch, when the sea level was up to 120 m below the present-day level (Voris, 2000) and a land bridge was exposed. In the meantime, rats that

lived in the southern side of the Isthmus of Kra (Sundaic Thai Gulf clade) might have been restricted in their distribution range by the Nakhon Si Thammarat mountain range in southern Thailand. Therefore, the migration of north-south populations was hampered.

*Rattus rattus* lineage IV occurs to be morphologically alike with *R. tanezumi*, which their distributions are extended throughout the Asian region (Aplin et al., 2011; Pagès et al., 2010; Robins et al., 2007). Inconclusive results may arise if the classification of species depends entirely on the morphological traits which might be dominated by the surrounded environmental factors. Rats in this group are normally large, have a longer tail, are good at climbing, and occupy primary forest to urban areas. Through the field surveys in urban areas, the observed rats have greyish brown to greyish black fur on the back and brownish white or grey on the belly. However, variation in fur colour within or between populations might be due to the environmental adaptations (habitat preference) and biological factors (age and sex) (Rios & Álvarez-Castañeda, 2012). The type locality of the species *R. diardi* is Java, so it is not obvious that this name would apply to these animals. This group may be an additional native species and deserves additional taxonomic attention.

***Malaysian wood rat***

For *R. rattus* complex lineage VI group, *R. tiomanicus* with associated karyotypes of 2n = 42 is responsible within this group. *Rattus tiomanicus* is the preferable name used by several taxonomists (Pagès et al., 2010; Robins et al., 2007; Musser & Carleton, 2005; Aplin et al., 2003). Aplin et al. (2011) recommended the co-occurrence of two taxa in this group, *R. tiomanicus* and *R. baluensis*. Both of this species are closely related (Musser, 1986). Yet, the present study treated both taxa as a separate taxonomic name to further elucidate their relationships. Although *R. baluensis* is known only from Mount Kinabalu

in Northern Borneo at a higher elevation, it has a very close relationship with the *R. tiomanicus* trapped at the gardens with a bushy area in a lower elevation (Camacho-Sanchez et al., 2018). It is believed that this species might be adaptable to habitat modification and the isolated populations preventing the genes exchange which contributes to allopatric speciation. The present findings demonstrated that *R. tiomanicus* and *R. baluensis* shared several familiar features that clustered them as sister-species taxa.

*Rattus tiomanicus* have morphological appearances of brown fur at the upper parts and light grey to white at the belly part. This is a widespread Sundaland endemic species, which is not known to be invasive. Rats in this group generally are medium to large size, have long tail, and scansorial. They usually live in primary forest, gardens or plantations areas, and occasionally in villages. There is no controversy about its specific status (Musser & Califia, 1982).

***Ricefield rat***

This species is tightly associated with agriculture and has a disjunct distribution in Indochina and the Malaysian Peninsula. From the data observed, the Malay Peninsula rats are very closely related to animals of Vietnam. In the glacial period, some *R. argentiventer* from southern Indochina were favoured expansion and emigration by the huge continental shelf of Sundaland. They may have then migrated towards the south-west across the land bridge and colonised the southernmost part of Thailand and northern part of Malay peninsula, represented by the current *R. argentiventer* populations. The scenario of *R. argentiventer* migration across the land bridge of Sundaland in the Pleistocene epoch is also supported by previous studies that found a relationship between the macaques of Indochinese and Peninsular Malaysian populations (Abdul-Latiff et al., 2014). Together with the fact that in the glacial period the northern latitude was cooler

and drier compared with the equatorial zone, the forest area mainly remained on the Sundaic region, whilst the Indochinese region was largely covered by savanna (Meijaard, 2003; Heaney, 1991). These factors could induce the migration of *R. argentiventer* in Indochina towards the south, as proposed above, forming the close relationship between the Vietnam and Peninsular Malaysia populations. This could suggest that ricefield rats are also invasive in Peninsular Malaysia. It might be possible that the close relationship between these two populations is caused by the incomplete lineage sorting which may occur during the evolution process of the mtDNA (Bunlungsup et al., 2017).

***Pacific rat***

This species is widely invasive, and is perhaps the first species of *Rattus* to dramatically expand its range in association with human travellers. It has been suggested that it may also be a species complex (Wilson & Reeder, 2005). The sequences found in the Peninsular Malaysian animals were very similar to animals from distant parts of their range, suggesting a recent common ancestor. The native range of this species is unknown, but presumably somewhere in Indochina.

***Brown rat***

The brown rat, *R. norvegicus*, is not a member of the *R. rattus* species complex, and diverged quite early in the history of the genus (Pagès et al., 2010). Its native range is likely central or northeastern Asia, and it is widely invasive (Wilson & Reeder, 2005). The data from this study support the recent invasive status of this species in Peninsular Malaysia.

### 5.1.2 The disclosure of the *Rattus* spp.

Of the six species of *Rattus* present on the Malaysian Peninsula, three species (the brown rat, the Pacific rat, and the Asian house rat) are clearly invasive in this region. This leaves the unnamed black rat, the Malaysian wood rat, and the ricefield rat as the possibly native species in the genus *Rattus* present on the Malaysian Peninsula. This has evolutionary implications in that this low species diversity suggests that the diversification of this genus took place outside of the Malay Peninsula, north of the Isthmus of Kra in Indochina. Perhaps more importantly this has important public health implications because these invasive species of *Rattus* has been associated with the spread of disease and disease vectors on a wide scale.

Current molecular research expected an emergence of *Rattus* in the mid Pliocene around 3 million years ago (Mya) (Robins et al., 2008; Jansa et al., 2006; Verneau et al., 1998; Furano & Usdin, 1995). The number of *Rattus* species extended hastily in the Pleistocene (Chaimanee & Jaeger, 2001), and the black rat is believed to have a complex evolutionary history after experienced the great divergence of mitochondrial lineages event during the mid Pleistocene, around 1.0 Mya (Aplin et al., 2011). Oceanian form of *R. rattus* which originated in the Indian peninsula arrived Britain by the 3rd century AD from the origin (Innes, 1990). The dispersal of the rats throughout the world started in Europe along the shipping routes (Atkinson, 1985). Where the Asian form is native, the Oceanian form is limited to the coastal areas.

The phylogenetic trees underpin these taxonomic status; the *R. rattus* lineage I clade, exhibiting the Oceanian form, accommodates samples from Australia, Iran, France, Papua New Guinea, New Zealand, Brazil, Madagascar, USA, Venezuela, Senegal, Guyana, Japan, South Africa, South India, Oman, and Sri Lanka. Asian form of the *R. rattus*

lineage III clade representing Nepal and Pakistan, while the *R. rattus* lineage IV clade displaying Southeast Asia (e.g. Peninsular Malaysia, Philippines, Indonesia, Vietnam, Cambodia, Laos, southern Thailand), and Sri Lankan rats. From the plot of Sri Lankan rats in lineage I and lineage IV, an assumption of different chromosome numbers circulating Sri Lankan rats can be made, which also contributes to the complexity of the geographical distribution.

*Rattus rattus* from Southeast Asia yielded Cyt b sequences of the same lineage with the more often encountered lineage of 2n = 40, which emerges to be unique to Sri Lanka. Also, *R. tanezumi* found to be grouped with others Asian *R. tanezumi*, with 2n = 42. However, this finding has to be examined thoroughly with cautious. Since the karyotypes of rats was not examined in this study, there is no definite proof to annotate *R. rattus* lineage IV and *R. tanezumi* of having diploid chromosome numbers of 40 and 42, respectively. Within the *R. rattus* complex, observations indicate that a pattern of chromosome polymorphism exists and this polymorphism is either derived from one of the karyomorphs of *R. rattus* or differentiated progressively from that of the black rat (Baverstock et al., 1983). So, the karyotyping of the rats within Peninsular Malaysia is a most helpful procedure to elucidate this hypothesis.

Commensalism and sympatricity in *Rattus* species might facilitate to hybridisation that leads to the closeness of genetic diversity between species. Bradley & Baker (2001) suggested that range of 2%-11% of genetic disparity in Cyt b sequences may stipulate a species level divergence for several mammal groups. Though, it is controversial to categorise specimens of the unresolved taxonomic class merely counting on a mtDNA genetic species concept (Ferguson, 2002). The remarkably low interspecific difference of 1.9% between *R. tiomanicus* and *R. baluensis* was significantly comparable to the

intraspecific discrepancy exists among the most distinct *R. rattus* lineage IV individuals (up to 2.1% K2P distance). This signifies that *R. baluensis* which is considered as an endemic species to Mount Kinabalu in Borneo (Musser & Carleton 2005), and *R. tiomanicus* which live in lower elevation habitats appraised as two distinct species due to the significant adaptations to the different geographical area. Furthermore, rats identified initially as *R. tanezumi* are highly related to *R. rattus* lineage I and III than those from lineage IV. It is suggested that the genetic exchange among closest descendant ought to adaptation and promoting to the emergence of new species (Hobolth et al., 2011; Seehausen, 2004).

The genetic distance of 3.9%-6.9% found between *R. rattus* and *R. tanezumi* is nearly similar to the evaluation of Tollenaere et al. (2010), Pagès et al. (2013), Yasuda et al. (2014), and Adhikari et al. (2018). Separation of both species occurred with an estimated divergence time of around 0.48 Mya, comparable with Robins et al. (2008). Robins et al. (2008) and Rowe et al. (2011) classified *R. tanezumi* and *R. rattus* as a sister taxon. Lack et al. (2012) observed the occurrence of hybridisation with introgression in *R. tanezumi* and *R. rattus* in the United States. Although there are limited morphological distinctive attributes that appear within *R. rattus* and *R. tanezumi*, this present work disproved the concluding notes that *R. tanezumi* are synonyms to *R. rattus*. Due to high nucleotide variations, the results support the status of valid and distinct species of *R. rattus* which successfully diverged from *R. tanezumi* within the genus *Rattus*. The genetic data rectified the taxonomic argument of *R. rattus* and *R. tanezumi* in Peninsular Malaysia and reaffirmed their presence as a sympatric associated species.

In furtherance, the topologies of NJ and BI trees were congruent in portraying the divergence of the *Rattus* species. The trees likewise supported the hypothesis that *R.*

*rattus* lineage IV is diverged according to the general phylogeography pattern. Several unique characteristics may arise within the rats in Southeast Asia that discriminated them from rats of other regions. Identically, the same tree topologies were observed in *R. baluensis* and *R. tiomanicus* where these two species served as a sister taxon. The *R. baluensis* samples arising in *R. tiomanicus* clade almost probably show introgression between them. The Ceylonese form of *R. rattus* occurs in Sri Lanka get exclusive recognition as those with chromosome diploid numbers of 38, 40, and hybrids of 39. Thus, it is fascinating to mention that the two nominal samples of *R. rattus* from Sri Lanka fall within the *R. rattus* lineages I and IV clades. This result is in agreements with Aplin et al. (2011) and Yasuda et al. (2014). The well-supported clades of *argentiventer*, *exulans*, and *norvegicus* are all compatible with the taxonomy. This study demonstrated that *R. norvegicus* were the earliest species to diverge and *R. tiomanicus*-*R. baluensis* were the latest species split among the ingroups. Robins et al. (2008) remarked the divergence of *R. norvegicus* as the earliest at approximately 2.9 Mya.

The divergence time estimation proposed that the diversification of *Rattus* initiated in the Pliocene. The abundance of species clustering patterns during the Pleistocene evidenced the ongoing evolution of various types of mammals. The warmer global climates and the extensive steppe vegetation promoted to habitat fragmentation, and inspired allopatric speciation (Knowles, 2001). Based on the time divergence estimation acquired from this study, it was validated that *R. rattus* strongly deviated from *R. tanezumi*. Furthermore, the *R. rattus* lineages I, III, and IV also diverged according to the theoretical phylogeographic pattern. These lineages populate allopatric distributions, with *R. rattus* lineage I being mainly native to Oceanian, *R. rattus* lineage III related to Himalayan, and *R. rattus* lineage IV belongs to the Southeast Asia. According to the island biogeography theory, the Isthmus of Kra played a pivotal role in the divergence of these

three *Rattus* lineages (Macarthur & Wilson, 1967). The border between the Indochinese and Sundaland regions is declared to be situated at the Isthmus of Kra in Peninsular Thailand (Lekagul & McNeely, 1988), as obvious cluster of mammals (Corbet & Hill, 1992) have been noticed between the two ends of this barrier. The results also highlighted a slight differences of diversification timepoint from the previous studies by Robins et al. (2008) and Aplin et al. (2011). This is probably due to the utilisation of single maternal marker of Cyt b in this study. Multiple genetic markers are necessary to obtain a strong timing estimation of the phylogenetic diversification in this recently evolved species. An integration of various calibration points at different time scales would be probable to notably improve the dating precision (Endicott et al., 2009; Ho & Larson, 2006).

This present study enhances and highlights the effectiveness of phylogenetic identification of *Rattus* species in Peninsular Malaysia using Cyt b sequences. Though, there is a crucial necessity to explore several groups of Malaysia's *Rattus* that possess cryptic speciation, introgression, and incomplete lineage sorting by the morphology, karyology, and molecular means. Further study is needed for developing comprehensive knowledge related to the distribution of genus *Rattus* in Malaysia and surrounding countries. Additional usage of molecular markers from several mitochondrial and nuclear genes are necessary to demonstrate the complex dispersal and profound taxonomic revision within the groups.

## 5.2 Population Structure of *Rattus rattus* in Peninsular Malaysia Based on Mitochondrial and Nuclear DNA Data

### 5.2.1 Mitochondrial and nuclear DNA as potential markers in elucidating the population structure of *Rattus rattus*

Nuclear DNA (nDNA) is a parental marker, inherited from both parents and enables for the genetic recognition of an individual. However, mitochondrial DNA (mtDNA) is a matrilineal marker, allows for the tracing of kinship through the female line (Avise, 1994). The fundamental gap between them are nuclear DNA is discovered inside the nucleus of the cell while mitochondrial DNA is discovered only in the mitochondria of the cell. Mitochondria are established to contribute significantly in various cellular processes and serves as the principal source of adenosine triphosphate (ATP) (Kujoth et al., 2005). By employing both mtDNA and nDNA as molecular markers, the genetic relationships and population structure of targeted groups can be inferred more precisely.

The conserved gene of cytochrome oxidase I (COI), cytochrome b (Cyt b), and hypervariable loci of displacement loop (D-loop) of mtDNA, together with nDNA of interphotoreceptor retinoid binding protein (IRBP) gene has proven to be effective in resolving the generic and specific level of *Rattus* from multiple geographic origins (Pagès et al., 2013, 2010; Aplin et al., 2011; Robins et al., 2007). Hence, these markers were further utilised in this study to compare and contrast the resolution and suitability at the intraspecific level. The application of various molecular markers may contribute precious information regarding the processes that frame the population structure of a species. Geographical (distance and barriers) and anthropological (habitat destruction, deforestation, and land conversion) influences may play a crucial role in separation of *R. rattus* populations. Instead of being threatened due to an inability to survive in the face of habitat destruction, *R. rattus* group adapt to exist in areas neighbouring human

residences where they are accessible to farms, gardens, as well as houses to forage for food (Drazo et al., 2008; Hill, 1997).

*Rattus rattus* has a very wide geographical distribution in Peninsular and continental Southeast Asia (Aplin et al., 2011). The biogeographical relationships between separated populations, in particular, those in the continental and insular, remain unclear for widespread *R. rattus* in Southeast Asia. Geographic isolation restricted the gene flow between populations prompt to population differentiation during the long evolutionary period. The oriental biogeographical regions that include the insular and mainland are a vital part of the mysterious history of faunal distribution. The Isthmus of Kra is considered to play an essential biogeographic division on the Thai-Malay Peninsula of mainland Southeast Asia. The borderline between the mainland and island is declared to be very unique, as prominent group of mammals (Corbet & Hill, 1992) have been observed between the two edges of this barrier.

Since *R. rattus* is widely dispersed, the population structure of *R. rattus* in Peninsular Malaysia is crucial for investigation to infer the relationships among the populations. This was the first study emphasising on the genetic structure of *R. rattus* in Peninsular Malaysia, which have not ever been explored, and the evolutionary status of this species are limited and urgent. In general, studies concerning population genetics of house rats are lacking globally, due to the status of the species which commonly known as pest and received less conservation attention. It was hypothesised that the population structure of *R. rattus* is undergoing population expansion. However, the patterns of segregation of the populations is further deduced to elucidate whether it is according to theoretical phylogeography pattern. Mainly known as vector for various zoonotic diseases, it is very

important to investigate the population structure of *R. rattus* in order to manage the dispersal of this entity.

### 5.2.2 Sequence and phylogenetic analysis

In this study, sequence analysis of *Rattus rattus* involved the concatenated mitochondrial DNA (mtDNA) of COI, Cyt b, and D-loop loci and a single nuclear DNA (nDNA) of IRBP. Both mtDNA and nDNA loci were highly conserved which yielded only 45 (1.86%) and 32 (2.61%) variable sites and parsimony informative characters across *R. rattus*, respectively. The result signified that mitochondrial DNA sequences are slightly having a higher rate of base substitution than nuclear DNA sequences. However, the results still remarked that both markers are sustained over *R. rattus*, and suitable to be used in discovering population structure of various species. The composition of Adenine (A) was high in mitochondrial DNA of *R. rattus* and Guanine (G) was high in nuclear DNA of *R. rattus*. These showed that the two-carbon nitrogen ring bases of purines (adenine and guanine) are dominant in the sequences of *R. rattus*.

The pairwise genetic distances between populations of *R. rattus* in Peninsular Malaysia ranged from 0.2%-1.2%, based on both mtDNA and nDNA sequences, which proposed a pattern of homogeneity among *R. rattus* populations. Also, pairwise genetic distance analysis indicated populations of Penang island and Terengganu were the most distant from other populations (1.1%-1.2%). The geographical distance of both populations is only 321 KM. However, these two populations reflected the lowest connection compared to other populations. The lowest genetic distance within populations are shown by Penang mainland-Perak (0.2%; 125 KM) for mtDNA sequences, and Kelantan-Terengganu *R. rattus* (0.5%; 131 KM) for nDNA sequences. Both Penang mainland and Perak are adjacent cities, similar to Kelantan and Terengganu. Based on the sampling localities

involved in this study, Penang mainland-Penang island had the least geographical distances of only ten KM. However, the pairwise genetic distance between these two populations are in the range of 0.5%-1.0%, as shown by both markers.

The results of pairwise genetic distances are highly expected, considering the vicariance theory on population disjunction as the central argument. In this situation, the island population and mainland population are divided by the Penang Strait; connected with the Strait of Malacca, to the north and south. After *R. rattus* expanded throughout Peninsular Malaysia, gene flow between the insular and other continents was cut-off over time. This might be due to the sea level rise in the interglacial period and caused insular populations to become isolated from those of other regions. This division probably induced the restriction of gene migration between both populations, triggering them to accrue the major genetic differences discovered in this study. Yet, only a single island population was observed in this study. According to Sun et al. (1999), geographically close habitats can be distinct, while geographical distant habitats can be similar in their environmental surroundings. In addition, the pairwise genetic distance between *Leopoldamys sabanus* with *R. rattus* is higher, thus resulted in distant genetic relationships, therefore, this species is suitable to be the outgroup in this study.

Mitochondrial DNA sequences revealed the highest nucleotide diversity ($\pi$) and net nucleotide divergence (*Da*) between populations of Penang island-Pahang. The lowest values of $\pi$ and *Da* were found consistent between Penang mainland-Perak populations. In contrast, nuclear DNA sequences revealed the highest $\pi$ and *Da* between populations of Penang island-Selangor. The lowest values of $\pi$ and *Da* were found consistent between Kelantan-Terengganu populations. Low level of genetic divergence was observed between populations of Penang mainland-Perak and Kelantan-Terengganu. Still,

population of *R. rattus* from Penang island with highest *Da* symbolised a gap with other populations of mainland, suggesting high level of genetic divergence.

Haplotype analysis based on mtDNA sequences showed that only a single haplotype is common representing northern populations of *R. rattus* from Kedah, Penang mainland, and Perak. For nDNA sequences, the shared haplotypes are reflecting populations of *R. rattus* from northern region of Perlis-Kedah, Perlis-Penang mainland, Penang mainland-Perak, southern region of Melaka-Johor, and east coast region of Kelantan-Terengganu. The phenomenon of local haplotypes also demonstrating the connection between the adjacent cities, where the closer populations are linked to each other. A lower genetic distance among populations could explain a higher genetic information sharing in the populations of *R. rattus* in Peninsular Malaysia. Still, population of Penang island was an exception herein. Overall, haplotype analysis presented that the *R. rattus* samples from each population are highly endemic based on the mtDNA and nDNA sequences, as most of the observed haplotypes were unique. An isolated population or populations with fewer gene flow contributed to the uniqueness of the haplotypes (Aboim et al., 2005). A relationship of one local haplotype with others in lower frequencies or unique haplotypes is a trend usually assigned to populations that have experienced recent range expansion (Rogers & Harpending, 1992; Slatkin & Hudson, 1991).

