



Faculty of Resource Science and Technology

**Prevalence of Potentially Pathogenic Bacteria in Rodent Populations
Near Natural Water Bodies From Sarawak**

Muhammad Amin Iman bin Azmi

**Master of Science
2024**

Prevalence of Potentially Pathogenic Bacteria in Rodent Populations Near
Natural Water Bodies From Sarawak

Muhammad Amin Iman bin Azmi

A thesis submitted

In fulfillment of the requirements for the degree of Master of Science

(Genetics)

Faculty of Resource Science and Technology

UNIVERSITI MALAYSIA SARAWAK

2024

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.



.....

Signature

Name: Muhammad Amin Iman bin Azmi

Matric No.: 21020455

Faculty of Resource Science and Technology

Universiti Malaysia Sarawak

Date: 15/4/2024

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my supervisor, Associate Professor Dr. Faisal Ali bin Anwarali Khan. He has given me numerous opportunities to expand my knowledge and has supported me throughout my master's journey. His valuable insight and willingness to aid me have allowed me to carry out my research smoothly. Furthermore, I would like to pay my gratitude to Associate Professor Dr. Tay Meng Guan for his guidance and knowledge on water quality sampling. I would also like to extend special thanks to Mr. Huzal, Mr. Shafri and Ms. Aktikah for their support during both field and laboratory work. Moreover, I wish to extend my appreciation to my postgraduate colleagues for their constant encouragement, especially to my labmates for their continuous support and cooperation throughout the sampling process.

I am profoundly grateful to UNIMAS for granting me the Zamalah Graduate Scholarship, which has enabled me to pursue my studies without financial constraints. Also, I am indebted to Sarawak Forestry Corporation for the permission to conduct this study throughout Sarawak (Permit No: WL 21/2021; WL 09/2022). I am also thankful for Sarawak Research and Development Council (SRDC) as this study was supported by their grant (SRDC-RDCRG02/CAT/2020/_33).

Last but not least, I am forever grateful to my family for their unwavering support throughout my postgraduate journey. Words cannot express how thankful I am for the sacrifices they have made for me to pursue my studies far from home. I am truly fortunate and blessed to have a loving family that supports my aspirations, and I would not have achieved this milestone in my life without them.

ABSTRACT

The gut microbiota of rodents forms a dynamic ecosystem within the gastrointestinal tract, actively contributing to the host's physiological processes. The gut microbiome composition of wild rodents in particular could be influenced by external environmental factors. However, the specific determinants within the environments that directly affect the gut environment of wild rodents remain uncertain. One such influential factor could be natural water bodies, where bacterial communities can exploit the readily available organic matter for growth and circulate with the surrounding rodent populations. Additionally, rodents are known to be natural reservoirs of pathogenic bacteria due to their synanthropic nature and can transmit them through various pathways. Hence, this study aims to characterise the gut microbiome of wild rodents near natural water bodies and to identify pathogenic bacteria within the gut environment. To achieve these objectives, sampling was conducted across forested areas throughout Sarawak. Faecal samples of wild rodents were collected from a total of five sampling sites located near human settlements, accompanied by water quality sampling of natural water bodies across these localities. Nanopore sequencing was performed on the extracted faecal DNA for the characterisation of the gut microbiome of 16 wild rodent individuals belonging to six different species. A total of 1052 bacterial genera were detected, contributing to the gut microbiota of rodents at varying compositions and abundances. The most prominent bacterial families include Lachnospiraceae, Lactobacillaceae and Oscillospiraceae, belonging to the phylum Bacillota. They could be considered members of the core rodent gut microbiome due to their high prevalence across various rodent individuals. Also, potentially pathogenic bacteria were detected, and they represent 11% of the total overall abundance. 51 potentially pathogenic bacterial genera, comprising of 774 species, that can cause zoonotic diseases were detected from all sampling

sites, notably from the genera *Clostridium*, *Klebsiella* and *Helicobacter*. Additionally, while water quality parameters did not exhibit a significant association with the alpha diversity of the gut microbiome or the abundance of pathogenic bacteria, the genus *Enterococcus* displayed a positive correlation with biochemical oxygen demand. Overall, the characterisation of the gut microbiome of wild rodents across different localities revealed distinct variations. The site-specific bacterial species detected in this study highlights the potential environmental influences on the gut microbiota. This includes the detection of pathogenic bacteria that could potentially transmit to nearby human settlements, implying considerable health concerns.