MSN analysis clearly illustrated the relationship and connectivity between the entire populations of *R. rattus* in Peninsular Malaysia. However, the linkage of the populations was mixed and no obvious geographic cluster of *R. rattus* populations in Peninsular Malaysia was shown, as all the populations were not separated to the phylogeographical patterns (i.e. northern-southern, east-west). This is precedented, as there were no major barriers identified across the mainland populations of Peninsular Malaysia. The

separation of *R. rattus* population of Penang island was also weak and unresolved through the MSN analysis.

Phylogenetically, both markers potrayed the same projections of NJ tree. According to the tree topologies, all *R. rattus* samples were excluded from the outgroup. The divergence of *Leopoldamys sabanus* from *R. rattus* were sufficient because *L. sabanus* came from the different genus of *Leopoldamys*. In the ingroups, the homogenous patterns of population structure were observed. The segregation of *R. rattus* populations were not emphasising on the island and mainland separation. The distribution of the haplotypes in the tree topologies also generally did not reflects the geographic separation of the samples.

### 5.2.3 Population genetics analysis

The genetic differentiation (*Nst*, *Fst*, and *Nm*) for populations of *R. rattus* in Peninsular Malaysia revealed the highest *Nst* and *Fst* value between Penang island-Pahang populations, based on both mtDNA and nDNA sequences. The lowest *Nst* and *Fst* value were discovered between populations of Kuala Lumpur-Melaka and Penang mainland-Perak for mtDNA and nDNA sequences, respectively. A higher *Fst* indicates the population might have undergone reproductive isolation and lower *Fst* is due to the exact sharing of genetic materials between populations. Here, the gene flow (*Nm*) justified both *Fst* and *Nst* analysis, in which Penang island-Pahang demonstrated the lowest *Nm*, suggesting a nearly cut-off gene flow involving both populations. The analysis of genetic differentiation on Peninsular Malaysia's *R. rattus* disclosed a higher separation of mainland and island populations ($Fst > 0.25$). However, the observed findings were probably due to the barrier of island-mainland (i.e. Penang Strait) and the isolation-by-distance (IBD) model were disproved in this situation. The outcomes of *Nst*, *Fst*, and *Nm* supported a closer relationship between mainland populations; e.g. Penang mainland-

Perak and Kuala Lumpur-Melaka, showing that populations of *R. rattus* are having a wide dispersal throughout mainland of Peninsular Malaysia. High dispersal abilities with restricted physical barriers might led to the high degree of genetic exchange. Thus, populations of *R. rattus* in mainland may include a panmictic population with relatively similar genetic identities.

The genetic population structure of many species had been specified by a pattern of IBD (Ansmann et al., 2012; Hoelzel et al., 2007; Natoli et al., 2005). But, the results from Mantel test rejected the IBD model for populations of *R. rattus* in Peninsular Malaysia with $P > 0.05$. There was no significant correlation observed between genetic differentiation and geographic distance among populations of *R. rattus*. For example, based on distance observation, Penang mainland is the nearest city to Penang island (geographical distance of ten KM). Yet, a natural barrier limiting the gene flow between the two populations although both cities are located in the adjacent of northern part of Peninsular Malaysia. Hypothetically, the populations may come from the same centre of origins, but the dispersal mechanism enables them to disperse (some may disperse to great distances). In some cases, they are still in viable range that enable them to reproduces and contribute to gene flow of populations, but in some other cases they disperse too far until disjunction of populations occur and cut-off the gene flow.

The population genetic structure ascribed by the fixation indices ($F_{ST}$) of 0.55 was observed in the populations of *R. rattus* from Peninsular Malaysia, based on mtDNA sequences. However, nDNA sequences only revealed the $F_{ST}$ value of 0.19 for the variation among populations. A high genetic differentiation is secured with the $F_{ST}$ value higher than 0.25, due to the genetic structure. In this context, the result from mtDNA sequences supported a high genetic differentiation among populations of *R. rattus* in

Peninsular Malaysia, with significant levels of genetic structure, however, the result from nDNA sequences signalling a significant moderate genetic differentiation and was not in concordance with the result of mtDNA sequences. The inconsistent findings were due to the information provided from the different molecular markers, as the mtDNA marker revealed the genetic structure from maternal lineage while nDNA marker disclosed the genetic structure from parental lineage. A high mutation and fast evolutionary rates and lack of recombination in the mtDNA marker versus those of the nuclear marker might led to these findings. Also, distinct genes and markers are subjected to distinct selective pressures, which may initiate to the inconsistency of population structure.

Overall, the entire populations of *R. rattus* from Peninsular Malaysia reported high level of genetic diversity. It assists as a way for populations to adapt to changing environments. Island populations usually exhibit lower genetic diversity compared to their equivalent mainland populations (Cao et al., 2012; Takahashi et al., 2007). The findings from mtDNA sequences utilised in this study disclosed that the genetic diversity (nucleotide and haplotype diversities) of *R. rattus* were the lowest in the population of Penang island compared to the mainland which could indicate a bottleneck or small founder size (Tajima's $D$ = 0.70767; Fu's $Fs$ = 1.420) following colonisation from the mainland. All six samples were analysed within Penang island, and low level of diversity was noticed from mitochondrial haplotypes which signified that Penang island populations maybe deviated from the mainland populations via inbreeding. Inbreeding triggers homozygosity, which can increase the probabilities of offspring being affected by deleterious or recessive traits, as well as reducing disease resistance. The role of the landscape in structuring populations has been emphasised in many studies, especially, the extent to which islands offer opportunities for isolation and the consistency of such an effect across lineages.

Additionally, the historic demographic analysis supported with the neutrality test illustrated a unimodal bell-shaped curve for entire populations of *R. rattus* in Peninsular Malaysia, signaturing that the populations is undergoing an expansion model. Fu & Li's *D** and *F** tests introduced significant deviation from neutrality for populations of *R. rattus* in Peninsular Malaysia, thus stipulating the possible occurrence of background selection. A positive Tajima's *D* indicates low levels of both low and high frequency polymorphisms in the populations of *R. rattus*, while a negative Fu's *Fs* was an indication of an excess number of alleles, as would be anticipated from a recent population expansion of *R. rattus*. When neutrality test was conducted on the entire populations of *R. rattus* in Peninsular Malaysia, a high negative Fu's *Fs* index and a significant Tajima's *D* value were found, providing a strong evidence of population expansion, and eliminate the probability of genetic hitchhiking and background selection, and evolutionary force that generate a model comparable to population expansion (Okello et al., 2005; Fu, 1997; Fu & Li, 1993). The star-like model in MSN also indicated a population expansion pattern (Slatkin & Hudson, 1991). The sum of square deviation (*SSD*) (Schneider & Excoffier, 1999) and the Harpending's raggedness index (Harpending, 1994) were employed to test the hypothesis that the observed data fit the sudden expansion model. The outcomes of non-significant values for *SSD* signified that the observed data do not deviate from that expected under the model of expansion. Non-significant raggedness index also stipulates the existence of population expansion event. The remarks of non-significant value in goodness-of-fit distribution for entire populations of *R. rattus* propose that population expansion took place recently (Rogers, 1995). A recent population expansion was also consistent with small differences detected between population before expansion ($\theta_0$) and population after expansion ($\theta_1$). With the absence of mutation rate in both mtDNA and nDNA genes of *R. rattus*, the expansion time cannot be estimated in this study.

*Rattus rattus* from Penang island were significantly differentiated from those in its adjacent mainland. Yet, there is no species complexity discovered between populations of *R. rattus* in Peninsular Malaysia, as the pairwise genetic distances supported the close relationships between entire populations. Based on several population analyses in this study, it was suggested that house rat of Penang island possessed unique genetic characteristics. It has been hypothesised that an isolated island and their populations might have unique sequences, allowing adaptation of rodents between its environment. The unique haplotypes detected in the populations (particularly the island population) would be a good indicator to conserve the genetic uniqueness of *R. rattus* in Penang island. Yet, additional samples from other island regions of Malay Peninsula are required to validate the segregation of island-mainland populations, in which might disclose the theory of island biogeography.

The outcomes of this study are directly pertinent to the management approaches of *R. rattus* in Peninsular Malaysia, mainly in terms of the eradication process. The human-rodent conflict has worsened in Malaysia due to habitat degradation and habituation of the species. It is crucial to recognise the genetic diversity between populations of *R.* rattus, to enable the management of the unique evolutionary lineages of the species. So, the target populations of *R. rattus* can be managed within the similar gene pool of the other populations (i.e. the mainland-island). This management is vital to prevent the loss of genetic diversity that cannot be recovered, or, in extreme cases, can lead to population hybridisation and potential outbreeding depression of the populations (DeSalle & Amato, 2004). These findings would also be valuable to other field of studies; such as, on the co-evolution of rodents as vector of diseases (i.e. evolution of leptospirosis) and biogeographical history of the Malay Peninsula.

Overall discoveries augmented a better insight on the population structure of *R. rattus* in Peninsular Malaysia and highlighted several important sampling localities for further observation which will contribute significantly to ongoing studies. The mitochondrial markers of COI, Cyt b, D-loop and nuclear marker of IRBP were shown to be suitable in resolving intraspecies relationships and served as reliable markers in elucidating the population structure of *R. rattus*. Yet, further conclusions can only be made with extensive sampling supported by an increase number of individuals from each locality as well as from a wider geographical area. An additional of molecular markers is needed to extend the present understanding.

## 5.3 Morphometric Analysis of Skull Characteristics of *Rattus rattus*

Conventional morphometrics approach enabled the description of the sizes of skull variability in *Rattus rattus* populations in Peninsular Malaysia. The principal components are created such that the PC1 accounts for the maximum variation, PC2 for as much as the remaining, and so forth. Based on the PCA plots, the PC1 scores were fairly negative indicating that the first principal component score increases with the decrease in all the 14 variables studied. This implies that the high PC1 score is associated with smaller measurements of all the 14 measurements studied. The eigenvalues for PC2 and PC3 were much smaller compared to those of PC1 showing that PC1 accounts for majority of the craniodental measurements considered. The highest negative correlation in HBC with PC2 reflects that the size of HBC decreases as PC2 increases and the decrease in HBC has the greatest influence in the variability of PC2. For PC3, the increase in the size of ML and M1-M3 is responsible for the variation in PC3. The similar interpretation is applicable for the case of the female *R. rattus* involving HBC for the variation of PC2 and PC3 (highest negative correlation) as well as LD and LBP for the variation of PC2 (highest positive correlation).

Variability was evident with the variation in craniodental size among three clades associated to age, even though the present samples did not show skulls with age classes belonging to stages C0, C1, and C5. It is difficult to trap individuals of ages C0 and C1 as individuals of these age groups are at infant stages, have restricted movements, and are still dependent on their mothers. On the other hand, individuals of age C5 belong to the very elderly group where most tend to be sickly and die. Three age groups i.e. C2, C3, and C4 related to the three clades are evident in the PCA and LDA plots. The outcomes imply statistically significant age-related variations in craniodental morphology in *R. rattus*. The existence of three clusters which represent the age groups was seen in the male and female specimens with ages C2 and C3 sharing some similarities with C4 portraying extensive variability compared to the aforesaid age groups. There were gradual changes from C2 and C3, to C4, proving that the sizes are increasing during the growth process. Nonetheless, every tooth-wear class is morphologically different in shape and size. Moreover, within *R. rattus*, individuals of tooth-wear classes C2 and C3 are presumably subadults while the ones with tooth-wear class C4 may represent adults. *Rattus rattus* belonging to age groups C2 and C3 shared some similarity in craniodental measurements and it was anticipated that there would not be much difference in terms of size between these age groups. The significant difference could be seen clearly in age group C4 because this was the adult group. The dentition of *R. rattus* is monophyodont where they only possess one generation of teeth throughout their entire life. Their molars are never replaced. Moreover, it has been described that distinct age classes have different fundamental developmental attributes (Cesaroni et al., 1997). Thus, natural selection performed upon specific traits of the skull may differ, relying on the age at which it occurs (Zelditch et al., 1992).

The presence of sexual dimorphism was also obviously recognised. In this study, the skull sizes of male *R. rattus* is slightly bigger compared to the females. It is in accordance with several studies by Abdel-Rahman et al. (2008) and Martínez et al. (2014), which demonstrated that some rodents may show sexual dimorphism. However, previous published studies on morphometric variation of *R. rattus* through conventional morphometrics demonstrated the absence of sexual dimorphism and age (ontogenetic) variation in both populations of Congreso Island and Tunisia (Ben Faleh et al., 2012; Ventura & López-Fuster, 2000). The absence of huge variability between populations based on ontogenetic and sexual dimorphism was also demonstrated in various rodents such as *Jaculus jaculus* (Ben Faleh et al., 2013, 2010), *Mastomys natalensis* (Lalis et al., 2009), and *Malacomys edwardsi* (Bohoussou et al., 2014).

The utilisation of conventional morphometric approaches in evaluating sexual dimorphism and ontogenetic variation in the current study demonstrated statistically significant variations among craniodental sizes of *R. rattus*. These outcomes are similar to that of Abdel-Rahman (2005). This investigation proposes that distinctive growth patterns within the sexes may prompt to diverse phenotypic patterns throughout the ontogeny of *R. rattus*. This recommends that sexual dimorphism in craniodental size could be due to natural selection influencing separately in the two sexes during growth development leading to contrary outcomes (Bronson, 1989). Behaviours do have an influence on sexual dimorphism in rats (Wickens, 1998; Zelditch et al., 1992). The reason for males to be larger in size is yet to be known due to the lack of common ecological inspections on *R. rattus* in Peninsular Malaysia.

As hypothesised by Lalis et al. (2009), the variation in the skull features might be due to the local adaptation to different environments. *Rattus rattus* in Peninsular Malaysia

may have encountered a rapid morphological change because of restricted migration of the species. Different local environmental pressures could have resulted in morphological differentiation within *R. rattus*. It is speculated that the existing skull variation within *R. rattus* in Peninsular Malaysia is a result of competition for food and space as well as predation. It is probable that the local variations of the habitats may have also led to skull variation in *R. rattus* especially rapid urbanisation and industrialisation.

Also, the remarkable skull variation in the populations of *R. rattus* is caused by diet preferences throughout the habitats. Previous studies on diet response of *R. rattus* in distinctive conditions showed that black rats are very demanding feeders although they feed on various animals and plants (Clark, 1982, 1981). The diet preference of *R. rattus* also varies substantially due to habitat and seasonal changes (Clark, 1982). Various investigations have demonstrated the presence of variability in the size and shape of the skull at the population level in mammals (Lalis et al., 2009; Polly, 2007).

However, no evidence of geographic variation related to the skull variability was illustrated in the populations of *R. rattus* from Peninsular Malaysia. The trends of variability were consistent without concerning to the geographical factors (i.e. latitude, longitude, and precipitation). Previous studies emphasised that size variation is commonly evaluated to be easily altered to environmental slopes (Patton & Brylski, 1987; Thorpe, 1976). Besides, few recent studies examined the real pattern of size variation related with a latitudinal factor as formerly stated for other rodents (Ben Faleh et al., 2016, 2012; Martínez & Di Cola, 2011; Macholán et al., 2008; Reis et al., 2002). However, this present study only covered limited areas in Peninsular Malaysia, and it is possible that the dispersal of *R. rattus* is unlimited beyond their current range limits because of no specific physical geographic barriers were discovered. As geographically known, the

climate in Peninsular Malaysia is equatorial, being hot and humid throughout the year. It is known that Peninsular Malaysia does not have a strong climatic gradient. Thus, as supported by the results from this study, common geographical factors such as latitude, longitude, and precipitation do not have an impact on the craniodental measurements of *R. rattus*.

*Rattus rattus*' survival strongly depends on various conditions but not much is notable about their ecology or exact habitats in localities. However, it cannot be examined here due to the lack of common ecological inspections on *R. rattus* in Peninsular Malaysia. Different aspects of intraspecific variability continue to be examined in this species, yet, the small samples sizes restricted the geographical elucidations of the observed findings. This present study could potentially be a yardstick if comparison of specimens can be done with several countries in Southeast Asia since no detailed morphometric study has been done to date, to establish a detailed morphological variation in *R. rattus* throughout the regions, possibly with an evidence of geographic variation in future studies.

### 5.4 Prevalence and Genetic Divergence of *Leptospira* spp. in *Rattus rattus*

Leptospires are globally distributed, however tropical and subtropical regions as Malaysia attracts the bacteria for longest survival (Levett, 2001). Due to the diversity of the bacteria living in different environments, human and animals are exposed to a continuous risk of infection. From this study, the discovery of pathogenic *Leptospira* spp. in rats was predictable as a previous study by Noguchi (1917) first documented the isolation of this spirochete from rats. The findings of *L. interrogans* and *L. borgpetersenii* infecting rats are in accordance with previous work by Benacer et al. (2016). The two species pose a serious health risk as both of them are commonly known as the main causative agent of leptospirosis globally, which can lead to fatality (Levett et al., 2006).

The identification of disease reservoirs is paramount in assisting the planning of effective prevention and control measures, limiting the proliferation rate of the spirochetes in the environment, as well as minimising the exposure risk of human comes into contact with the infected animals.

Molecular approach, in particularly PCR, is regards as a potent technique in detection of leptospiral infection in rats as demonstrated in this study. The incorporation of both 16S rRNA and LipL32 protein genes in the screening not only detect the presence of the bacteria, but also discriminate between pathogenic and non-pathogenic *Leptospira* spp. In addition, by knowing the nature of leptospiral colonisation in an infected host, selective screening on both liver and kidney tissues reveals the disease's development stage, be it early or late infection, respectively. So, PCR holds as a rapid tool for diagnosis.

Of the 51 rats recognised as leptospirosis positive, it can be seen that the prevalence of leptospires in subadult (39.1%) and adult (39.3%) rats do not differ significantly. The aggressiveness of both age groups in fighting for resources may had led to the transmission and circulation of *Leptospira* spp. among them. However, the result from this study discovered that the infection is more prevalent in male rats compared to the females. This finding is in agreement with earlier study by Mosallanejad et al. (2013) who discovered higher prevalence of *Leptospira* spp. in male compared to female rats. This is may be due to the different behaviour and lifestyle of both male and female rats, and are due entirely to exposure-related bias.

Based on the observation of *Leptospira* spp. prevalence in 13 cities in Peninsular Malaysia, rats of Penang island revealed a higher rate of infection. The island which is famous as a tourist spot attracts the rats to forage due to the disposal and accumulation of

garbage and rubbish by the public everywhere near the seaside. The area was considered as less hygienic during the sampling session, and the rats have easy access to the contaminated environment and stagnant water (Pui et al., 2017). Irregular sanitation greatly stimulated the spreading of *Leptospira* spp. (Benacer et al., 2013). On the other hand, rats caught from Terengganu disclosed a lower infection rate as compared to other cities. A proper waste management and sanitation were visible near the sampling niches at the housing area.

In furtherance, the sampling niches of fresh markets, housing areas, and seaside disclosed the pervasive distribution of rats within the areas. The leftovers and the irregular garbage management attracted the rats and accelerated their breeding and growth. The rapid urbanisation and the emerging of a run-down area with insufficient facilities and poor sanitation bring the dispersion of rats closer to human and enhance the possibility of dissemination of rodent-borne diseases. The results indicated that the rats caught in fresh markets were highly exposed to the leptospiral infection. The condition of the fresh markets with abundant resources for the rats to scavenge, the improper maintenance of the drainage system congested with rubbish, and the wet containers left exposed create a risk of feasible contamination with rat's urine. The moist surroundings promote the survival of pathogenic leptospires outside the hosts, aiding as a catalyst of leptospiral infection in both humans and animals.