Keywords: 16S gene, gut microbiome, next-generation sequencing, Rodentia, Borneo

***Kepelbagaian Bakteria Berpotensi Patogenik di dalam Populasi Tikus Berdekatan
dengan Sumber Air Semulajadi di Sarawak***

ABSTRAK

Mikrobiota usus tikus menghasilkan sebuah ekosistem yang dinamik di dalam saluran gastrousus, secara tidak langsung menyumbang kepada proses fisiologi perumahnya. Komposisi mikrobiom usus tikus liar boleh dipengaruhi oleh beberapa faktor luaran persekitaran. Namun, faktor khusus daripada persekitaran yang mempengaruhi keadaan usus tikus-tikus liar masih belum dikenalpasti. Antara kemungkinan berpunca daripada sumber air semulajadi, di mana komuniti bakteria menggunakan bahan-bahan organik yang sedia ada untuk membiak dan seterusnya tersebar dalam kalangan populasi tikus liar di sekeliling sumber air. Selain itu, tikus juga terkenal sebagai pembawa bagi bakteria patogenik kerana kejadian sinantropik yang menjuruskan kepada penularan bakteria tersebut melalui pelbagai cara. Oleh itu, kajian ini bertujuan untuk mengklasifikasikan mikrobiom di dalam usus tikus liar yang dijumpai berdekatan dengan sumber air semulajadi dan seterusnya mengenalpasti bakteria patogenik di dalam usus tikus-tikus tersebut. Selaras dengan objektif kajian, pengambilan sampel dijalankan di beberapa kawasan hutan di Sarawak. Sampel najis tikus liar diambil daripada lima lokasi yang berhampiran dengan penempatan manusia, berserta dengan sampel kualiti air daripada sumber air semulajadi di lokasi-lokasi tersebut. Penjujukan nanopore dilakukan ke atas DNA najis tikus yang telah diekstrak untuk mengklasifikasikan jenis mikrobiom usus bagi 16 individu tikus liar daripada enam spesis berbeza. Sejumlah 1052 genus bakteria dikesani yang membentuk mikrobiota usus tikus dengan pelbagai komposisi dan kekayaan yang berbeza. Keluarga bakteria paling utama terdiri daripada keluarga Lachnospiraceae, Lactobacillaceae dan Oscillospiraceae yang tergolong di dalam filum Bacillota. Keluarga tersebut dianggap sebagai teras

mikrobiom usus tikus disebabkan kekerapan kejadian di dalam individu tikus. Malah, bakteria yang berpotensi patogenik telah dikesan, dan ia mewakili 11% daripada jumlah keseluruhan. 51 genus bakteria yang berpotensi patogenik, terdiri daripada 774 spesies, yang boleh menyebabkan penyakit zoonotik dikenalpasti daripada kesemua lokasi untuk pengambilan sampel, terutamanya dari genus Clostridium, Klebsiella dan Helicobacter. Meskipun parameter kualiti air tidak menunjukkan hubungan yang signifikan dengan kepelbagaian alfa mikrobiom usus atau kepelbagaian bakteria patogenik, salah satu genus bakteria iaitu Enterococcus menunjukkan terdapat perkaitan dengan permintaan oksigen biokimia. Secara keseluruhan, pengklasifikasi mikrobiom usus tikus liar dari pelbagai lokasi menunjukkan terdapatnya perbezaan. Spesies bakteria yang dikesan di dalam satu lokasi daripada kajian ini menunjukkan terdapat kesan persekitaran terhadap mikrobiota usus tikus. Ini termasuk penemuan bakteria patogenik yang berpotensi merebak ke penempatan manusia yang berdekatan, membawa kepada kebimbangan dalam isu kesihatan.