*Rattus rattus* are mostly asymptomatic during the leptospiral infection and predominantly harbour the pathogen within their renal tubules. These infected animals continue to shed the bacteria into the environment through urination. The results of this study demonstrated a higher occurrence of the pathogen within the kidney tissues compared to the liver tissues. An assumption can be made herein, where kidney is the

most favourable place which promotes continuous colonisation of *Leptospira* spp. A study by Gomes-Solecki et al. (2017) proposed that the proximal tubule of the kidney represents a secure place for *Leptospira* spp. to stay as they are shielded from the activity of the immune system. The incident of the leptospires in the liver tissues are most likely to happen during the early stage of bacterial infection as the bacteria spread comprehensively to nearly all tissues throughout the body (Athanazio et al., 2008). However, with the production of antibody within days, the pathogens are rapidly eliminated from most tissues with the help of the immune defence, except the kidney (Athanazio et al., 2008; Monahan et al., 2008).

The phylogenetic reconstructions of 16S rRNA gene revealed that the genetic make-up of all the three groups of *Leptospira* spp. (pathogenic, intermediate, and saprophytic) are unique and discriminative among them. The results of genetic distance also support the separation of the groups. In addition, the genetic constitution between both pathogenic species of *L. interrogans* and *L. borgpetersenii* disclosed a huge species variation. The comparisons of genetic distance between *L. interrogans* and *L. borgpetersenii* remarked LipL32 gene as the best tool for the recognition of pathogenic *Leptospira* spp. due to greater divergence compared to 16S rRNA gene.

The results of genetic variability of the two closely related species of *L. interrogans* and *L. borgpetersenii* indicated that several genetic variations existed within and between species, which further disclosed significant levels of genetic structure. It is surprising that the result of pathogenic *Leptospira* spp. identified from rats disclosed a strong relationship with *Leptospira* spp. spotted from another host, such as *Sus scrofa* (Rettinger et al., 2012). This indicates that the *Leptospira* spp. inhabiting animals, humans, and environment were genetically identical and have a broad host range. To date, there is no

proof that certain genetics composition only response to a particular host and reservoir (Sumanta et al., 2015).

Notably, the amplification of 16S rRNA gene highlighted the individuals of *L. interrogans* shows no variation among them, while individuals of *L. borgpetersenii* shows huge variation. This indicates that *L. borgpetersenii* is highly mutated and easily evolved, contributing to the sequence variations as compared to *L. interrogans*. However, through the amplification of LipL32 gene, the variation among individuals of *L. borgpetersenii* become less prominent and comparable with *L. interrogans*. The conserved region of LipL32 gene demonstrated a slower mutation and evolutionary rates. The findings from this study also implied that gene flows between *L. interrogans* and *L. borgpetersenii* were limited, which show that the segment of DNA of the bacteria is not transferred between them, engenders to a unique identity of both species. The genetic analysis of both genes demonstrated a high level of genetic divergence between the two species (*L. interrogans* and *L. borgpetersenii*) studied. Ancestral population is expected to demonstrate a higher degree of haplotype and nucleotide diversity (Nei & Tajima, 1981; Nei, 1973). Throughout this study, higher degree of both haplotype and nucleotide diversities were found in *L. borgpetersenii*, proposing that *L. borgpetersenii* is an older ancestral species while *L. interrogans* is the more recent species.

## 5.5 Transcriptome Study of *Rattus rattus* Infected with *Leptospira* spp.

Humans or animals infected with pathogenic *Leptospira* spp. exhibit a variety of clinical signs ranging from asymptomatic to severe multiple organ failure and fatal diseases. It is not yet known the mechanisms by which leptospires cause damage to the host tissue or, conversely, the pathways used by infected hosts to prevent organ failure. Even though presently, using molecular biology techniques, leptospires can be classified

into species, yet, classification using serology is still frequently used. More than 300 pathogenic serovars are reported and it is still being questioned whether different serovars can cause different severity of disease. Several considerations such as virulence, number of organisms, and host factors may be involved to enable *Leptospira* spp. in triggering different degrees of symptoms (Silva et al., 2008).

The kidney is a primary target of leptospires in which renal tubular conditions favour leptospiral survival during both acute and chronic infections (Yang et al., 2001; Faine et al., 1999). In the previous work by Athanazio et al. (2008), after the rat (*Rattus norvegicus*) was experimentally infected with *L. interrogans* serovar Copenhageni, a wide circulation of pathogen to nearly all tissues is detected in the early phases of infection. This is proceeded by clearance of leptospires from most tissues within days, with the exception of selective survival and propagation of leptospires in the kidneys, followed by leptospiruria on the seventh day. This clearance from most tissues is probably aided by circulating anti-leptospiral immunoglobulin (IgM and IgG), which is noticeable in rats seven days after infection, strengthen the justification that the kidney is therefore to be somewhat immune privileged and thus facilitates persistent leptospiral colonisation (Athanazio et al., 2008; Monahan et al., 2008; Faine et al., 1999). Since leptospires firstly circulate to all tissues and not specifically the kidney, this does not support the postulation that the kidneys are specifically targeted by leptopires due to tropism. It is difficult to ascertain from the literature the mechanisms that facilitate this privileged immune state.

Pro-inflammatory cytokines comprise interleukins (IL)-1β, IL-6, interferons (IFNs), and tumor necrosis factors (TNFs), as well as chemokines, serve as chemo attractants to assemble leukocytes to the area of tissue damage or infection (Turner et al., 2014; Tisoncik et al., 2012). As lately revised by Werts (2017), data display that leptospiral

PAMPs activate innate immunity via multiple PRRs to produce cytokines and initiate the inflammatory cascade. The immune response to an infection is now considered rather than simply initiated, to be precisely regulated. Expression levels of cytokines are mostly studied in the context of septic shock and severe sepsis, both for their potential prognostic value (Bozza et al., 2007; Sfeir et al., 2001; Arnalich et al., 2000; van Dissel et al., 1998) and as a means to expand understanding of host-pathogen interactions. Mechanisms underlying the specific inflammation activated by leptospiral PAMPs/PRR have not yet been completely clarified. Fascinatingly, unlike other bacterial lipopolysaccharide (LPS) (i.e. *Escherichia coli*) that stimulates the TLR4 signalling pathway classically, Nahori et al. (2005) identified leptospiral LPS in mice by both TLR2 and TLR4. In addition, Gomes-Solecki et al. (2017) revealed that TLR4-deficient mice demonstrate clinical signs of severe leptospirosis, demonstrating that murine TLR4 protects mice against leptospirosis. This species specificity in the sensing of leptospiral PAMPs also proposes that hosts may activate various inflammatory responses based on their susceptibility to leptospirosis. The TLR signalling pathway demonstrated by the cytokines and chemokines observed in this study probably led to the resistance of *Rattus rattus* towards leptospirosis.

Various studies demonstrated the gene expression of hosts experimentally infected with *Leptospira* spp. The present study, in contrast, investigated the interactions between *Leptospira* spp. and wild rats, specifically *R. rattus* which were naturally infected with the bacteria. This gave an opportunity to compare, via transcriptome analysis, the gene expression in rats that continue to survive the leptospiral infection with that of healthy rats, to improve an insight of the host-bacteria interactions. Several genes were discovered to be differentially expressed in the leptospiral-infected rats compared to the uninfected controls. Still, the differentially expressed genes (DEGs) demonstrated variation in the

expression level depending on the organs involved. From the transcriptome results, a higher number of host genes were significantly up-regulated and a lower number of genes were significantly down-regulated by infection with leptospirosis in the kidney tissue, based on significant differences in the expression of relative transcript abundance. In contrast, liver tissue revealed the highest amount of down-regulated genes and the lowest amount of up-regulated genes during the bacterial infection. In the present study, none of the rats got infected in both kidney and liver tissues simultaneously. The leptospiral DNA presented in either kidney or liver of an individual at each time point. However, most of the livers were found to be clear of the bacteria compared to the kidneys. This finding confirmed the preference of leptospires to endure for a longer time in the kidney and supported the clearance of bacteria in the liver at the early phase of leptospiral infection. Although the quantity and time of bacterial infection toward the rats was not counted and lack in this study, the wild rats act as an asymptomatic carrier of the bacteria without any sign of illness, which shows a high level of tolerance for the disease.

Most immune response study on host-bacteria interaction only focuses on the major organ of kidney for immunity (Ferrer et al., 2014; da Silva et al., 2012; Lowanitchapat et al., 2010; Vielhauer et al., 2004). Limited studies have concentrated on gene expression in the liver of animal models of leptospirosis. In the current study, the liver tissue was also included together with the kidney tissue to gain a broad overview of leptospirosis infection in the house rat. Moreover, six immune-related genes were selected to elucidate the host-bacteria interaction in house rat infected with *Leptospira* spp. by transcriptome, providing a comprehensive summary of the cascade of genes respond towards bacterial infection. The six differentially expressed *R. rattus* immune genes; tumor necrosis factor-α (TNF-α), interleukin (IL)-1β, IL-6, IL-10, macrophage inflammatory protein-1α (MIP-1α/C-C-type chemokine ligand 3, CCL3), and gamma interferon (IFN-γ)-inducible

protein 10 kDa (IP-10/C-X-C-type chemokine ligand 10, CXCL10) were explored in the wild rats naturally infected with leptospires. Those genes were greatly observed and reported in many previous studies on various hosts experimentally infected with *Leptospira* spp. Thus, in the present study, by incorporating multiple organs that are involved in the leptospiral infection, the expression and progress of host cell defence against *Leptospira* spp. could be observed.

In the response to leptospirosis, various cytokine productions or gene expressions have been shown. In the present study, cytokines and chemokines gene expression were examined in the livers and kidneys within the carrier state of rats, because kidney and liver are the organs regularly revealed to be injured in leptospirosis. The results showed that pro-inflammatory IL-1β, IL-6, TNF-α, MIP-1α, IP-10, and anti-inflammatory of IL-10 were up-regulated in kidneys. Yet, expression levels of IL-1β, TNF-α, MIP-1α, IP-10, and IL-10 but not IL-6 were significantly different in livers during the carrier state. Only IL-6 was up-regulated during the infection in the livers. The activation of the inflammatory response, particularly by the assembly of cytokines, is vital for the primary destruction of bacteria during infection. Yet, Cagliero et al. (2018) reported that unregulated cytokine secretion can cause a cytokine storm process, which can promote to sepsis and organ damages.

The results demonstrated the upregulation of IL-1β in the kidneys of *R. rattus* and downregulation of IL-1β in the livers during leptospires infection. It is in agreement with the results from Chou et al. (2018) which signifies the upregulation of IL-1β in damaged kidneys of mice promoted by infection of pathogenic *Leptospira* spp. IL-1β stimulates the initiation of immune cells as macrophages and subsequent release of immuno-regulators as well as pro-inflammatory factors that boost the inflammatory response. High

circulation level of IL-1β is reviewed to be the indicators of a cytokine storm and its role in sepsis pathophysiology has been extensively studied. These information mark IL-1β as the main conductor of the wrong inflammatory response and the sepsis discovered in severe leptospirosis after cytokine storm-induced. Besides, both kidneys and livers of *R. rattus* observed in this study highly expressed IL-6 during leptospiral infection. Induced by TNF-α and IL-1β, pro-inflammatory IL-6 is produced by a huge number of immune and non-immune cells. High IL-6 concentration is a sign of septic shock and associates with severity of leptospirosis (Papa & Kotrotsiou, 2015; Reis et al., 2013; Schulte et al., 2013). IL-6 can also play an uncertain role in leptospirosis immunopathogenesis.

In this present study, the expression of inflammatory TNF-α was also quantified in both kidneys and livers, however, a difference was observed between organs. TNF-α was upregulated in the kidneys and downregulated in the livers. The result obtained in this work was in accordance with the study by Lowanitchapat et al. (2010), which demonstrated an upregulation of TNF-α in renal of hamsters induced with *L. interrogans* serovar Pyrogenes. In the kidneys of three mouse models studied by da Silva et al. (2012), inflammatory TNF-α was examined to challenge renal modifications following infection with leptospirosis. Transcript level after the early days of infection was disclosed to be upregulated and maintained for up to one week after inoculation. According to Athanazio et al. (2008), when challenging leptospiral infection, mice deficient in TNF-α receptor production displayed more renal lesions signifying a defensive effect of this cytokine during chronic leptospirosis. While TNF-α was the first cytokine related with severity of leptospirosis (Tajiki & Salomão, 1996; Estavoyer et al., 1991), its exact role in the pathophysiology of the disease is still under discussion.

This study also inspected the regulation of chemokines MIP-1α/CCL3 and IP-10/CXCL-10, and significant differences in the transcript level were described in both rat kidneys and livers. MIP-1α was upregulated in the kidneys while in rat livers it was significantly downregulated. Earlier studies by Fujita et al. (2015) demonstrated overexpression of MIP-1α in the hamster kidneys after four days of infection with a virulent leptospiral strain. Hence, expression of MIP-1α in the rat kidneys observed in this study could clarify the involvement of this cytokine in rodent leptospirosis. Moreover, the level of IP-10 transcript was earlier quantified in hamster kidneys infected with virulent leptospires, and an increase in gene expression was presented during the acute infection stage (Fujita et al., 2015; Lowanitchapat et al., 2010). In this present study, during the carrier state, IP-10 was also upregulated in the rat kidneys, with downregulation in the livers; which reaffirmed the earlier note that IP-10 was involved in leptospirosis.

Anti-inflammatory effectors serve as a vital role to counterbalance the inflammatory cytokine effects. IL-10 with pleiotropic effects in immunoregulation and inflammation is typically known as one of the anti-inflammatory cytokines. IL-10 is a main innate immunity regulator, reviewed to contribute to septic shock immunosuppression (Mege et al., 2006; Sfeir et al., 2001). This cytokine also defends against deleterious tissue injuries due to high-level of cytokine release (Moore et al., 2001; Gesser et al., 1997). In this study, high expression of IL-10 was noticed in rat kidneys, however, IL-10 was down-regulated in the livers. Zhang et al. (2013) demonstrated a low-level of IL-10 expression results in a worsen glomerulonephritis in transgenic mice. Therefore, upregulation of IL-10 can compensate for the overexpression of inflammatory cytokines that play a role against local inflammatory processes and prevent acute kidney lesions. This recommends that IL-10 overexpression could effectively control and suppress the pro-inflammatory

response with minimum rat lesions, but also successfully regulate B cells and the production of specific antibodies, resulting in the elimination of the bacteria from the tissues of rats.

In furtherance, study of Vernel-Pauillac & Merien (2006) described cytokine mRNA expression in blood of leptospiral infected hamsters. They detected an obvious TNF-α and IP-10 expression within very early after infection (1 hour and 8 hours, respectively). IL-10 expression was notable after second and third days of infections in delayed hamsters. The cytokine expression in livers and kidneys has been observed in the present study. The results seem to correlate with those reported by the Vernel-Pauillac & Merien (2016). Throughout this study, the expression of TNF-α was not detected in the livers and was expressed only in the kidneys. The IL-10 was only obviously expressed in the kidneys that correlated with data reported by this group as well. In a study by Tajiki et al. (1997), it was found that the high IL-10 or TNF-α ratio was related with lower severity of disease. However, Kyriakidis et al. (2011) demonstrated a high IL-10 or TNF-α ratio has been related with the fatal outcome of patients with leptospirosis. These data proposed the role of pro-inflammatory as well as anti-inflammatory cytokines in leptospirosis disease outcomes. However, it is necessary to further explore the application of these two cytokines as a disease marker in larger groups of infected rats. Rapid production of IL-10 anti-inflammatory cytokine could control the production of pro-inflammatory cytokines assisting in mitigation of tissue injury. Thus, a dramatic imbalance in cytokine release following leptospiral infection may play a significant role in severe leptospirosis development.

The production of inflammatory cytokines and chemokines is often studied to establish the role of immune response in organ damage observed in infectious diseases. The results

of the present study evidenced that chemokines and inflammatory cytokines were rapidly upregulated in the kidneys of rats compared to the livers during leptospiral infection, proving the colonisation of leptospires in the liver only during an early stage. It has also been proposed that these mediators play a role in pathogenesis of leptospirosis. These results propose that inflammatory cytokines should be released rapidly for bacterial eradication. Delayed expression can enable colonisation and proliferation of bacteria. Inflammatory cytokine release, however, should be properly modulated. Prolonged secretion could encourage further damage to the tissue. Leptospiral toxins may directly cause damage to the tissue that is further encouraged by inflammatory response. In short, a high cellular immune response took place during leptospirosis in the disease's initial stage followed by subsequent humoral response. The discoveries from this study propose that cytokine production such as IL-1β, IL-6, IL-10, TNF-α, MIP-1α, and IP-10 should be associated with leptospirosis immunopathogenesis. Overall, this study showed major differences in cytokine expression profiles depending on organs and host resistance to leptospirosis, and these may be part of a host tissue damage elucidation.

# CHAPTER 6: CONCLUSION

In conclusion, the molecular complexities in *Rattus* do exist at the genus level. Molecular barcoding utilised in this study to confirm species identity of all *Rattus* sampled successfully led to the identification of six species: Asian house rat (*R. tanezumi*), Malaysian wood rat (*R. tiomanicus*), Norway/brown rat (*R. norvegicus*), Pacific rat (*R. exulans*), ricefield rat (*R. argentiventer*), and an unnamed black/house rat (*R. rattus cf*. lineage IV sensu Aplin et al., 2011). This suggests a total of six species from the genus *Rattus* are present in Peninsular Malaysia. Comparison of the sequences with sequences available in the literature suggests that the Asian house rat, brown rat, and Pacific rat are obviously invasive in this region.

Molecular analysis via Cyt b gene illustrated distant relationship between Southeast Asian *Rattus rattus* (*R. rattus cf*. lineage IV sensu Aplin et al., 2011) towards *R. rattus* sensu stricto (lineage I) and Himalayan *R. rattus* (lineage III). The Isthmus of Kra which served as a biogeographical barrier may have been an utmost factor in the divergence of these three *Rattus* lineages. Considering that the common karyotype of *R. rattus* globally was 2n = 38/40/42, it can be accepted presumably that this molecular difference discovered between *R. rattus* lineage IV with *R. rattus* lineages I and III populations are also followed by karyological or chromosomal variation. However, this assumption requires to be additionally proved later on cytogenetic study to evaluate the divergence of the lineages. Also, this study exhibited the presence of two complex sympatric species of *R. rattus cf*. lineage IV sensu Aplin et al., 2011 and *R. tanezumi* in Peninsular Malaysia. The results disclosed a huge variability between them which could be defined in perspective of evolution.

In addition, the morphological analyses suggest the presence of sexual dimorphism and ontogenetic variation in populations of *R. rattus* in Peninsular Malaysia. This study recommends that the failure to ideally assess age variation and sexual dimorphism may prompt in faulty conclusions in consequent analyses of the depiction of taxa. Besides, the findings from the population structure study are directly relevant to the management approaches of *R. rattus* in Peninsular Malaysia, mainly in terms of the eradication process. It is crucial to recognise the genetic diversity between populations of *R. rattus*, to enable the management of the unique evolutionary lineages of the species. However, further conclusions can only be made with extensive sampling from a broader geographical area. It is proposed that future comparative studies should aim to utilise several molecular markers in order to resolve the complexities within *Rattus* species, to develop the current understanding. Additional research is also needed to determine the species name of the black rat and to determine the distribution and biology of all of these species both native and invasive in order to better conserve native wildlife and control disease.

In furtherance, this study presents an insight on essential surveillance information on the prevalence of *Leptospira* spp. from rats in Peninsular Malaysia. Two pathogenic species of *Leptospira*, namely *L. interrogans* and *L. borgpetersenii* were discovered from *R. rattus* with a large dissemination, emphasising a critical public health concern. To a greater extent, this study suggests that PCR method is a reliable tool in molecular diagnostic which succeed in the discovery of leptospiral DNA from rats. The diagnosis of leptospirosis can be facilitated using 16S rRNA and LipL32 genes as molecular markers of infection. Rats should be marked as the most important host and reservoir of leptospirosis. The risk of leptospirosis infection can be diminished with numerous preventive measures such as rodent control programme. The findings from this study

could provide hints for subsequent molecular, clinical, and epidemiological research to execute advance diagnostic strategies.

Moreover, the contribution of cytokines to the host response to leptospires was also established by wild rat models of leptospirosis. The results of this study showed the upregulation of chemokines and inflammatory cytokines in the kidneys of *R. rattus* infected with leptospirosis. Thus, best modulation of chemokines and inflammatory cytokines in rat kidneys possibly signifies limited pathological lesions within this resistant model. This study offered an insight into the mechanism of antibacterial in *R. rattus* and the role of differentially expressed immune genes in response to leptospiral infection. Understanding the rats' defence mechanisms can contribute to improving leptospirosis management. This study has also generated an abundant list of transcripts from *R. rattus* which will offer a fundamental basis for future genomics research in this field, inclusive of host-pathogen interactions. The tissue specific transcripts identified in the present study can be implicated in the rat immune responses as candidate genes, and providing additional genomic resources for subsequent study on *R. rattus*. Elucidating the functions of the genes and their corresponding products during chronic disease would further reveal the pathogenic mechanisms of chronic leptospirosis, altering therapeutic and control strategies development.

## REFERENCES


Abdel-Rahman, E. H. (2005). *Taxonomic status of Arvicanthis lesson, 1842 from Sudan*. (Doctoral dissertation, Ph. D. Thesis, University of Kwa-Zulu Natal, Durban, South Africa).

Abdel-Rahman, E. H., Taylor, P. J., Contrafatto, G., Lamb, J. M., Bloomer, P., & Chimimba, C. T. (2008). Geometric craniometric analysis of sexual dimorphism and ontogenetic variation: A case study based on two geographically disparate species, *Aethomys ineptus* from southern Africa and *Arvicanthis niloticus* from Sudan (Rodentia: Muridae). *Mammalian Biology*, *74*, 361-373.

Abdul-Latiff, M. A. B., Ruslin, F., Faiq, H., Hairul, M. S., Rovie-Ryan, J. J., Abdul-Patah, P., . . . Md-Zain, B. M. (2014). Continental monophyly and molecular divergence of Peninsular Malaysia's *Macaca fascicularis fascicularis*. *BioMed Research International*, *2014*, Article#897682.

Abhyankar, A., Park, H. B., Tonolo, G., & Luthman, H. (2009). Comparative sequence analysis of the non-protein-coding mitochondrial DNA of inbred rat strains. *PLoS ONE*, *4*(12), Article#e8148.

Aboim, M. A., Menezes, G. M., Schlitt, T., & Rogers, A. D. (2005). Genetic structure and history of populations of the deep-sea fish *Helicolenus dactylopterus* (Delaroche, 1809) inferred from mtDNA sequence analysis. *Molecular Ecology*, *14*(5), 1343-1354.

Adhikari, P., Han, S. H., Kim, Y. K., Kim, T. W., Thapa, T. B., Subedi, . . . Oh, H. S. (2018). New record of the oriental house rat, *Rattus tanezumi*, in Nepal inferred from mitochondrial cytochrome b gene sequences. *Mitochondrial DNA Part B*, *3*(1), 386-390.

Adkins, R. M., Gelke, E. L., Rowe, D., & Honeycutt, R. L. (2001). Molecular phylogeny and divergence time estimates for major rodent groups: Evidence from multiple genes. *Molecular Biology and Evolution*, *18*(5), 777-791.

Al-Ashy, R., Chakroun, I., El-Sabban, M. E., & Homaidan, F. R. (2006). The role of NF-κB in mediating the anti-inflammatory effects of IL-10 in intestinal epithelial cells. *Cytokine*, *36*(1-2), 1-8.

Alexandrino, J., Arntzen, J. W., & Ferrand, N. (2002). Nested clade analysis and the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Heredity*, *88*(1), 66-74.

Allaya, H., Faleh, A. B., Zrelli, S., Hajjej, G., Hattour, A., Quignard, J. P., & Trabelsi, M. (2016). Morphological variation of bullet tuna *Auxis rochei* (Risso, 1810) from Tunisian Waters. *Acta Zoologica Bulgarica*, *68*(3), 433-442.

Ansmann, I. C., Parra, G. J., Lanyon, J. M., & Seddon, J. M. (2012). Fine-scale genetic population structure in a mobile marine mammal: Inshore bottlenose dolphins in Moreton Bay, Australia. *Molecular Ecology*, *21*(18), 4472-4485.

Aplin, K. P., Chesser, T., & Have, J. T. (2003). Evolutionary biology of the genus *Rattus*: Profile of an archetypal rodent pest. *ACIAR Monograph Series*, *96*, 487-498.

Aplin, K. P., Suzuki, H., Chinen, A. A., Chesser, R. T., Have, J. T., Donnellan, S. C., . . . Catzeflis, F. (2011). Multiple geographic origins of commensalism and complex dispersal history of black rats. *PLoS ONE*, *6*(11), Article#e26357.

Arnalich, F., Garcia-Palomero, E., López, J., Jiménez, M., Madero, R., Renart, J., . . . Montiel, C. (2000). Predictive value of nuclear factor κB activity and plasma cytokine levels in patients with sepsis. *Infection and Immunity*, *68*(4), 1942-1945.

Athanazio, D. A., Silva, E. F., Santos, C. S., Rocha, G. M., Vannier-Santos, M. A., McBride, A. J., . . . Reis, M. G. (2008). *Rattus norvegicus* as a model for persistent renal colonization by pathogenic *Leptospira interrogans*. *Acta Tropica*, *105*, 176-180.

Atkinson, I. A. (1985). The spread of commensal species of *Rattus* to Oceanic Islands and their effects on Island Avifaunas. *ICPB Technical Publication*, *3*, 35-81.

Audic, S., & Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Research*, *7*, 986-995.

Avise, J. C. (1994). Molecular markers, natural history and evolution. New York, NY: Chapman & Hall.

Backstedt, B. T., Buyuktanir, O., Lindow, J., Wunder Jr, E. A., Reis, M. G., Usmani-Brown, S., . . . Pal, U. (2015). Efficient detection of pathogenic leptospires using 16S ribosomal RNA. *PLoS ONE*, *10*(6), Article#e0128913.

Bahaman, A. R., & Ibrahim, A. L. (1988). A review of leptospirosis in Malaysia. *Veterinary Research Communications*, *12*, 179-189.

Baker, R. J., & Bradley, R. D. (2006). Speciation in mammals and the genetic species concept. *Journal of Mammalogy*, *87*(4), 643-662.

Balakirev, A. E., & Rozhnov, V. V. (2012). Contribution to the species composition and taxonomic status of some *Rattus* inhabiting Southern Vietnam and Sundaland. *Russian Journal of Theriology*, *11*(1), 33-45.

Balakirev, A. E., Abramov, A. V., & Rozhnov, V. V. (2013). Revision of the genus *Leopoldamys* (Rodentia, Muridae) as inferred from morphological and molecular data, with a special emphasis on the species composition in continental Indochina. *Zootaxa*, *3640*(4), 521-549.

Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, *16*(1), 37-48.

Banfi, E., Cinco, M., Bellini, M., & Soranzo, M. R. (1982). The role of antibodies and serum complement in the interaction between macrophages and leptospires. *Journal of General Microbiology*, *128*, 813-816.

Barbosa, A. S., Abreu, P. A., Vasconcellos, S. A., Morais, Z. M., Gonçales, A. P., Silva, A. S., . . . Isaac, L. (2009). Immune evasion of *Leptospira* species by acquisition of human complement regulator C4BP. *Infection and Immunity*, *77*(3), 1137-1143.

Barbosa, S., Pauperio, J., Searle, J. B., & Alves, P. C. (2013). Genetic identification of Iberian rodent species using both mitochondrial and nuclear loci: application to non-invasive sampling. *Molecular Ecology Resources*, *13*(1), 43-56.

Bastos, A. D., Nair, D., Taylor, P. J., Brettschneider, H., Kirsten, F., Mostert, E., . . . Contrafatto, G. (2011). Genetic monitoring detects an overlooked cryptic species and reveals the diversity and distribution of three invasive *Rattus* congeners in south Africa. *BMC Genetics*, *12*(1), Article#26.

Baverstock, P. R., Adams, M., Maxson, L. R., & Yosida, T. H. (1983). Genetic differentiation among karyotypic forms of the black rat, *Rattus rattus*. *Genetics*, *105*(4), 969-983.

Begon, M., Hazel, S. M., Telfer, S., Bown, K., Carslake, D., Cavanagh, R., . . . Bennett, M. (2003). Rodents, cowpox virus and islands: Densities, numbers and thresholds. *Journal of Animal Ecology*, *72*(2), 343-355.

Bellinvia, E., Munclinger, P., & Flegr, J. (1999). Application of the RAPD technique for a study of the phylogenetic relationships among eight species of the genus *Apodemus*. *Folia Zoologica Praha*, *48*(4), 241-248.

Belmain, S. R., Meyer, A. N., Timbrine, R., & Penicela, L. (2003). Managing rodent pests in households and food stores through intensive trapping. *ACIAR Monograph Series*, *96*, 440-445.

Ben Faleh, A. R., Cosson, J. F., Tatard, C., Ben Othmen, A., Said, K., & Granjon, L. (2010). Are there two cryptic species of the lesser jerboa *Jaculus jaculus* (Rodentia: Dipodidae) in Tunisia? Evidence from molecular, morphometric, and cytogenetic data. *Biological Journal of the Linnean Society*, *99*, 673-686.

Ben Faleh, A. R., Annabi, A., & Said, K. (2012). Morphometric variation in Black Rat *Rattus rattus* (Rodentia: Muridae) from Tunisia. *Acta Zoologica Bulgarica*, *64*, 381-387.

Ben Faleh, A. R., Cornette, R., Annabi, A., Said, K., & Denys, C. (2013). Patterns of size and shape skull variability in Tunisian populations of *Jaculus jaculus* (Rodentia: Dipodidae). *Acta Zoologica Bulgarica*, *65*, 217-223.

Ben Faleh, A. R., Allaya, H., Quignard, J. P., Shahin, A. A. A. B., & Trabelsi, M. (2016). Patterns of skull shape variation in relation to some geoclimatic conditions in the greater jerboa *Jaculus orientalis* (Rodentia, Dipodidae) from Tunisia. *Turkish Journal of Zoology*, *40*, 1-10.

Benacer, D., Mohd Zain, S. N., Amran, F., Galloway, R. L., & Thong, K. L. (2013). Isolation and molecular characterization of *Leptospira interrogans* and *Leptospira borgpetersenii* isolates from the urban rat populations of Kuala Lumpur, Malaysia. *American Journal of Tropical Medicine and Hygiene*, *88*(4), 704-709.

Benacer, D., Mohd Zain, S. N., Sim, S. Z., Mohd Khalid, M. K., Galloway, R. L., Souris, M., & Thong, K. L. (2016). Determination of *Leptospira borgpetersenii* serovar Javanica and *Leptospira interrogans* serovar Bataviae as the persistent *Leptospira* serovars circulating in the urban rat populations in Peninsular Malaysia. *Parasites and Vectors*, *9*(1), Article#117.

Bezerra, A. M. R., & de Oliveira, J. A. (2010). Taxonomic implications of cranial morphometric variation in the genus *Clyomys* Thomas, 1916 (Rodentia: Echimyidae). *Journal of Mammalogy*, *91*, 260-272.

Bharti, A. R., Nally, J. E., Ricaldi, J. N., Matthias, M. A., Diaz, M. M., Lovett, M. A., . . . Vinetz, J. M. (2003). Leptospirosis: A zoonotic disease of global importance. *The Lancet Infectious Diseases*, *3*, 757-771.

Bohoussou, H. K., Akpatou, B. K., Kadjo, B., Soulemane, O., N'Goran, E. K., & Nicolas, V. (2014). Morphometric variation in the forest rodent *Malacomys edwardsi* (Rodentia: Muridae) in Côte d'Ivoire. *Journal of Applied Biosciences*, *80*, 7014-7023.

Bouga, M., Harizanis, P. C., Kilias, G., & Alahiotis, S. (2005). Genetic divergence and phylogenetic relationships of honey bee *Apis mellifera* (Hymenoptera: Apidae) populations from Greece and Cyprus using PCR-RFLP analysis of three mtDNA segments. *Apidologie*, *36*(3), 335-344.

Bourhy, P., Collet, L., Brisse, S., & Picardeau, M. (2014). *Leptospira mayottensis* sp. nov., a pathogenic species of the genus *Leptospira* isolated from humans. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 4061-4067.

Bozza, F. A., Salluh, J. I., Japiassu, A. M., Soares, M., Assis, E. F., Gomes, R. N., . . . Bozza, P. T. (2007). Cytokine profiles as markers of disease severity in sepsis: A multiplex analysis. *Critical Care*, *11*(2), Article#R49.

Bradley, R. D., & Baker, R. J. (2001). A test of the genetic species concept: Cytochrome b sequences and mammals. *Journal of Mammalogy*, *82*, 960-973.

Bronner, G. N., Merwe, M. V. D., & Njobe, K. (2007). Nongeographic cranial variation in two medically important rodents from south Africa, *Matomys natalensis* and *Mastomys coucha*. *Journal of Mammalogy*, *88*, 1179-1194.

Bronson, F. H. (1989). Mammalian reproductive biology. Chicago, US: University of Chicago Press.

Buckle, A. P. (1994). Rodent control methods: Chemical. In A. P. Buckle, & R. H. Smith (Eds.), Rodent pests and their control (pp. 127-160). Wallingford, UK: CABI Publishing.

Buckle, A. P., & Smith, R. H. (2015). Rodent pests and their control. Wallingford, UK: CABI Publishing.

Buckley, D. H., & Schmidt, T. M. (2002). Exploring the biodiversity of soil—a microbial rain forest. *Biodiversity of Microbial Life*, *2002*, 183-208.

Bucklin, A., Guarnieri, M., Hill, R. S., Bentley, A. M., & Kaartvedt, S. (1999). Taxonomic and systematic assessment of planktonic copepods using mitochondrial COI sequence variation and competitive, species-specific PCR. In J. P. Zehr, & M. A. Voytek (Eds.), Molecular ecology of aquatic communities (pp. 239-254). Dordrecht, Netherlands: Springer Science+Business Media.

Bunlungsup, S., Imai, H., Hamada, Y., Matsudaira, K., & Malaivijitnond, S. (2017). Mitochondrial DNA and two Y-chromosome genes of common long-tailed macaques (*Macaca fascicularis fascicularis*) throughout Thailand and vicinity. *American Journal of Primatology*, *79*(2), Article#e22596.

Cagliero, J., Villanueva, S. Y. A., & Matsui, M. (2018). *Leptospira* pathophysiology: Into the storm of cytokines. *Frontiers in Cellular and Infection Microbiology*, *8*, Article#204.

Caimano, M. J., Sivasankaran, S. K., Allard, A., Hurley, D., Hokamp, K., Grassmann, A. A., . . . Nally, J. E. (2014). A model system for studying the transcriptomic and physiological changes associated with mammalian host-adaptation by *Leptospira interrogans* serovar Copenhageni. *PLoS Pathogens*, *10*(3), Article#e1004004.

Camacho-Sanchez, M., Leonard, J. A., Fitriana, Y., Tilak, M. K., & Fabre, P. H. (2017). The generic status of *Rattus annandalei* (Bonhote, 1903) (Rodentia, Murinae) and its evolutionary implications. *Journal of Mammalogy*, *98*(5), 1340-1355.

Camacho-Sanchez, M., Quintanilla, I., Hawkins, M. T. R., Tuh, F. Y. Y., Wells, K., Maldonado, J. E., & Leonard, J. A. (2018). Interglacial refugia on tropical mountains: novel insights from the summit rat (*Rattus baluensis*), a Borneo mountain endemic. *Diversity and Distributions*, *24*, 1252-1266.

Cann, R. L., Stoneking, M., & Wilson, A. C. (1987). Mitochondrial DNA and human evolution. *Nature*, *325*(6099), Article#31.

Cantatore, P., Roberti, M., Pesole, G., Ludovico, A., Milella, F., Gadaletal, M. N., & Saccone, C. (1994). Evolutionary analysis of cytochrome b sequences in some Perciformes: Evidence for a slower rate of evolution than in mammals. *Journal of Molecular Evolution*, *39*(6), 589-597.

Cao, L. F., Zheng, H. Q., Hu, C. Y., He, S. Y., Kuang, H. O., & Hu, F. L. (2012). Phylogeography of *Apis dorsata* (Hymenoptera: Apidae) from China and neighboring Asian areas. *Annals of the Entomological Society of America*, *105*(2), 298-304.

Catzeflis, F. M., Hanni, C., Sourrouille, P., & Douzery, E. (1995). Molecular systematics of hystricognath rodents: The contribution of sciurognath mitochondrial 12S rRNA sequences. *Molecular Phylogenetics and Evolution*, *3*(4), 357-360.

Cerqueira, G. M., McBride, A. J., Queiroz, A., Pinto, L. S., Silva, É. F., Hartskeerl, R. A., . . . Dellagostin, O. A. (2010). Monitoring *Leptospira* strain collections: The need for quality control. *The American Journal of Tropical Medicine and Hygiene*, *82*(1), 83-87.

Cesaroni, D., Matarazzo, P., Allegrucci, G., & Sbordoni, V. (1997). Comparing patterns of geographic variation in cave crickets by combining geostatistic methods and Mantel tests. *Journal of Biogeography*, *24*, 419-431.

Chaimanee, Y., & Jaeger, J. J. (2001). Evolution of *Rattus* (Mammalia, Rodentia) during the Plio-Pleistocene in Thailand. *Historical Biology*, *15*(1-2), 181-191.

Chan, E. Y. (2005). Advance in sequencing technology. *Mutation Research*, *573*, 13-40.

Chassin, C., Picardeau, M., Goujon, J. M., Bourhy, P., Quellard, N., Darche, S., . . . Buzoni-Gatel, D. (2009). TLR4- and TLR2-mediated B cell responses control the clearance of the bacterial pathogen, *Leptospira interrogans*. *The Journal of Immunology*, *183*(4), 2669-2677.

Chelomina, G. N. (1996). DNA restriction polymorphism in European and Asiatic species of wood mice of the genus *Apodemus*. *Genetika*, *32*(10), 1381-1386.

Chelomina, G. N. (1998). Molecular phylogeny of forest and field mice of the genus *Apodemus* (Muridae, Rodentia) based on the data on restriction analysis of total nuclear DNA. *Genetika*, *34*(9), 1286-1292.

Chelomina, G. N., Suzuki, H., Tsuchiya, K., Moriwaki, K., Liapunova, E. A., & Vorontsov, N. N. (1998). Sequencing of the mtDNA cytochrome b gene and reconstruction of the matriarchal relationships between wood and field mice of the genus *Apodemus* (Muridae, Rodentia). *Genetika*, *34*(5), 650-661.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., & Thompson, J. D. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Research*, *31*, 3497-3500.

Chimimba, C. T. (2001). Infraspecific morphometric variation in *Aethomys namaquensis* (Rodentia: Muridae) from southern Africa. *Journal of Zoology*, *253*, 191-210.

Chinen, A. A., Suzuki, H., Aplin, K. P., Tsuchiya, K., & Suzuki, S. (2005). Preliminary genetic characterization of two lineages of black rats (*Rattus rattus* sensu lato) in Japan, with evidence for introgression at several localities. *Genes and Genetic Systems*, *80*(5), 367-375.

Chingangbam, D. S., Laishram, J. M., & Suzuki, H. (2015). Molecular phylogenetic characterization of common murine rodents from Manipur, Northeast India. *Genes and Genetic Systems*, *90*(1), 21-30.

Chou, L. F., Chen, T. W., Yang, H. Y., Chang, M. Y., Hsu, S. H., Tsai, C. Y., . . . Yang, C. W. (2018). Murine renal transcriptome profiles upon leptospiral infection: Implications for chronic kidney diseases. *The Journal of Infectious Diseases*, *218*(9), 1411-1423.

Clark, D. A. (1981). Foraging patterns of black rats across a desert-montane forest gradient in the Galapagos Islands. *Biotropica*, *13*, 182-194.

Clark, D. A. (1982). Foraging behavior of a vertebrate omnivore (*Rattus rattus*): Meal structure, sampling, and diet breadth. *Ecology*, *63*, 763-772.

Clarridge, J. E. (2004). Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clinical Microbiology Reviews*, *17*(4), 840-862.

Cock, P. J., Fields, C. J., Goto, N., Heuer, M. L., & Rice, P. M. (2010). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, *38*(6), 1767-1771.

Conroy, C. J., Rowe, K. C., Rowe, K. M., Kamath, P. L., Aplin, K. P., Hui, L., . . . Patton, J. L. (2013). Cryptic genetic diversity in *Rattus* of the San Francisco Bay region, California. *Biological Invasions*, *15*(4), 741-758.

Corbet, G. B., & Hill, J. E. (1992). The mammals of the Indomalayan region: A systematic review. New York, NY: Oxford University Press.

Corney, B. G., Slack, A. T., Symonds, M. L., Dohnt, M. F., McClintock, C. S., McGowan, M. R., & Smythe, L. D. (2008). *Leptospira weilii* serovar Topaz, a new member of the Tarassovi serogroup isolated from a bovine source in Queensland, Australia. *International Journal of Systematic and Evolutionary Microbiology*, *58*(10), 2249-2252.