Kata kunci: *gen 16S, mikrobiom usus, penjujukan generasi hadapan, Rodent, Borneo*

TABLE OF CONTENTS

	Page
DECLARATION	i
ACKNOWLEDGEMENT	ii
ABSTRACT	iii
<i>ABSTRAK</i>	v
TABLE OF CONTENTS	vii
LIST OF TABLES	x
LIST OF FIGURES	xii
LIST OF ABBREVIATIONS	xiv
CHAPTER 1: GENERAL INTRODUCTION	1
1.1 Introduction	1
1.2 Problem Statement	3
1.3 Objectives	4
1.4 Chapter Summary	4
CHAPTER 2: LITERATURE REVIEW	5
2.1 Gut Microbiota of Rodents	5
2.2 Role of Water Bodies on Bacterial Transmission	8

2.3	Pathogenic Bacteria in Rodents	10
2.4	Next-Generation Sequencing	13
CHAPTER 3: GUT MICROBIOME COMPOSITION OF WILD RODENTS		17
IN FORESTED AREAS OF SARAWAK		
3.1	Introduction	17
3.2	Materials and Methods	19
3.2.1	Study Sites	19
3.2.2	Sampling Methods	20
3.2.3	DNA Extraction and Molecular Species Identification	20
3.2.4	Next-Generation Sequencing	21
3.2.5	Data Analysis	22
3.3	Results	23
3.4	Discussion	32
3.5	Conclusion	36
3.6	Recommendation	36
CHAPTER 4: PREVALENCE OF POTENTIALLY PATHOGENIC BACTERIA		37
IN WILD RODENTS NEAR NATURAL WATER BODIES FROM		
SARAWAK		
4.1	Introduction	37
4.2	Materials and Methods	39

4.2.1	Study Sites	39
4.2.2	Sampling Methods	41
4.2.3	Water Quality Sampling	42
4.2.4	DNA Extraction and Next-Generation Sequencing	42
4.2.5	Data Analysis	43
4.3	Results	44
4.4	Discussion	52
4.5	Conclusion	57
4.6	Recommendation	58
CHAPTER 5: GENERAL DISCUSSION		59
5.1	Discussion	59
5.1.1	Gut Microbiome Composition	59
5.1.2	Prevalence of Pathogenic Bacteria	60
CHAPTER 6: CONCLUSION AND RECOMMENDATION		62
6.1	Conclusion	62
6.2	Recommendation	63
REFERENCES		64
APPENDICES		82

LIST OF TABLES

	Page
Table 2.1 Summary of Different Physiological and Environmental Factors That can Affect the Gut Microbiota of Rodents With Their Respective Citations	8
Table 2.2 Summary of Previous Studies in Sarawak That Highlight the Presence of Pathogenic Bacteria in Wild Rodent Populations With Their Respective Citations	12
Table 2.3 Summary of Previous Studies That Utilised Next-Generation Sequencing for the Analyses of Gut Microbiome in Rodents With Their Respective Citations	16
Table 3.1 Summary of Rodent Species Trapped in Sarawak	23
Table 3.2 Average Percentage of K2P Genetic Distance Values Among Rodent Species Based on Cytochrome <i>b</i> Gene Sequences	25
Table 3.3 Relative Abundance Table in Percentages of the Top 23 Most Abundant Bacterial Genera (>1% Overall Relative Abundance) According to Rodent Species	28
Table 4.1 Environmental Information and Description of Sampling Sites in Sarawak	40
Table 4.2 Abundance Table of the Top 23 Most Abundant Bacterial Genera (>1% Overall Relative Abundance) According to Different Localities. Potentially Pathogenic Bacterial Genera Marked With ‘*’	44

Table 4.3	Water Quality Parameters of Natural Water Bodies of Sampling Sites in Sarawak. Correlation Coefficients Between Shannon Diversity Indices and Water Quality Parameters are Shown in the Last Row	50
-----------	--	----

LIST OF FIGURES

	Page
Figure 3.1 Map of Sampling Sites Marked With Red Dots in Sarawak, Borneo	19
Figure 3.2 Photo of Captured <i>Niviventer cremoriventer</i> Individual in a Cage Trap	24
Figure 3.3 Photo of <i>Rattus tanezumi</i> Individual in the Wild	24
Figure 3.4 Maximum-Likelihood Phylogenetic Tree of Sequenced Rodent Individuals Highlighted in Blue. Red Diamonds Indicate Nodes With Bootstrap Values >90 (PM = Peninsular Malaysia, BO = Borneo, nBO = North Borneo, nwBO = Northwest Borneo, neBO = Northeast Borneo, wBO = West Borneo, eBO = East Borneo)	26
Figure 3.5 Sankey Chart of Combined Gut Microbiome Composition of All Rodent Individuals (n = 16) According to Different Taxa Levels. Only Bacterial Genera With Relative Abundance of >1% Were Shown. Number Above Nodes Indicate the Number of Reads Assigned for Each Taxon. The Most Prominent Taxon for Each Taxa Level was Circled in Red	29
Figure 3.6 Abundance Bar Plots of the Top 23 Most Abundant Bacterial Genera From Rodent Individuals Across Different Localities. The Top 23 Genera Were Selected as They Represent >1% of the Relative Abundance of the Total Bacterial Abundance. Potentially Pathogenic Bacterial Genera Marked With '*'. Juvenile <i>N. cremoriventer</i> Individual Outlined in Black. Abbreviations of Samples Follow Appendix A	30