Cosson, J. F., Picardeau, M., Mielcarek, M., Tatard, C., Chaval, Y., Suputtamongkol, Y., . . . Morand, S. (2014). Epidemiology of *Leptospira* transmitted by rodents in southeast Asia. *PLoS Neglected Tropical Diseases*, *8*(6), Article#e2902.

Croucher, N. J., & Thomson, N. R. (2010). Studying bacterial transcriptomes using RNA-seq. *Current Opinion in Microbiology*, *13*(5), 619-624.

da Silva, J. B., Carvalho, E., Covarrubias, A. E., Ching, A. T. C., Mattaraia, V. G., Paiva, D., . . . Zorn, T. T. (2012). Induction of TNF-alfa and CXCL-2 mRNAs in different organs of mice infected with pathogenic *Leptospira*. *Microbial Pathogenesis*, *52*(4), 206-216.

Darvish, J., Hajjar, T., Moghadam Matin, M., Haddad, F., & Akbary rad, S. (2008). New species of five-toed jerboa (Rodentia: Dipodidae: Allactaginae) from north-east Iran. *Journal of Sciences, Islamic Republic of Iran*, *19*(2), 103-109.

De Graaff, G. (1981). The rodents of southern Africa: Notes on their identification, distribution, ecology, and taxonomy. Oxford, UK: Butterworth-Heinemann.

DeBry, R. W., & Sagel, R. M. (2001). Phylogeny of Rodentia (Mammalia) inferred from the nuclear-encoded gene IRBP. *Molecular Phylogenetics and Evolution*, *19*(2), 290-301.

Dejtaradol, A., Renner, S. C., Karapan, S., Bates, P. J., Moyle, R. G., & Päckert, M. (2016). Indochinese-Sundaic faunal transition and phylogeographical divides north of the Isthmus of Kra in Southeast Asian Bulbuls (Aves: Pycnonotidae). *Journal of Biogeography*, *43*(3), 471-483.

DeSalle, R., & Amato, G. (2004). The expansion of conservation genetics. *Nature Reviews Genetics*, *5*(9), 702-712.

Donoghue, P. C., & Benton, M. J. (2007). Rocks and clocks: Calibrating the tree of life using fossils and molecules. *Trends in Ecology and Evolution*, *22*(8), 424-431.

Dötsch, A., Eckweiler, D., Schniederjans, M., Zimmermann, A., Jensen, V., Scharfe, M., . . . Häussler, S. (2012). The *Pseudomonas aeruginosa* transcriptome in planktonic cultures and static biofilms using RNA sequencing. *PLoS ONE*, *7*(2), Article#e31092.

Drazo, N. A., Kennis, J., Leirs, H., & Migimiru, D. A. (2008). Farmer survey in the hinterland of Kisangani (Democratic Republic of Congo) on rodent crop damage and rodent control techniques used. *Mammalia*, *72*(3), 192-197.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*(1), Article#214.

Duell, B. L., Tan, C. K., Carey, A. J., Wu, F., Cripps, A. W., & Ulett, G. C. (2012). Recent insights into microbial triggers of interleukin-10 production in the host and the impact on infectious disease pathogenesis. *FEMS Immunology and Medical Microbiology*, *64*(3), 295-313.

Ellerman, J. R. (1941). The families and genera of living rodents (Vol. 2). London, UK: British Museum of Natural History.

Ellerman, J. R. (1949). The families and genera of living rodents (Vol. 3). London, UK: British Museum of Natural History.

Ellerman, J. R., & Morrison-Scott, T.C.S. (1966). Checklist of Palaearctic and Indian mammals, 1758 to 1946 (2nd ed.). Oxford, UK: Alden Press.

Endicott, P., Ho, S. Y., Metspalu, M., & Stringer, C. (2009). Evaluating the mitochondrial timescale of human evolution. *Trends in Ecology and Evolution*, *24*(9), 515-521.

Erol, E., Jackson, C. B., Steinman, M., Meares, K., Donahoe, J., Kelly, N., . . . Carter, C. N. (2015). A diagnostic evaluation of real-time PCR, fluorescent antibody and microscopic agglutination tests in cases of equine leptospiral abortion. *Equine Veterinary Journal*, *47*(2), 171-174.

Estavoyer, J. M., Racadot, E., Couetdic, G., Leroy, J., & Grosperrin, L. (1991). Tumor necrosis factor in patients with leptospirosis. *Reviews of Infectious Diseases*, *13*(6), 1245-1246.

Evangelista, K. V., & Coburn, J. (2010). *Leptospira* as an emerging pathogen: A review of its biology, pathogenesis and host immune responses. *Future Microbiology*, *5*(9), 1413-1425.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*, 479-491.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver. 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564-567.

Faine, S., Adler, B., Bolin, C., & Perolat, P. (1999). *Leptospira* and Leptospirosis (2nd ed.). Melbourne, AU: Medisci Press.

Fanton d'Andon, M., Quellard, N., Fernandez, B., Ratet, G., Lacroix-Lamandé, S., Vandewalle, A., Boneca, I. G., Goujon, J. M., . . . Werts, C. (2014). *Leptospira interrogans* induces fibrosis in the mouse kidney through Inos-dependent, TLR- and NLR-independent signaling pathways. *PLoS Neglected Tropical Diseases*, *8*(1), Article#e2664.

Farr, R. W. (1995). Leptospirosis. *Clinical Infectious Diseases*, *21*(1), 7-8.

Ferguson, J. W. H. (2002). On the use of genetic divergence for identifying species. *Biological Journal of the Linnean Society*, *75*, 509-516.

Ferrer, M. F., Scharrig, E., Alberdi, L., Cedola, M., Pretre, G., Drut, R., . . . Gomez, R. M. (2014). Decay-accelerating factor 1 deficiency exacerbates leptospiral-induced murine chronic nephritis and renal fibrosis. *PLoS ONE*, *9*(7), Article#e102860.

Filiatrault, M. J. (2011). Progress in prokaryotic transcriptomics. *Current Opinion in Microbiology*, *14*(5), 579-586.

Fletcher, W. (1928). Recent work on leptospirosis, tsutsugamushi disease and tropical typhus in the Federated Malay States. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *21*, 265-288.

Forman, S., Hungerford, N., Yamakawa, M., Yanase, T., Tsai, H. J., Joo, Y. S., . . . Nha, J. J. (2008). Climate change impacts and risks for animal health in Asia. *Revue Scientifique et Technique (International Office of Epizootics)*, *27*, 581-597.

Fraga, T. R., Isaac, L., & Barbosa, A. S. (2016). Complement evasion by pathogenic *Leptospira*. *Frontiers in Immunology*, *7*(623), 1-7.

Francis, C. M. (2008). A field guide to the mammals of south-east Asia. London, UK: New Holland Publishers Limited.

Francoy, T. M., Wittmann, D., Drauschke, M., Müller, S., Steinhage, V., Bezerra-Laure, M. A. F., . . . Gonçalves, L. S. (2008). Identification of africanized honey bees through wing morphometrics: Two fast and efficient procedures. *Apidologie*, *39*(5), 488-494.

Freeland, J. R. (2005). Molecular ecology. New Jersey, US: John Wiley & Sons.

Frese, S. A., MacKenzie, D. A., Peterson, D. A., Schmaltz, R., Fangman, T., Zhou, Y., . . . Juge, N. (2013). Molecular characterization of host-specific biofilm formation in a vertebrate gut symbiont. *PLoS Genetics*, *9*(12), Article#e1004057.

Fu, Y. X., & Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics*, *133*(3), 693-709.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*(2), 915-925.

Fujishima, S., Sasaki, J., Shinozawa, Y., Takuma, K., Kimura, H., Suzuki, M., . . . Kanazawa, M. (1996). Serum MIP-1α and IL-8 in septic patients. *Intensive Care Medicine*, *22*(11), 1169-1175.

Fujita, R., Koizumi, N., Sugiyama, H., Tomizawa, R., Sato, R., & Ohnishi, M. (2015). Comparison of bacterial burden and cytokine gene expression in golden hamsters in early phase of infection with two different strains of *Leptospira interrogans*. *PLoS ONE*, *10*(7), Article#e0132694.

Furano, A. V., & Usdin, K. (1995). DNA fossils and phylogenetic analysis using L1 (Line-1, long interspersed repeated) DNA to determine the evolutionary history of mammals. *Journal of Biological Chemistry*, *270*(43), 25301-25304.

Galan, M., Pagès, M., & Cosson, J. F. (2012). Next-generation sequencing for rodent barcoding: species identification from fresh, degraded and environmental samples. *PLoS ONE*, *7*(11), Article#e48374.

Gesser, B., Leffers, H., Jinquan, T., Vestergaard, C., Kirstein, N., Sindet-Pedersen, S., . . . Larsen, C. G. (1997). Identification of functional domains on human interleukin 10. *Proceedings of the National Academy of Sciences*, *94*(26), 14620-14625.

Gharizadeh, B., Ohlin, A., Mölling, P., Bäckman, A., Aminia, B., Olcénc, P., & Nyréna, P. (2003). Multiple group-specific sequencing primers for reliable and rapid DNA sequencing. *Molecular and Cellular Probes*, *17*, 203-210.

Gilbert, S. F. (2003). The morphogenesis of evolutionary developmental biology. *International Journal of Developmental Biology*, *47*(7-8), Article#467.

Goios, A., Pereira, L., Bogue, M., Macaulay, V., & Amorim, A. (2007). MtDNA phylogeny and evolution of laboratory mouse strains. *Genome Research*, *17*(3), 293-298.

Gomes-Solecki, M., Santecchia, I., & Werts, C. (2017). Animal models of leptospirosis: Of mice and hamsters. *Frontiers in Immunology*, *8*, Article#58.

Goncalves-de-Albuquerque, C. F., Burth, P., Silva, A. R., Younes-Ibrahim, M., Castro-Faria-Neto, H. C., & Castro-Faria, M. V. (2012). *Leptospira* and inflammation. *Mediators of Inflammation*, *2012*, Article#317950.

Gorog, A. J., Sinaga, M. H., & Engstrom, M. D. (2004). Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteheadi*). *Biological Journal of Linnean Society*, *81*, 91-109.

Gouy, M., Guindon, S., & Gascuel, O. (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, *27*(2), 221-224.

Grassmann, A. A., McBride, A. J. A., Nally, J. E., & Caimano, M. J. (2015). Generation of mammalian host-adapted *Leptospira interrogans* by cultivation in peritoneal dialysis membrane chamber implantation in rats. *Bio-protocol*, *5*(14), Article#e1536.

Graybeal, A. (1994). Evaluating the phylogenetic utility of genes: A search for genes informative about deep divergences among vertebrates. *Systematic Biology*, *43*(2), 174-193.

Gross, P. S., Bartlett, T. C., Browdy, C. L., Chapman, R. W., & Warr, G. W. (2001). Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific white shrimp, *Litopenaeus vannamei*, and the Atlantic white shrimp, *L. setiferus*. *Developmental and Comparative Immunology*, *25*(7), 565-577.

Groth, J. G., & Barrowclough, G. F. (1999). Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Molecular Phylogenetics and Evolution*, *12*(2), 115-123.

Grzimek, B. (2003). Grzimek's animal life encyclopedia: Mammals. In N. Schlager, D. Olendorf, & M. McDade (Eds.), Order: Rodentia (Vol. 16) (pp. 126-128). Farmington Hills, MI: Gale Group.

Gueńet, J. L. (2005). Assessing the genetic component of the susceptibility of mice to viral infections. *Briefings in Functional Genomics*, *4*(3), 225-240.

Guernier, V., Richard, V., Nhan, T., Rouault, E., Tessier, A., & Musso, D. (2017). *Leptospira* diversity in animals and humans in Tahiti, French Polynesia. *PLoS Neglected Tropical Diseases*, *11*(6), Article#e0005676.

Gülbitti, O. S., & Sümer, S. (2003). Protein and DNA in systematic biology. *Turkish Journal of Biology*, *27*, 47-55.

Hamond, C., Pinna, A., Martins, G., & Lilenbaum, W. (2014). The role of leptospirosis in reproductive disorders in horses. *Tropical Animal Health and Production*, *46*(1), 1-10.

Harpending, H. C. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, *66*(4), 591-600.

Heaney, L. R. (1991). A synopsis of climatic and vegetational change in Southeast Asia. *Climatic Change*, *19*, 53-61.

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

Hebert, P. D., Stoeckle, M. Y., Zemlak, T. S., & Francis, C. M. (2004). Identification of birds through DNA barcodes. *PLoS Biology*, *2*(10), Article#e312.

Hill, C. M. (1997). Crop-raiding by wild vertebrates: The farmer's perspective in an agricultural community in western Uganda. *International Journal of Pest Management*, *43*(1), 77-84.

Hingston, M., Goodman, S. M., Ganzhorn, J. U., & Sommer, S. (2005). Reconstruction of the colonization of southern Madagascar by introduced *Rattus rattus*. *Journal of Biogeography*, *32*(9), 1549-1559.

Ho, S. Y., & Larson, G. (2006). Molecular clocks: When times are a-changin'. *Trends in Genetics*, *22*(2), 79-83.

Hobolth, A., Dutheil, J. Y., Hawks, J., Schierup, M. H., & Mailund, T. (2011). Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection. *Genome Research*, *21*, 349-356.

Hoelzel, A. R., Hey, J., Dahlheim, M. E., Nicholson, C., Burkanov, V., & Black, N. (2007). Evolution of population structure in a highly social top predator, the killer whale. *Molecular Biology and Evolution*, *24*(6), 1407-1415.

Hookey, J. V. (1992). Detection of Leptospiraceae by amplification of 16S ribosomal DNA. *FEMS Microbiology Letters*, *90*, 267-274.

Huang, J., Liang, X., Xuan, Y., Geng, C., Li, Y., Lu, H., . . . Sun, N. (2017). A reference human genome dataset of the BGISEQ-500 sequencer. *GigaScience*, *6*(5), Article#gix024.

Huchon, D., Catzeflis, F. M., & Douzery, E. J. (1999). Molecular evolution of the nuclear von Willebrand factor gene in mammals and the phylogeny of rodents. *Molecular Biology and Evolution*, *16*(5), 577-589.

Huchon, D., Catzeflis, F. M., & Douzery, E. J. (2000). Variance of molecular datings, evolution of rodents and the phylogenetic affinities between Ctenodactylidae and Hystricognathi. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*(1441), 393-402.

Huchon, D., & Douzery, E. J. (2001). From the old world to the new world: A molecular chronicle of the phylogeny and biogeography of hystricognath rodents. *Molecular Phylogenetics and Evolution*, *20*(2), 238-251.

Huchon, D., Madsen, O., Sibbald, M. J., Ament, K., Stanhope, M. J., Catzeflis, F., . . . Douzery, E. J. (2002). Rodent phylogeny and a timescale for the evolution of Glires: evidence from an extensive taxon sampling using three nuclear genes. *Molecular Biology and Evolution*, *19*(7), 1053-1065.

Hudson, R. R., Boos, D. D., & Kaplan, N. L. (1992). A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution*, *9*(1), 138-151.

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, *17*(8), 754-755.

Hughes, J. B., Round, P. D., & Woodruff, D. S. (2003). The Indochinese-Sundaic faunal transition at the Isthmus of Kra: an analysis of resident forest bird species distributions. *Journal of Biogeography*, *30*(4), 569-580.

Innes, J. G. (1990). Ship rat. In C. M. King (Ed.), The handbook of New Zealand mammals (pp. 206-225). Oxford, UK: Oxford University Press.

Irwin, D. M., Kocher, T. D., & Wilson, A. C. (1991). Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution*, *32*(2), 128-144.

Isabel, N., Beaulieu, J., Thériault, P., & Bousquet, J. (1999). Direct evidence for biased gene diversity estimates from dominant random amplified polymorphic DNA (RAPD) fingerprints. *Molecular Ecology*, *8*, 477-483.

Jansa, S. A., & Weksler, M. (2004). Phylogeny of muroid rodents: Relationships within and among major lineages as determined by IRBP gene sequences. *Molecular Phylogenetics and Evolution*, *31*(1), 256-276.

Jansa, S. A., Barker, F. K., & Heaney, L. R. (2006). The pattern and timing of diversification of Philippine endemic rodents: Evidence from mitochondrial and nuclear gene sequences. *Systematic Biology*, *55*(1), 73-88.

Jansa, S. A., Soarimalala, V., Goodman, S. M., & Keith, F. B. (2008). Morphometric variation and phylogeographic structure in *Macrotarsomys bastardi* (Rodentia: Nesomyidae), an endemic Malagasy dry forest rodent. *Journal of Mammalogy*, *89*, 316-324.

Jittapalapong, S., Herbreteau, V., Hugot, J. P., Arreesrisom, P., Karnchanabanthoeng, A., Rerkamnuaychoke, W., & Morand, S. (2009). Relationship of parasites and pathogens diversity to rodents in Thailand. *Kasetsart Journal of Natural Science*, *43*, 106-117.

Jost, B. H., Adler, B., Vinh, T., & Faine, S. (1986). A monoclonal antibody reacting with a determinant on leptospiral lipopolysaccharide protects guinea pigs against leptospirosis. *Journal of Medical Microbiology*, *22*, 269-275.

Kahle, D., & Wickham, H. (2013). Ggmap: Spatial visualization with ggplot2. *The R Journal*, *5*(1), 144-161.

Khaghani, R. (2007). The economic and health impact of rodent in urban zone and harbours and their control methods. *Annals of Military and Health Sciences Research*, *4*(4), 1071-1078.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357-360.

Kirk, J. L., Beaudette, L. A., Hart, M., Moutoglis, P., Klironomos, J. N., Lee, H., & Trevors, J. T. (2004). Methods of studying soil microbial diversity. *Journal of Microbiological Methods*, *58*(2), 169-188.

Knowles, L. L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, *10*(3), 691-701.

Ko, A. I., Galvao, R. M., Ribeiro Dourado, C. M., Johnson, W. D. Jr., & Riley, L. W. (1999). Urban epidemic of severe leptospirosis in Brazil. Salvador leptospirosis study group. *The Lancet*, *354*, 820-825.

Koay, T. K., Nirmal, S., Noitie, L., & Tan, E. (2004). An epidemiological investigation of an outbreak of leptospirosis associated with swimming, Beaufort, Sabah. *Medical Journal of Malaysia*, *59*(4), Article#5.

Koh, H. S., Yoo, M. H., Lee, B. G., & Park, J. G. (2004). Molecular DNA systematic analyses of east Asian mammals: Sequence variation of cytochrome b gene and control region of mitochondrial DNA of common otter, *Lutra lutra lutra* L. (Mammalia, Carnivora) from Korea. *Korean Journal of Biological Sciences*, *8*(3), 231-233.

Kolbert, C. P., & Persing, D. H. (1999). Ribosomal DNA sequencing as a tool for identification of bacterial pathogens. *Current Opinion in Microbiology*, *2*(3), 299-305.

Kong, L., Zhang, Y., Ye, Z. Q., Liu, X. Q., Zhao, S. Q., Wei, L., & Gao, G. (2007). CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Research*, *35*(2), W345-W349.

Kosoy, M., Khlyap, L., Cosson, J. F., & Morand, S. (2015). Aboriginal and invasive rats of genus *Rattus* as hosts of infectious agents. *Vector-Borne and Zoonotic Diseases*, *15*(1), 3-12.

Kruckenhauser, L., Pinsker, W., Haring, E., & Arnold, W. (1999). Marmot phylogeny revisited: Molecular evidence for a diphyletic origin of sociality. *Journal of Zoological Systematics and Evolutionary Research*, *37*(1), 49-56.

Kujoth, G. C., Hiona, A., Pugh, T. D., Someya, S., Panzer, K., Wohlgemuth, S. E., . . . Morrow, J. D. (2005). Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science*, *309*(5733), 481-484.

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

Kumazawa, Y., & Nishida, M. (2000). Molecular phylogeny of osteoglossoids: A new model for Gondwanian origin and plate tectonic transportation of the Asian arowana. *Molecular Biology and Evolution*, *17*(12), 1869-1878.

Kuznetsov, G. V. (2006). Mammals of Vietnam. Moscow, RU: KMK Scientific Press.

Kyriakidis, I., Samara, P., & Papa, A. (2011). Serum TNF-α, sTNFR1, IL-6, IL-8 and IL-10 levels in Weil's syndrome. *Cytokine*, *54*(2), 117-120.

Lack, J. B., Greene, D. U., Conroy, C. J., Hamilton, M. J., Braun, J. K., Mares, M. A., & Van Den Bussche, R. A. (2012). Invasion facilitates hybridization with introgression in the *Rattus rattus* species complex. *Molecular Ecology*, *21*(14), 3545-3561.

Lacroix-Lamandé, S., d'Andon, M. F., Michel, E., Ratet, G., Philpott, D. J., Girardin, S. E., . . . Werts, C. (2012). Downregulation of the Na/K-ATPase pump by leptospiral glycolipoprotein activates the NLRP3 inflammasome. *The Journal of Immunology*, *188*, 2805-2814.

Lagadec, E., Gomard, Y., Le Minter, G., Cordonin, C., Cardinale, E., Ramasindrazana, B., . . . Dellagi, K. (2016). Identification of *Tenrec ecaudatus*, a wild mammal introduced to Mayotte Island, as a reservoir of the newly identified human pathogenic *Leptospira mayottensis*. *PLoS Neglected Tropical Diseases*, *10*(8), Article#e0004933.

Lalis, A., Evin, A., & Denys, C. (2009). Morphological identification of sibling species: The case of west African *Mastomys* (Rodentia: Muridae) in sympatry. *Comptes Rendus Biologies*, *332*, 480-488.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357-359.

Latifah, I., Abdul Halim, A., Rahmat, M. S., Nadia, M. F., Ubil, Z. E, Asmah, H., . . . Nasir, M. A. (2017). Isolation by culture and PCR identification of LipL32 gene of pathogenic *Leptospira* spp. in wild rats of Kuala Lumpur. *The Malaysian Journal of Pathology*, *39*(2), Article#161.

Leclerc-Madlala, S., & Janowski, M. (2004). Perceptions and practices relevant to the transmission of plague, leptospirosis and toxoplasmosis. *Natural Resources Institute Report*, *2780*, 3-69.

Lecompte, E., Aplin, K., Denys, C., Catzeflis, F., Chades, M., & Chevret, P. (2008). Phylogeny and biogeography of African Murinae based on mitochondrial and nuclear gene sequences, with a new tribal classification of the subfamily. *BMC Evolutionary Biology*, *8*(199), 1-21.

Leessmith, D., Clark, G., Vonhaartman, L., Howell, T., Keast, J., King, B., . . . Schreiber, R. W. (1988). Indomalayan Region-a substitute name for Wallace Oriental Region. *Ibis*, *130*(3), 447-448.

Lekagul, B., & McNeely, J. A. (1988). Mammals of Thailand (2nd ed.). Bangkok, TH: Saha Karn Bhaet.

Levett, P. N. (2001). Leptospirosis. *Clinical Microbiology Reviews*, *14*, 296-326.

Levett, P. N., Morey, R. E., Galloway, R. L., & Steigerwalt, A. G. (2006). *Leptospira broomii* sp. nov., isolated from humans with leptospirosis. *International Journal of Systematic and Evolutionary Microbiology*, *56*(3), 671-673.

Li, B., & Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, *12*(1), Article#323.

Li, J., Zheng, X., Cai, Y., Zhang, X., Yang, M., Yue, B., & Li, J. (2015). DNA barcoding of Murinae (Rodentia: Muridae) and Arvicolinae (Rodentia: Cricetidae) distributed in China. *Molecular Ecology Resources*, *15*, 153-167.

Li, K., Hong, W., Jiao, H., Wang, G. D., Rodriguez, K. A., Buffenstein, R., . . . Zhao, H. (2015). Sympatric speciation revealed by genome-wide divergence in the blind mole rat *Spalax*. *Proceedings of the National Academy of Sciences*, *112*(38), 11905-11910.

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

Liu, S. A., Jiang, R. S., Chen, F. J., Wang, W. Y., & Lin, J. C. (2012). Somatic mutations in the D-loop of mitochondrial DNA in oral squamous cell carcinoma. *European Archives of Oto-Rhino-Laryngology*, *269*(6), 1665-1670.

Liu, D. Y., Liu, J., Liu, B. Y., Liu, Y. Y., Xiong, H. R., Hou, W., & Yang, Z. Q. (2017). Phylogenetic analysis based on mitochondrial DNA sequences of wild rats, and the relationship with Seoul virus infection in Hubei, China. *Virologica Sinica*, *32*(3), 235-244.

Lohman, D. J., de Bruyn, M., Page, T., von Rintelen, K., Hall, R., Ng, P. K., . . . von Rintelen, T. (2011). Biogeography of the Indo-Australian archipelago. *Annual Review of Ecology, Evolution, and Systematics*, *42*, 205-226.

Long, J. L. (2003). Introduced mammals of the world: Their history, distribution and influence. Wallingford, UK: CABI.

López, M., Foronda, P., Feliu, C., & Hernández, M. (2013). Genetic characterization of black rat (*Rattus rattus*) of the Canary Islands: origin and colonization. *Biological Invasions*, *15*(11), 2367-2372.

Lowanitchapat, A., Payungporn, S., Sereemaspun, A., Ekpo, P., Phulsuksombati, D., Poovorawan, Y., & Chirathaworn, C. (2010). Expression of TNF-α, TGF-β, IP-10 and IL-10 mRNA in kidneys of hamsters infected with pathogenic *Leptospira*. *Comparative Immunology, Microbiology, and Infectious Diseases*, *33*(5), 423-434.

Lowe, A., Harris, S., & Ashton, P. (2004). Ecological genetics: Design, analysis, and application. Oxford, UK: Blackwell.

Lu, L., Chesters, D., Zhang, W., Li, G., Ma, Y., Ma, H., . . . Liu, Q. (2012). Small mammal investigation in spotted fever focus with DNA-barcoding and taxonomic implications on rodents species from Hainan of China. *PLoS ONE*, *7*(8), Article#e43479.

Lunde, D. P., & Son, N. T. (2001). An identification guide to the rodents of Vietnam. New York, NY: American Museum of Natural History, Center for Biodiversity and Conservation.

Lydeard, C., Holznagel, W. E., Garner, J., Hartfield, P., & Pierson, J. M. (1997). A molecular phylogeny of mobile river drainage basin pleurocerid snails (Caenogastropoda: Cerithioidea). *Molecular Phylogenetics and Evolution*, *7*(1), 117-128.

Lynch, M., & Crease, T. J. (1990). The analysis of population survey data on DNA sequence variation. *Molecular Biology and Evolution*, *7*(4), 377-394.

MacArthur, R. H., & Wilson, E. O. (1967). The theory of island biogeography. Princeton, US: Princeton University Press.

Macholán, M., Mikula, O., & Vohralík, V. (2008). Geographic phenetic variation of two eastern-Mediterranean non-commensal mouse species, *Mus macedonicus* and *M. cypriacus* (Rodentia: Muridae) based on traditional and geometric approaches to morphometrics. *Zoologischer Anzeiger*, *247*, 67-80.

Madsen, K., Cornish, A., Soper, P., McKaigney, C., Jijon, H., Yachimec, C., . . . De Simone, C. (2001). Probiotic bacteria enhance murine and human intestinal epithelial barrier function. *Gastroenterology*, *121*(3), 580-591.

Magalhaes, J., Teruszkin, I., Sutter, F., Dias, A., Hillen, L., & Marina, M. (2010). Multiplex PCR-based detection of *Leptospira* in environmental water samples obtained from a slum settlement. *Memórias do Instituto Oswaldo Cruz*, *105*(3), 353-355.

Malkin, K. (1984). Enhancement of *Leptospira* hardjo agglutination titers in sheep and goat serum by heat inactivation. *Canadian Journal of Comparative Medicine*, *48*, 208-210.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2), 209-220.

Marinho, M., Langoni, H., Oliveira, S. L., Lima, V. M. F., Peiró, J. R., Perri, S. H. V., & Carreira, R. (2005). Role of cytokines, NO, and H2O2 on the immunopathology of Leptospirosis in genetically selected mice. *Journal of Venomous Animals and Toxins Including Tropical Diseases*, *11*(2), 198-212.

Marinho, M., Silva, C., Lima, V. M. F. D., Peiró, J. R., & Perri, S. H. V. (2006). Cytokine and antibody production during murine leptospirosis. *Journal of Venomous Animals and Toxins Including Tropical Diseases*, *12*(4), 595-603.

Martin, Y., Gerlach, G., Schlötterer, C., & Meyer, A. (2000). Molecular phylogeny of European muroid rodents based on complete cytochrome b sequences. *Molecular Phylogenetics and Evolution*, *16*(1), 37-47.

Martin, J. A., & Wang, Z. (2011). Next-generation transcriptome assembly. *Nature Reviews Genetics*, *12*, 671-682.

Martínez, J. J., & Di Cola, V. (2011). Geographic distribution and phenetic skull variation in two close species of *Graomys* (Rodentia, Cricetidae, Sigmodontinae). *Zoologischer Anzeiger*, *250*, 175-194.

Martínez, J. J., Millien, V., Simone, I., & Priotto, J. W. (2014). Ecological preference between generalist and specialist rodents: spatial and environmental correlates of phenotypic variation. *Biological Journal of the Linnean Society*, *112*, 180-203.

Martinez-Lopez, D. G., Fahey, M., & Coburn, J. (2010). Responses of human endothelial cells to pathogenic and non-pathogenic *Leptospira* species. *PLoS Neglected Tropical Diseases*, *4*, Article#e918.

Matisoo-Smith, E., & Robins, J. (2009). Mitochondrial DNA evidence for the spread of Pacific rats through Oceania. *Biological Invasions*, *11*(7), 1521-1527.

Matsui, S. (2005). Protecting human and ecological health under viral threats in Asia. *Water Science and Technology*, *51*, 91-97.

Matsui, M., Rouleau, V., Bruyere-Ostells, L., & Goarant, C. (2011). Gene expression profiles of immune mediators and histopathological findings in animal models of leptospirosis: Comparison between susceptible hamsters and resistant mice. *Infection and Immunity*, *79*, 4480-4492.

Matsui, M., Soupé, M. E., Becam, J., & Goarant, C. (2012). Differential in vivo gene expression of major *Leptospira* proteins in resistant or susceptible animal models. *Applied and Environmental Microbiology*, *78*(17), 6372-6376.

Matsui, M., Roche, L., Geroult, S., Soupé-Gilbert, M. E., Monchy, D., Huerre, M., & Goarant, C. (2016). Cytokine and chemokine expression in kidneys during chronic leptospirosis in reservoir and susceptible animal models. *PLoS ONE*, *11*(5), Article#e0156084.

Matthee, C. A., Burzlaff, J. D., Taylor, J. F., & Davis, S. K. (2001). Mining the mammalian genome for artiodactyl systematics. *Systematic Biology*, *50*(3), 367-390.

Matthias, M. A., Ricaldi, J. N., Cespedes, M., Diaz, M. M., Galloway, R. L., Saito, M., . . . Gilman, R. H. (2008). Human leptospirosis caused by a new, antigenically unique *Leptospira* associated with a *Rattus* species reservoir in the Peruvian Amazon. *PLoS Neglected Tropical Diseases*, *2*(4), Article#e213.

Mayer, W. E., Herrmann, M., & Sommer, R. J. (2007). Phylogeny of the nematode genus *Pristionchus* and implications for biodiversity, biogeography and the evolution of hermaphroditism. *BMC Evolutionary Biology*, *7*(1), Article#104.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., . . . DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297-1303.

Meerburg, B., Singleton, G., & Kijlstra, A. (2009). Rodent-borne diseases and their risks for public health. *Critical Reviews in Microbiology*, *35*(3), 221-270.

Meester, J., & Setzer, H. W. (1971). The mammals of Africa - an identification manual. Washington, DC: Smithsonian Institution Press.

Meester, J., Rautenbach, I. L., Dippenaar, N. J., & Baker, C. M. (1986). Classification of southern African mammals. *Transvaal Museum Monographs*, *5*, 1-359.

Mege, J. L., Meghari, S., Honstettre, A., Capo, C., & Raoult, D. (2006). The two faces of interleukin 10 in human infectious diseases. *The Lancet Infectious Diseases*, *6*(9), 557-569.

Meijaard, E. (2003). Mammals of south-east Asian islands and their Late Pleistocene environments. *Journal of Biogeography*, *30*, 1245-1257.

Meri, T., Murgia, R., Stefanel, P., Meri, S., & Cinco, M. (2005). Regulation of complement activation at the C3-level by serum resistant leptospires. *Microbial Pathogenesis*, 39, 139-147.

Metzker, M. (2010). Sequencing technologies - the next generation. *Nature Reviews Genetics*, *11*, 31-46.

Meyer, A., Kocher, T. D., Basasibwaki, P., & Wilson, A. C. (1990). Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature*, *347*, 550-553.

Michaux, J. R., Libois, R., Ramalhinho, M. G., & Maurois, C. (1998). On the mtDNA restriction patterns variation of the Iberian wood mouse (*Apodemus syluaticus*). Comparison with other west Mediterranean populations. *Hereditas*, *129*(3), 187-194.

Michaux, J. R., Reyes, A., & Catzeflis, F. (2001). Evolutionary history of the most speciose mammals: Molecular phylogeny of muroid rodents. *Molecular Biology and Evolution*, *18*(11), 2017-2031.

Michaux, J. R., Chevret, P., Filippucci, M. G., & Macholan, M. (2002). Phylogeny of the genus *Apodemus* with a special emphasis on the subgenus *Sylvaemus* using the nuclear IRBP gene and two mitochondrial markers: Cytochrome b and 12S rRNA. *Molecular Phylogenetics and Evolution*, *23*(2), 123-136.

Millan, J., Candela, M. G., Lopez-Bao, J. V., Pereira, M., Jimenez, M. A., & Leon-Vizcaino, L. (2008). Leptospirosis in wild and domestic carnivores in natural areas in Andalusia, Spain. *Vector Borne and Zoonotic Diseases*, *9*(5), 549-554.

Miller, S. I., Ernst, R. K., & Bader, M. W. (2005). LPS, TLR4 and infectious disease diversity. *Nature Reviews Microbiology*, *3*, 36-46.

Mogensen, T. H. (2009). Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical Microbiology Reviews*, *22*, 240-273.

Monahan, A. M., Callanan, J. J., & Nally, J. E. (2008). Proteomic analysis of *Leptospira interrogans* shed in urine of chronically infected hosts. *Infection and Immunity*, *76*, 4952-4958.

Monteiro, L. R., Duarte, L. C., & dos Reis, S. F. (2003). Environmental correlates of geographical variation in skull and mandible shape of the punaré rat *Thrichomys apereoides* (Rodentia: Echimyidae). *Journal of Zoology*, *261*, 47-57.

Moore, K. W., de Waal Malefyt, R., Coffman, R. L., & O'Garra, A. (2001). Interleukin-10 and the interleukin-10 receptor. *Annual Review of Immunology*, *19*(1), 683-765.

Morand, S., Bordes, F., Chen, H. W., Claude, J., Cosson, J. F., Galan, M., . . . Ribas, A. (2015). Global parasite and *Rattus* rodent invasions: The consequences for rodent-borne diseases. *Integrative Zoology*, *10*(5), 409-423.

Moreno, L. Z., Miraglia, F., Lilenbaum, W., Neto, J. S., Freitas, J. C., Morais, Z. M., . . . Moreno, A. M. (2016). Profiling of *Leptospira interrogans*, *L. santarosai*, *L. meyeri* and *L. borgpetersenii* by SE-AFLP, PFGE and susceptibility testing-a continuous attempt at species and serovar differentiation. *Emerging Microbes and Infections*, *5*(3), Article#e17.

Morey, R. E., Galloway, R. L., Bragg, S. L., Steigerwalt, A. G., Mayer, L. W., & Levett, P. N. (2006). Species-specific identification of Leptospiraceae by 16S rRNA gene sequencing. *Journal of Clinical Microbiology*, *44*(10), 3510-3516.

Morozova, O., & Marra, M. A. (2008). Applications of next-generation sequencing technologies in functional genomics. *Genomics*, *92*(5), 255-264.

Mosallanejad, B., Ghorbanpour Najafabadi, M., Avizeh, R., & Abdollahpour, G. (2013). A serological survey on leptospiral infection among wild rats (*Rattus rattus*) of Ahvaz district, southwest of Iran: A preliminary study. *Jundishapur Journal of Microbiology*, *6*(10), 66-71.

Mu, Y., Ding, F., Cui, P., Ao, J., Hu, S., & Chen, X. (2010). Transcriptome and expression profiling analysis revealed changes of multiple signaling pathways involved in immunity in the large yellow croaker during *Aeromonas hydrophila* infection. *BMC Genomics*, *11*(1), Article#506.

Murphy, W. J., Eizirik, E., O'brien, S. J., Madsen, O., Scally, M., Douady, C. J., . . . Springer, M. S. (2001). Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science*, *294*(5550), 2348-2351.

Musser, G. G. (1981). Results of the Archbold Expedition No. 105: Notes on systematics of Indo-Malayan murid rodents and descriptions of new genera and species from Ceylon, Sulawesi and the Philippines. *Bulletin of the American Museum of Natural History*, *168*(3), 225-334.

Musser, G. G., & Califia, D. (1982). Identities of rats from Pulau Maratua and other islands of East Borneo. *American Museum Novitates*, *2726*, 1-30.

Musser, G. G., & Newcomb, C. (1983). Malaysian murids and the giant rat of Sumatra. *Bulletin of the American Museum of Natural History*, *174*(4), 327-598.

Musser, G. G. (1986). Sundaic *Rattus*: Definitions of *Rattus baluensis* and *Rattus korinchi*. *American Museum Novitates*, *2862*, 1-24.

Musser, G. G., & Carleton, M. D. (1993). Family Muridae. In D. E. Wilson, & D. M. Reeder (Eds.), Mammal species of the world: A taxonomic and geographic reference (pp. 501-756). Washington, DC: Smithsonian University Press.

Musser, G. G., & Carleton, M. D. (2005). Superfamily Muroidea. In D. E. Wilson, & D. M. Reeder (Eds.), Mammal species of the world: A taxonomic and geographic reference (pp. 894-1531). Baltimore, US: Johns Hopkins University Press.

Musser, G. G., Lunde, D. P., & Son, N. T. (2006). Description of a new genus and species of rodent (Murinae, Muridae, Rodentia) from the tower karst region of northeastern Vietnam. *American Museum Novitates*, *3517*, 1-41.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), Article#853.

Nahori, M. A., Fournié-Amazouz, E., Que-Gewirth, N. S., Balloy, V., Chignard, M., Raetz, C. R., . . . Werts, C. (2005). Differential TLR recognition of leptospiral lipid A and lipopolysaccharide in murine and human cells. *The Journal of Immunology*, *175*(9), 6022-6031.

Nakamura, I., Hang'ombe, B. M., Sawa, H., Kobayashi, S., Orba, Y., Ishii, A., . . . Takada, A. (2013). Cross-reactivity of secondary antibodies against African rodents and application for sero-surveillance. *Journal of Veterinary Medical Science*, *75*(6), 819-825.

Nally, J. E., Grassmann, A. A., Planchon, S., Sergeant, K., Renaut, J., Seshu, J., . . . Caimano, M. J. (2017). Pathogenic leptospires modulate protein expression and post-translational modifications in response to mammalian host signals. *Frontiers in Cellular and Infection Microbiology*, *7*, Article#362.

Natoli, A., Birkun, A., Aguilar, A., Lopez, A., & Hoelzel, A. R. (2005). Habitat structure and the dispersal of male and female bottlenose dolphins (*Tursiops truncatus*). *Proceedings of the Royal Society B: Biological Sciences*, *272*(1569), 1217-1226.

Nedbal, M. A., Allard, M. W., & Honeycutt, R. L. (1994). Molecular systematics of hystricognath rodents: Evidence from the mitochondrial 125 rRNA gene. *Molecular Phylogenetics and Evolution*, *3*(3), 206-220.

Nedbal, M. A., Honeycutt, R. L., & Schilitter, D. A. (1996). Higher-level systematics of rodents (Mammalia, Rodentia): Evidence from the mitochondrial 12S rRNA gene. *Journal of Mammalian Evolution*, *3*(3), 201-237.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA*, *70*, 3321-3323.

Nei, M., & Tajima, F. (1981). DNA polymorphism detectable by restriction endonucleases. *Genetics*, *97*, 145-163.

Nicolas, V., Wendelen, W., Barriere, P., Dudu, A., & Colyn, M. (2008). Morphometric variation in *Hylomyscus alleni* and *H. stella* (Rodentia: Muridae), and description of a new species. *Journal of Mammalogy*, *89*, 222-231.

Nilsson, M. A., Gullberg, A., Spotorno, A. E., Arnason, U., & Janke, A. (2003). Radiation of extant marsupials after the K/T boundary: evidence from complete mitochondrial genomes. *Journal of Molecular Evolution*, *57*(1), S3-S12.

Noguchi, H. (1917). Spirochaeta icterohaemorrhagiae in American wild rats and its relation to the Japanese and European strains: First paper. *Journal of Experimental Medicine*, *25*(5), 755-763.

Nowak, R. M. (1999). Walker's mammals of the world (6th ed.). Baltimore, US: Johns Hopkins University Press.

Nunez, A., Gómez-Villamandos, J. C., Sánchez-Cordón, P. J., de Marco, M. F., Pedrera, M., Salguero, F. J., & Carrasco, L. (2005). Expression of proinflammatory cytokines by hepatic macrophages in acute classical swine fever. *Journal of Comparative Pathology*, *133*(1), 23-32.

Odigie, E. A., Ekeolu, O. K., Asemota, O. D., Uwagie-Ero, A. E., Aighewi, T. I., Ighedosa, . . . Fagboya, E. T. (2017). Comparative non-metric and morphometric analyses of rats at residential halls of the University of Benin campus, Nigeria. *Journal of Infection and Public Health*, *11*(3), 412-417.

O'Grady, N. P., Tropea, M., Preas, H. L., Reda, D., Vandivier, R. W., Banks, S. M., & Suffredini, A. F. (1999). Detection of macrophage inflammatory protein (MIP)-1α and MIP-β during experimental endotoxemia and human sepsis. *The Journal of Infectious Diseases*, *179*(1), 136-141.

Okello, J. B. A., Nyakaana, S., Masembe, C., Siegismund, H. R., & Arctander, P. (2005). Mitochondrial DNA variation of the common hippopotamus: Evidence for a recent population expansion. *Heredity*, *95*(3), 206-215.

Opal, S. M. (2007). The host response to endotoxin, anti-lipopolysaccharide strategies, and the management of severe sepsis. *International Journal of Medical Microbiology*, *297*, 365-377.

Oshida, T., & Masuda, R. (2000). Phylogeny and zoogeography of six squirrel species of the genus *Sciurus* (Mammalia, Rodentia), inferred from cytochrome b gene sequences. *Zoological Science*, *17*(3), 405-410.

Oshida, T., Lin, L. K., Yanagawa, H., Endo, H., & Masuda, R. (2000). Phylogenetic relationships among six flying squirrel genera inferred from mitochondrial cytochrome b gene sequences. *Zoological Science*, *17*(4), 485-490.

Ostfeld, R. S., & Keesing, F. (2000). Biodiversity series: The function of biodiversity in the ecology of vector-borne zoonotic diseases. *Canadian Journal of Zoology*, *78*(12), 2061-2078.

Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. *Science*, *276*(5313), 734-740.

Pagès, M., Chaval, Y., Herbreteau, V., Waengsothorn, S., Cosson, J. F., Hugot, J. P., . . . Michaux, J. (2010). Revisiting the taxonomy of the Rattini tribe: A phylogeny-based delimitation of species boundaries. *BMC Evolutionary Biology*, *10*, Article#184.

Pagès, M., Corbet, G., Orth, A., Volobouev, V., Michaux, J., & Catzeflis, F. (2011). Morphological, chromosomal, and genic differences between sympatric *Rattus rattus* and *Rattus satarae* in south India. *Journal of Mammalogy*, *92*(3), 659-670.

Pagès, M., Bazin, E., Galan, M., Chaval, Y., Claude, J., Herbreteau, V., . . .Cosson, J. F. (2013). Cytonuclear discordance among southeast Asian black rats (*Rattus rattus* complex). *Molecular Ecology*, *22*(4), 1019-1034.

Palasio, R. G. S., de Almeida Guimarães, M. C., Ohlweiler, F. P., & Tuan, R. (2017). Molecular and morphological identification of *Biomphalaria* species from the state of São Paulo, Brazil. *ZooKeys*, *668*, 11-32.

Papa, A., & Kotrotsiou, T. (2015). Cytokines in human leptospirosis. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *109*(12), 749-754.

Paramasvaran, S., Mohd Zain, S. N., Jeffery, J., Selvanesan, S., Hassan, L., Sani, R. A., . . . Amal, N. M. (2013). Distribution and morphological measurements of wild and urban rodents from four habitats in the states of Selangor and Negeri Sembilan, Malaysia. *Malaysian Journal of Veterinary Research (Malaysia)*, *4*(2), 1-12.

Parnell, J. (2013). The biogeography of the Isthmus of Kra region: a review. *Nordic Journal of Botany*, *31*(1), 001-015.

Patel, J. M., Vihol, P. D., Prasad, M. C., Patel, J. H., Raval J. K., & Kalyani, I. H. (2017). Polymerase chain reaction in the diagnosis of spontaneous leptospirosis in bovines. *International Journal of Current Microbiology and Applied Sciences*, *6*(12), 1723-1728.

Patton, J. L., Yang, S. Y., & Myers, P. (1975). Genetic and morphological divergence among introduced rat populations (*Rattus rattus*) of the Galapagos Archipelago, Ecuador. *Systematic Zoology*, *24*, 296-310.

Patton, J. L., & Brylski, P. V. (1987). Pocket gophers in alfalfa fields: Causes and consequences of habitat-related body size variation. *The American Naturalist*, *130*, 493-506.

Patwardhan, A., Ray, S., & Roy, A. (2014). Molecular markers in phylogenetic studies-a review. *Journal of Phylogenetics and Evolutionary Biology*, *2*(2), Article#131.

Pavlinov, I. Y. (2005). Systematic of recent mammals (2nd ed.). Moskva, RU: Izdatel'stvo Moskovskogo Universiteta.

Pergams, O. R., & Ashley, M. V. (2001). Microevolution in island rodents. *Genetica*, *112*(1), 245-256.

Pergams, O. R., Barnes, W. M., & Nyberg, D. (2003). Mammalian microevolution: rapid change in mouse mitochondrial DNA. *Nature*, *423*(6938), Article#397.

Pergams, O. R., & Lacy, R. C. (2008). Rapid morphological and genetic change in Chicago-area *Peromyscus*. *Molecular Ecology*, *17*(1), 450-463.

Pergams, O. R., & Kareiva, P. (2009). Support services: a focus on genetic diversity. *Generations*, *7*(5), Article#6.

Pergams, O. R., Byrn, D., Lee, K. L. Y., & Jackson, R. (2015). Rapid morphological change in black rats (*Rattus rattus*) after an island introduction. *PeerJ*, *3*, Article#e812.

Perolat, P., Chappel, R. J., Adler, B., Baranton, G., Bulach, D. M., Billinghurst, M. L., . . . Serrano, M. S. (1998). *Leptospira fainei* sp. nov., isolated from pigs in Australia. *International Journal of Systematic Bacteriology*, *48*(3), 851-858.

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, *33*(3), 290-295.

Petrovic-Djergovic, D., Popovic, M., Chittiprol, S., Cortado, H., Ransom, R. F., & Partida-Sánchez, S. (2015). CXCL10 induces the recruitment of monocyte-derived macrophages into kidney, which aggravate puromycin aminonucleoside nephrosis. *Clinical and Experimental Immunology*, *180*(2), 305-315.

Philippart, F., & Cavaillon, J. M. (2007). Sepsis mediators. *Current Infectious Disease Reports*, *9*, 358-365.

Picardeau, M. (2013). Diagnosis and epidemiology of leptospirosis. *Médecine et Maladies Infectieuses*, *43*(1), 1-9.

Pimsai, U., Pearch, M. J., Satasook, C., Bumrungsri, S., & Bates, P. J. (2014). Murine rodents (Rodentia: Murinae) of the Myanmar-Thai-Malaysian Peninsula and Singapore: Taxonomy, distribution, ecology, conservation status, and illustrated identification keys. *Bonn Zoological Bulletin*, *63*, 15-114.

Plata-Luis, J., Foronda, P., Martín-Alonso, A., Feliu, C., Alves, J., Gil, H., & Valladares, B. (2016). *Leptospira interrogans* in rodents from Cape Verde. *Vector-Borne and Zoonotic Diseases*, *16*(11), 731-733.

Polly, P. D. (2007). Phylogeographic differentiation in *Sorex araneus*: Morphology in relation to geography and karyotype. *Russian Journal of Theriology*, *6*, 73-84.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., . . . Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595-609.

Posada, D., & Crandall, K. A. (1998). Modeltest: Testing the model of DNA substitution. *Bioinformatics*, *14*(9), 817-818.

Post, J. C., & Ehrlich, G. D. (2000). The impact of the polymerase chain reaction in clinical medicine. *The Journal of the American Medical Association*, *283*(12), 1544-1546.

Postic, D., Merien, F., Perolat, P., & Baranton, G. (2000). Biological diagnosis. Leptospirosis-Lyme Borreliosis (2nd ed.). Paris, FR: Collection des Laboratoires de Reference et d'Expertise.

Pui, C. F., Bilung, L. M., Apun, K., & Su'ut, L. (2017). Diversity of *Leptospira* spp. in rats and environment from urban areas of Sarawak, Malaysia. *Journal of Tropical Medicine*, *2017*, Article#3760674.

Punyadeera, C., Schneider, E. M., Schaffer, D., Hsu, H. Y., Joos, T. O., Kriebel, F., . . . Verhaegh, W. F. (2010). A biomarker panel to discriminate between systemic

inflammatory response syndrome and sepsis and sepsis severity. *Journal of Emergencies, Trauma and Shock*, *3*(1), 26-35.

Que-Gewirth, N. L., Ribeiro, A. A., Kalb, S. R., Cotter, R. J., Bulach, D. M., Adler, B., . . . Raetz, C. R. (2004). A methylated phosphate group and four amide-linked acyl chains in *Leptospira interrogans* lipid A. The membrane anchor of an unusual lipopolysaccharide that activates TLR2. *Journal of Biological Chemistry*, *279*(24), 25420-25429.

Quezada-Euán, J. J. G., Paxton, R. J., Palmer, K. A., de May-Itzá, W. J., Tay, W. T., & Oldroyd, B. P. (2007). Morphological and molecular characters reveal differentiation in a neotropical social bee, *Melipona beecheii* (Apidae: Meliponini). *Apidologie*, *38*(3), 247-258.

R Core Team. (2015). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. Retrieved on May 18, 2018 from http://www.R-project.org/.

Rambaut, A. (2009). TRACER v1. 5. Retrieved on March 4, 2018 from http://tree. bio. ed. ac. uk/software/tracer/.

Reis, E. A., Hagan, J. E., Ribeiro, G. S., Teixeira-Carvalho, A., Martins-Filho, O. A., Montgomery, R. R., . . . Reis, M. G. (2013). Cytokine response signatures in disease progression and development of severe clinical outcomes for leptospirosis. *PLoS Neglected Tropical Diseases*, *7*(9), Article#e2457.

Reis, S. F., Duarte, L. C., Monteiro, L. R., & Von Zuben, F. J. (2002). Geographic variation in cranial morphology in *Thrichomys apereoides* (Rodentia: Echimyidae). I. Geometric descriptors and patterns of variation in shape. *Journal of Mammalogy*, *83*(2), 333-344.

Renaud, S., & Millien, V. (2001). Intra- and interspecific morphological variation in the field mouse species *Apodemus argenteus* and *A. speciosus* in the Japanese Archipelago: The role of insular isolation and biogeographic gradients. *Biological Journal of the Linnean Society*, *74*, 557-569.

Rettinger, A., Krupka, I., Grünwald, K., Dyachenko, V., Fingerle, V., Konrad, R., . . . Huber, I. (2012). *Leptospira* spp. strain identification by MALDI TOF MS is an equivalent tool to 16S rRNA gene sequencing and multi locus sequence typing (MLST). *BMC Microbiology*, *12*(1), Article#185.

Richard, A. B., Allan, R. W., Ray, B., Thomas, P. M., & Alex, D. (2018). Maps: Draw geographical maps. R package version 3.3.0.

Rios, E., & Álvarez-Castañeda, S. T. (2012). Pelage colour variation in pocket gophers (Rodentia: Geomyidae) in relation to sex, age and differences in habitat. *Mammalian Biology-Zeitschrift für Säugetierkunde*, *77*(3), 160-165.

Robalino, J., Almeida, J. S., McKillen, D., Colglazier, J., Trent III, H. F., Chen, Y. A., . . . Gross, P. S. (2007). Insights into the immune transcriptome of the shrimp *Litopenaeus vannamei*: tissue-specific expression profiles and transcriptomic responses to immune challenge. *Physiological Genomics*, *29*(1), 44-56.

Robins, J. H., Hingston, M., Matisoo-Smith, E., & Ross, H. (2007). Identifying *Rattus* species using mitochondrial DNA. *Molecular Ecology Notes*, *7*, 717-729.

Robins, J. H., McLenachan, P. A., Phillips, M. J., Craig, L., Ross, H. A., & Matisoo-Smith, E. (2008). Dating of divergences within the *Rattus* genus phylogeny using whole mitochondrial genomes. *Molecular Phylogenetics and Evolution*, *49*(2), 460-466.

Robins, J. H., McLenachan, P. A., Phillips, M. J., McComish, B. J., Matisoo-Smith, E., & Ross, H. A. (2010). Evolutionary relationships and divergence times among the native rats of Australia. *BMC Evolutionary* Biology, *10*(1), Article#375.

Rogers, A. R., & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, *9*, 552-569.

Rogers, A. R. (1995). Genetic evidence for a Pleistocene population explosion. *Evolution*, *49*(4), 608-615.

Rohlf, F. J. (1990). Morphometrics. *Annual Review of Ecology and Systematics*, *21*(1), 299-316.

Rohlf, F. J., & Marcus, L. F. (1993). A revolution morphometrics. *Trends in Ecology and Evolution*, *8*(4), 129-132.

Rosli, M. K. A., Syed-Shabthar, S. M. F., Abdul-Patah, P., Abdul-Samad, Z., Abdul, S. N., Burhanuddin, M. N., . . . Sekiguchi, T. (2014). A new subspecies identification and population study of the Asian small-clawed otter (*Aonyx cinereus*) in Malay Peninsula and southern Thailand based on fecal DNA method. *The Scientific World Journal*, *2014*, Article#457350.

Ross, H. A., Lento, G. M., Dalebout, M. L., Goode, M., Ewing, G., McLaren, P., . . . Baker, C. S. (2003). DNA surveillance: Web-based molecular identification of whales, dolphins, and porpoises. *Journal of Heredity*, *94*(2), 111-114.

Rowe, K. C., Aplin, K. P., Baverstock, P. R., & Moritz, C. (2011). Recent and rapid speciation with limited morphological disparity in the genus *Rattus*. *Systematic Biology*, *60*(2), 188-203.

Rueda, M., Rodríguez, M. Á., & Hawkins, B. A. (2013). Identifying global zoogeographical regions: lessons from Wallace. *Journal of Biogeography*, *40*(12), 2215-2225.

Rumbo-Feal, S., Gómez, M. J., Gayoso, C., Álvarez-Fraga, L., Cabral, M. P., Aransay, A. M., . . . Bou, G. (2013). Whole transcriptome analysis of *Acinetobacter baumannii* assessed by RNA-sequencing reveals different mRNA expression profiles in biofilm compared to planktonic cells. *PLoS ONE*, *8*(8), Article#e72968.

Saccone, C., De Giorgi, C., Gissi, C., Pesole, G., & Reyes, A. (1999). Evolutionary genomics in Metazoa: The mitochondrial DNA as a model system. *Gene*, *238*(1), 195-209.

Saito, M., Villanueva, S. Y., Chakraborty, A., Miyahara, S., Segawa, T., Asoh, T., . . . Yoshida, S. I. (2013). Comparative analysis of *Leptospira* strains isolated from environmental soil and water in the Philippines and Japan. *Applied and Environmental Microbiology*, *79*(2), 601-609.

Sakuma, Y., Ranorosoa, M. C., Kinoshita, G., Shimoji, H., Tsuchiya, K., Ohdachi, S. D., . . . Suzuki, H. (2016). Variation in the coat-color-controlling genes, Mc1r and Asip, in the house mouse *Mus musculus* from Madagascar. *Mammal Study*, *41*(3), 131-141.

Sambrook, J. F., & Russell, D.W. (2001). Molecular cloning: A laboratory manual (3rd ed.). New York, NY: Cold Spring Harbor Laboratory Press.

Scheffé, H. (1959). The analysis of variance. New York, NY: John Wiley & Sons.

Schlager, S. (2017). Morpho and Rvcg-shape analysis in R. In G. Zheng, S. Li, & G. Szekely (Eds.), Statistical shape and deformation analysis (pp. 217-256). London, UK: Academic Press.

Schlegel, M., Ali, H. S., Stieger, N., Groschup, M. H., Wolf, R., & Ulrich, R. G. (2012). Molecular identification of small mammal species using novel cytochrome b gene-derived degenerated primers. *Biochemical Genetics*, *50*(5-6), 440-447.

Schlick, N. E., Jensen-Seaman, M. I., Orlebeke, K., Kwitek, A. E., Jacob, H. J., & Lazar, J. (2006). Sequence analysis of the complete mitochondrial DNA in 10 commonly used inbred rat strains. *American Journal of Physiology-Cell Physiology*, *291*(6), C1183-C1192.

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols*, *3*(6), Article#1101.

Schneider, S., & Excoffier, L. (1999). Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. *Genetics*, *152*(3), 1079-1089.

Schneider, S., Roessli, D., & Excoffier, L. (2000). Arlequin: A software for population genetics data analysis, version 2.000. Switzerland, CH: Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva.

Schulte, W., Bernhagen, J., & Bucala, R. (2013). Cytokines in sepsis: Potent immunoregulators and potential therapeutic targets—an updated view. *Mediators of Inflammation*, *2013*, Article#165974.

Seehausen, O. (2004). Hybridization and adaptive radiation. *Trends in Ecology and Evolution*, *4*, 198-207.

Serizawa, K., Suzuki, H., & Tsuchiya, K. (2000). A phylogenetic view on species radiation in *Apodemus* inferred from variation of nuclear and mitochondrial genes. *Biochemical Genetics*, *38*(1-2), 27-40.

Sfeir, T., Saha, D. C., Astiz, M., & Rackow, E. C. (2001). Role of interleukin-10 in monocyte hyporesponsiveness associated with septic shock. *Critical Care Medicine*, *29*(1), 129-133.

Shen, S., Park, J. W., Lu, Z. X., Lin, L., Henry, M. D., Wu, Y. N., . . . Xing, Y. (2014). rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proceedings of the National Academy of Sciences*, *111*(51), E5593-E5601.

Silva, E. F., Medeiros, M. A., McBride, A. J., Matsunaga, J., Esteves, G. S., Ramos, J. G., . . . Haake, D. A. (2007). The terminal portion of leptospiral immunoglobulin-like protein LigA confers protective immunity against lethal infection in the hamster model of leptospirosis. *Vaccine*, *25*(33), 6277-6286.

Silva, É. F., Santos, C. S., Athanazio, D. A., Seyffert, N., Seixas, F. K., Cerqueira, G. M., . . . Ko, A. I. (2008). Characterization of virulence of *Leptospira* isolates in a hamster model. *Vaccine*, *26*(31), 3892-3896.

Singleton, G. (2003). Impacts of rodents on rice production in Asia. *IRRI Discussion Paper Series*, *45*, Article#30.

Singleton, G. R., Hinds, L. A., Krebs, C. J., & Spratt, D. M. (2003). Rats, mice and people: An interwoven relationship. Rats, mice and people: Rodent biology and management. *ACIAR Monograph*, *96*, 9-10.

Slack, A. T., Kalambaheti, T., Symonds, M. L., Dohnt, M. F., Galloway, R. L., Steigerwalt, A. G., . . . Smythe, L. D. (2008). *Leptospira wolffii* sp. nov., isolated from a human with suspected leptospirosis in Thailand. *International Journal of Systematic and Evolutionary Microbiology*, *58*(10), 2305-2308.

Slack, A. T., Khairani-Bejo, S., Symonds, M. L., Dohnt, M. F., Galloway, R. L., Steigerwalt, A. G., . . . Smythe, L. D. (2009). *Leptospira kmetyi* sp. nov., isolated from an environmental source in Malaysia. *International Journal of Systematic and Evolutionary Microbiology*, *59*(4), 705-708.

Slatkin, M., & Hudson, R. R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, *129*, 555-562.

Smith, C. G., Turner, L. H., Harrison, J. L., & Broom, J. C. (1961). Animal leptospirosis in Malaya: Methods, zoogeographical background, and broad analysis of results. *Bulletin of the World Health Organization*, *24*(1), Article#5.

Smits, B. M., & Cuppen, E. (2006). Rat genetics: The next episode. *Trends in Genetics*, *22*(4), 232-240.

Springer, M. S., Teeling, E. C., Madsen, O., Stanhope, M. J., & de Jong, W. W. (2001). Integrated fossil and molecular data reconstruct bat echolocation. *Proceedings of the National Academy of Sciences*, *98*(11), 6241-6246.

Sriskandan, S., & Altmann, D. M. (2008). The immunology of sepsis. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, *214*(2), 211-223.

Stanhope, M. J., Waddell, V. G., Madsen, O., De Jong, W., Hedges, S. B., Cleven, G. C., . . . Springer, M. S. (1998). Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proceedings of the National Academy of Sciences*, *95*(17), 9967-9972.

Steppan, S. J., Akhverdyan, M. R., Lyapunova, E. A., Fraser, D. G., Vorontsov, N. N., Hoffmann, R. S., & Braun, M. J. (1999). Molecular phylogeny of the marmots (Rodentia: Sciuridae): Tests of evolutionary and biogeographic hypotheses. *Systematic Biology*, *48*(4), 715-734.

Steppan, S. J., Zawadzki, C., & Heaney, L. R. (2003). Molecular phylogeny of the endemic Philippine rodent *Apomys* (Muridae) and the dynamics of diversification

in an oceanic archipelago. *Biological Journal of the Linnean Society*, *80*(4), 699-715.

Steppan, S. J., Storz, B. L., & Hoffmann, R. S. (2004). Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1. *Molecular Phylogenetics and Evolution*, *30*(3), 703-719.

Steppan, S. J., & Schenk, J. J. (2017). Muroid rodent phylogenetics: 900-species tree reveals increasing diversification rates. *PLoS ONE*, *12*(8), Article#e0183070.

Sumanta, H., Wibawa, T., Hadisusanto, S., Nuryati, A., & Kusnanto H. (2015). Spatial analysis of *Leptospira* in rats, water and soil in Bantul District Yogyakarta. *Open Journal of Epidemiology*, *5*, 22-31.

Sun, G. L., Díaz, O., Salomon, B., & von Bothmer, R. (1999). Genetic diversity in *Elymus caninus* as revealed by isozyme, RAPD, and microsatellite markers. *Genome*, *42*(3), 420-431.

Suzán, G., Marcé, E., Giermakowski, J. T., Mills, J. N., Ceballos, G., Ostfeld, R. S., . . . Yates, T. L. (2009). Experimental evidence for reduced rodent diversity causing increased hantavirus prevalence. *PLoS ONE*, *4*(5), Article#e5461.

Suzuki, H., Tsuchiya, K., & Takezaki, N. (2000). A molecular phylogenetic framework for the Ryukyu endemic rodents *Tokudaia osimensis* and *Diplothrix legata*. *Molecular Phylogenetics and Evolution*, *15*(1), 15-24.

't Hoen, P. A. C., Friedländer, M. R., Almlöf, J., Sammeth, M., Pulyakhina, I., Anvar, S. Y., . . . van Ommen, G. J. B. (2013). Reproducibility of high-throughput mRNA and small RNA sequencing across laboratories. *Nature Biotechnology*, *31*(11), Article#1015.

Tajiki, M. H., & Salomão, R. (1996). Association of plasma levels of tumor necrosis factor α with severity of disease and mortality among patients with leptospirosis. *Clinical Infectious Diseases*, *23*(5), 1177-1178.

Tajiki, M. H., Satie, A. N., & Salomão, R. (1997). The ratio of plasma levels of IL-10/TNF-alpha and its relationship to disease severity and survival in patients with leptospirosis. *The Brazilian Journal of Infectious Diseases: An Official Publication of the Brazilian Society of Infectious Diseases*, *1*(3), 138-141.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585-595.

Takahashi, A., Kato, T., & Ota, H. (2007). A new species of the genus *Geoemyda* (Chelonii: Geoemydidae) from the Upper Pleistocene of Tokunoshima Island, the central Ryukyus, Japan. *Current Herpetology*, *26*(1), 1-11.

Tamrin, N. A. M., & Abdullah, M. T. (2011). Molecular phylogenetics and systematic of five genera of Malaysian murine rodents (*Maxomys*, *Sundamys*, *Leopoldamys*, *Niniventer,* and *Rattus*) inferred from partial mitochondrial cytochrome c oxidase subunit I (COI) gene. *Journal of Science and Technology in the Tropics*, *7*, 75-86.

Tangkanakul, W., Tharmaphornpil, P., Plikaytis, B. D., Bragg, S., Poonsuksombat, D., Choomkasien, P., . . . Ashford, D. A. (2000). Risk factors associated with leptospirosis in northeastern Thailand, 1998. *The American Journal of Tropical Medicine and Hygiene*, *63*(3), 204-208.

Tansuphasiri, U., Chanthadee, R., Phulsuksombati, D., & Sangjun, N. (2006). Development of a duplex-polymerase chain reaction for rapid detection of pathogenic *Leptospira*. *Southeast Asian Journal of Tropical Medicine and Public Health*, *37*, 297-308.

Tassanakajon, A., Klinbunga, S., Paunglarp, N., Rimphanitchayakit, V., Udomkit, A., Jitrapakdee, S., . . . Tang, S. (2006). *Penaeus monodon* gene discovery project: The generation of an EST collection and establishment of a database. *Gene*, *384*, 104-112.

Taylor, P. J., Arntzen, L., Hayter, M., Iles, M., Frean, J., & Belmain, S. (2008). Understanding and managing sanitary risks due to rodent zoonoses in an African city: Beyond the Boston Model. *Integrative Zoology*, *3*, 38-50.

Thorpe, R. S. (1976). Biometric analysis of geographic variation and racial affinities. *Biological Reviews*, *51*, 407-452.

Tian, Y. C., Chen, Y. C., Hung, C. C., Chang, C. T., Wu, M. S., Phillips, A. O., & Yang, C. W. (2006). Leptospiral outer membrane protein induces extracellular matrix accumulation through a TGF-β1/Smad-dependent pathway. *Journal of The American Society of Nephrology*, *17*(10), 2792-2798.

Tisoncik, J. R., Korth, M. J., Simmons, C. P., Farrar, J., Martin, T. R., & Katze, M. G. (2012). Into the eye of the cytokine storm. *Microbiology and Molecular Biology Reviews*, *76*(1), 16-32.

Tollenaere, C., Brouat, C., Duplantier, J. M., Rahalison, L., Rahelinirina, S., Pascal, M., . . . Cosson, J. F. (2010). Phylogeography of the introduced species *Rattus rattus* in the western Indian Ocean, with special emphasis on the colonization history of Madagascar. *Journal of Biogeography*, *37*(3), 398-410.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., . . . Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, *7*(3), 562-578.

Trewick, S. A. (2000). Mitochondrial DNA sequences support allozyme evidence for cryptic radiation of New Zealand Peripatoides (Onychophora). *Molecular Ecology*, *9*(3), 269-281.

Truong, T. T., Yoshimatsu, K., Araki, K., Lee, B. H., Nakamura, I., Endo, R., . . . Okumura, M. (2009). Molecular epidemiological and serological studies of hantavirus infection in northern Vietnam. *Journal of Veterinary Medical Science*, *71*(10), 1357-1363.

Tucunduva de Faria, M., Athanazio, D. A., Goncalves Ramos, E. A., Silva, E. F., Reis, M. G., & Ko, A. I. (2007). Morphological alterations in the kidney of rats with natural and experimental *Leptospira* infection. *Journal of Comparative Pathology*, *137*(4), 231-238.

Turner, M. D., Nedjai, B., Hurst, T., & Pennington, D. J. (2014). Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1843*(11), 2563-2582.

Udvardy, M. D. (1975). A classification of the biogeographical provinces of the world (Vol. 8). Morges, Switzerland: International Union for Conservation of Nature and Natural Resources.

Ullmann, L. S., & Langoni, H. (2011). Interactions between environment, wild animals and human leptospirosis. *Journal of Venomous Animals and Toxins Including Tropical Diseases*, *17*, 119-129.

van Dissel, J. T., Van Langevelde, P., Westendorp, R. G., Kwappenberg, K., & Frölich, M. (1998). Anti-inflammatory cytokine profile and mortality in febrile patients. *The Lancet*, *351*(9107), 950-953.

van Peenen, P. F., Ryan, P. F., & Light, R. H. (1969). Preliminary identification manual for mammals of south Vietnam. Washington, DC: Smithsonian Institution.

Vanasco, N. B., Sequeira, M. D., Sequeira, G., & Tarabla, H. D. (2003). Associations between leptospiral infection and seropositivity in rodents and environmental characteristics in Argentina. *Preventive Veterinary Medicine*, *60*(3), 227-235.

Vanasco, N. B., Jacob, P., Landolt, N., Chiani, Y., Schmeling, M. F., Cudos, C., . . . Lottersberger, J. (2016). Diagnostic accuracy of an IgM enzyme-linked immunosorbent assay and comparison with 2 polymerase chain reactions for early

diagnosis of human leptospirosis. *Diagnostic Microbiology and Infectious Disease*, *84*, 292-297.

Venables, W. N., & Ripley, B. D. (2002). Modern applied statistics (4th ed.). New York, NY: Springer-Verlag.

Ventura, J., & Lopez-Fuster, M. J. (2000). Morphometric analysis of the black rat, *Rattus rattus*, from Congreso Island (Chafarinas Archipelago, Spain). *Orsis*, *15*, 91-102.

Verneau, O., Catzeflis, F., & Furano, A. V. (1998). Determining and dating recent rodent speciation events by using L1 (LINE-1) retrotransposons. *Proceedings of the National Academy of Sciences*, *95*(19), 11284-11289.

Vernel-Pauillac, F., & Merien, F. (2006). Proinflammatory and immunomodulatory cytokine mRNA time course profiles in hamsters infected with a virulent variant of *Leptospira interrogans*. *Infection and Immunity*, *74*(7), 4172-4179.

Vernel-Pauillac, F., & Goarant, C. (2010). Differential cytokine gene expression according to outcome in a hamster model of leptospirosis. *PLoS Neglected Tropical Diseases*, *4*, Article#e582.

Victoriano, A. F. B., Smythe, L. D., Gloriani-Barzaga, N., Cavinta, L. L., Kasai, T., Limpakarnjanarat, K., . . .Yanagihara, Y. (2009). Leptospirosis in the Asia Pacific region. *BMC Infectious Diseases*, *9*(1), Article#147.

Vielhauer, V., Berning, E., Eis, V., Kretzler, M., Segerer, S., Strutz, F., . . . Anders, H. J. (2004). CCR1 blockade reduces interstitial inflammation and fibrosis in mice with glomerulosclerosis and nephrotic syndrome. *Kidney International*, *66*(6), 2264-2278.

Vilariño-Güell, C., Wider, C., Ross, O. A., Dachsel, J. C., Kachergus, J. M., Lincoln, S. J., . . . Behrouz, B. (2011). VPS35 mutations in Parkinson disease. *The American Journal of Human Genetics*, *89*(1), 162-167.

Voris, H. K. (2000). Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems and time durations. *Journal of Biogeography*, *27*, 1153-1167.

Wada, S., Oishi, M., & Yamada, T. K. (2003). A newly discovered species of living baleen whale. *Nature*, *426*, 278-281.

Wallace, A. R. (1869). Sir Charles Lyell on geological climates and the origin of species. *Quarterly Review*, *126*(252), 359-394.

Wallace, A. R. (1876). The geographical distribution of animals. London: Macmillan and Co.

Wang, B., Sullivan, J. A., Sullivan, G. W., & Mandell, G. L. (1984). Role of specific antibody in interaction of leptospires with human monocytes and monocyte-derived macrophages. *Infection and Immunity*, *46*, 809-813.

Watthanakaiwan, V., Klakhang, S., Promkerd, P., & Keawta, S. (2015). A study of genetic diversity in ricefield rat (*Rattus argentiventer* (Robinson & Kloss, 1916)) in Thailand based on mitochondrial cytochrome b gene. *The 12th National Plant Protection Conference*, *2015*, Article#96.

Werts, C., Tapping, R. I., Mathison, J. C., Chuang, T. H., Kravchenko, V., Saint Girons, I., . . . Underhill, D. M. (2001). Leptospiral lipopolysaccharide activates cells through a TLR2-dependent mechanism. *Nature Immunology*, *2*(4), 346-352.

Werts, C. (2017). Interaction of *Leptospira* with the innate immune system. In B. Adler (Ed.), Spirochete biology: The post genomic era (pp. 163-187). New York, NY: Springer.

White, T. J., Bruns, T., Lee, S. J. W. T., & Taylor, J. L. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols: A Guide to Methods and Applications*, *18*(1), 315-322.

Wickens, A. P. (1998). The causes of aging. Amsterdam, NL: Harwood Academic.

Wilson, D. E., & Reeder, D. M. (2005). Mammal species of the world: A taxonomic and geographic reference (3rd ed.). Baltimore, US: Johns Hopkins University Press.

Woese, C. R., & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences*, *74*(11), 5088-5090.

Woese, C. R., Stackebrandt, E., Macke, T. J., & Fox, G. E. (1985). A phylogenetic definition of the major eubacterial taxa. *Systematic and Applied Microbiology*, *6*, 143-151.

Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews*, *51*(2), Article#221.

Woodruff, D. S., & Turner, L. M. (2009). The Indochinese-Sundaic zoogeographic transition: a description and analysis of terrestrial mammal species distributions. *Journal of Biogeography*, *36*(5), 803-821.

Woodruff, D. S. (2010). Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodiversity and Conservation*, *19*(4), 919-941.

World Health Organization (WHO). (2003). Human leptospirosis: guidance for diagnosis, surveillance and control. *WHO Library*, *45*, 1-109.

World Health Organization (WHO). (2010). Report of the first meeting of the leptospirosis. Burden epidemiology reference group, Geneva: World Health Organization.

Xue, F., Dong, H., Wu, J., Wu, Z., Hu, W., Sun, A., . . . Yan, J. (2010). Transcriptional responses of *Leptospira interrogans* to host innate immunity: Significant changes in metabolism, oxygen tolerance, and outer membrane. *PLoS Neglected Tropical Diseases*, *4*(10), Article#e857.

Xue, F., Zhao, X., Yang, Y., Zhao, J., Yang, Y., Cao, Y., . . . Gu, J. (2013). Responses of murine and human macrophages to leptospiral infection: A study using comparative array analysis. *PLoS Neglected Tropical Diseases*, *7*(10), Article#e2477.

Yang, C. W., Wu, M. S., & Pan, M. J. (2001). Leptospirosis renal disease. *Nephrology Dialysis Transplantation*, *16*(5), 73-77.

Yasuda, S. P., Gamage, C. D., Koizumi, N., Nishio, S., Isozumi, R., Shimizu, K., . . . Arikawa, J. (2014). Distinct genetic characteristics of Sri Lankan *Rattus* and *Bandicota* (Murinae, Rodentia) inferred from mitochondrial and nuclear markers. *Genes and Genetic Systems*, *89*(2), 71-80.

Yiğit, N., Çolak, E., Sözen, M., Özkurt, S. (1998). The taxonomy and karyology of *Rattus norvegicus* (Berkenhout, 1769) and *Rattus rattus* (Linnaeus, 1758) (Rodentia: Muridae) in Turkey. *Turkish Journal of Zoology*, *22*, 203-212.

Yong, H. S. (1969). Karyotypes of Malayan rats (Rodentia-Muridae, genus *Rattus* Fischer). *Chromosoma*, *27*(3), 245-267.

Yosida, T. H., Tsuchiya, K., & Moriwaki, K. (1971). Karyotypic differences of black rats, *Rattus rattus*, collected in various localities of east and southeast Asia and Oceania. *Chromosoma*, *33*(3), 252-267.

Yosida, T. H., Kato, H., Tsuchiya, K., Sagai, T., & Moriwaki, K. (1974). Cytogenetical survey of black rats, *Rattus rattus*, in southwest and Central Asia, with special regard to the evolutional relationship between three geographical types. *Chromosoma*, *45*(1), 99-109.

Yosida, T. H. (1977). Supernumerary chromosomes in the black rat (*Rattus rattus*) and their distribution in three geographic variants. *Cytogenetic and Genome Research*, *18*(3), 149-159.

Yosida, T. H. (1978). A preliminary note on silver-stained nucleolar organizer regions in the black and Norway rats. *Proceedings of the Japan Academy, Series B*, *54*(7), 353-358.

Yosida, T. H. (1979). The karyotype of *Rattus villosissimus* resulted from Robertsonian fission of small metacentrics. *Proceedings of the Japan Academy, Series B*, *55*(10), 497-501.

Yosida, T. H. (1980). Cytogenetics of the black rat: Karyotype evolution and species differentiation. Tokyo, JP: University of Tokyo Press.

Yosida, T. H., Udagawa, T., Ishibashi, M., Moriwaki, K., Yabe, T., & Hamada, T. (1985). Studies on the karyotypes of the black rats distributed in the Pacific and south Pacific islands, with special regard to the border line of the Asian and Oceanian type black rats on the Pacific Ocean. *Proceedings of the Japan Academy, Series B*, *61*(2), 71-74.

Yue, G., & Wang, L. (2017). Current status of genome sequencing and its applications in aquaculture. *Aquaculture*, *468*, 337-347.

Zarei, B., Aliabadian, M., Drvish, J., Moghadam, F. Y., & Mohammadi, S. (2013). Geometric morphometric analyses of the short-tailed Bandicoot Rat (*Nesokia indica*) (Rodentia: Muridae) in the north and southeast of Iran. *Zoology and Ecology*, *23*(4), 260-265.

Zelditch, M. L., Bookstein, F. L., & Lundrigan, B. L. (1992). Ontogeny of integrated skull growth in the cotton rat *Sigmodon fulviventer*. *Evolution*, *46*, 1164-1180.

Zhang, P., Chen, Y. Q., Liu, Y. F., Zhou, H., & Qu, L. H. (2003). The complete mitochondrial genome of the Chinese giant salamander, *Andrias davidianus* (Amphibia: Caudata). *Gene*, *311*, 93-98.

Zhang, R., Li, Q., Chuang, P. Y., Lu, G., Liu, R., Yang, J., . . . He, J. C. (2013). Regulation of pathogenic Th17 cell differentiation by IL-10 in the development of glomerulonephritis. *The American Journal of Pathology*, *183*(2), 402-412.

Zuckerkandl, E., & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In V. Bryson, & H. J. Vogel (Eds.), Evolving genes and proteins (pp. 97-166). New York, NY: Academic Press.

## LIST OF PUBLICATIONS AND PAPERS PRESENTED

### PUBLICATIONS

1. **Mohamad Ikbal, N. H.,** Bhassu, S., Simarani, K., Uni, S., Chan, C. C., & Omar, H. (2019). Prevalence and genetic divergence of *Leptospira interrogans* and *Leptospira borgpetersenii* in house rats (*Rattus rattus*) from Peninsular Malaysia. *Asian Pacific Journal of Tropical Medicine*. (Accepted).

2. **Mohamad Ikbal, N. H.,** Pathmanathan, D., Bhassu, S., Simarani, K., & Omar, H. (2019). Morphometric analysis of craniodental characters of the house rat, *Rattus rattus* (Rodentia: Muridae) in Peninsular Malaysia. *Sains Malaysiana*. (Accepted).

### PAPERS PRESENTED

1. **Mohamad Ikbal, N. H.**, Omar, H., Simarani, K., Chan, C. C., & Bhassu, S. (2018). *A transcriptome study on Rattus rattus infected with pathogenic Leptospira spp.* Paper presented at the University of Malaya-Indonesian Universities Symposium (UMInd 2018), 8-9 November 2018, Kuala Lumpur, Malaysia.

2. **Mohamad Ikbal, N. H.**, Bhassu, S., Simarani, K., & Omar, H. (2017). *Cytochrome b gene based phylogenetic studies of the Rattus rattus species complex from Peninsular Malaysia*. Paper presented at the South Asian Conference on Small Mammals, 27-29 August 2017, Kathmandu, Nepal.

3. **Mohamad Ikbal, N. H.**, Bhassu, S., Simarani, K., & Omar, H. (2017). *Genetic diversity and population structure of black rat (Rattus rattus) populations in Peninsular Malaysia*. Paper presented at the South Asian Conference on Small Mammals, 27-29 August 2017, Kathmandu, Nepal.