Figure 3.7	Box Plot of Shannon Diversity Index of the Gut Microbiome of Rodent Individuals Across Different Study Sites. Juvenile <i>N. cremoriventer</i> Individual is Labelled and Circled in Black	31
Figure 4.1	Map of Sampling Sites Indicated by Red Dots in Sarawak, Borneo	40
Figure 4.2	Abundance Bar Plots of the Top 12 Most Abundant Potentially Pathogenic Genera Across Different Localities. The Selected Genera Represented >1% of the Relative Abundance of the Total Number of Pathogenic Bacteria While the Remaining Genera Below the Threshold Were Consolidated and Labelled as ‘Others’	47
Figure 4.3	A Venn Diagram Showing the Number of Bacterial Species Being Detected Between Five Localities of Study Sites	48
Figure 4.4	A Non-Metric Multidimensional Scaling (NMDS) Ordination Plot Using the Jaccard Index of all Rodent Individuals Colour Coded to Their Locality	49
Figure 4.5	Correlogram Showing Correlations Between Water Quality Parameters and Abundance of Pathogenic Bacteria. Positive Correlations are Shown in Blue and Negative Correlations are Displayed in Red With the Intensity of the Colour Being Proportional to Correlation Strength. Values Marked With Black Outline are Significant at $p < 0.05$	51

LIST OF ABBREVIATIONS

°C	Degree Celsius
BLAST	Basic Local Alignment Search Tool
BOD	Biochemical Oxygen Demand
DNA	Deoxyribonucleic Acid
DO	Dissolved Oxygen
GGNP	Gunung Gading National Park
LHNP	Lambir Hills National Park
MAE	Marup Atas Engkilili
MEGA	Molecular Evolutionary Genetics Analysis
mg/L	Milligram per Litre
ml	Millilitre
mm	Millimetre
ms/cm	Millisiemen per Centimetre
mtDNA	Mitochondrial Deoxyribonucleic Acid
n	Number
NCBI	National Center for Biotechnological Information
NGS	Next-Generation Sequencing
NMDS	Non-Metric Multidimensional Scaling
ONT	Oxford Nanopore Technologies
PCR	Polymerase Chain Reaction
ppt	Parts per Thousand
RNA	Ribonucleic Acid
rRNA	Ribosomal Ribonucleic Acid

SRDC	Sarawak Research and Development Council
SSK	Sungai Sibau Kapit
TDS	Total Dissolved Solids
UNIMAS	Universiti Malaysia Sarawak
UPK	Ulu Poi Kanowit

CHAPTER 1

GENERAL INTRODUCTION

1.1 Introduction

The gut microbiota of rodents forms a complex ecosystem within the gastrointestinal tract, actively contributing to host health through processes such as digestion, nutrient absorption, colonisation resistance, and immune system regulation (Kinross et al., 2011; Valdes et al., 2018; Hou et al., 2022). Essential microbiota within the gut environment are regarded as core members, responsible for shaping the microbiome and significantly influencing host physiology (Risely, 2020; Perlman et al., 2022; Sharon et al., 2022). For wild rodents in particular, their gut microbiota changes with age and reproductive status, while environmental factors further impact microbial structures, influenced by habitat types, diet variation, and seasonal shifts (Maurice et al., 2015; Viney, 2019; Fenn et al., 2023). These factors can lead to variations in the gut microbiota, potentially affecting microbial diversity among different rodent individuals. However, the specific ecological influences that shape the microbiota remain largely unknown, warranting further research.

Bacteria can thrive in diverse environments, including the natural water bodies within forested areas, where they exploit available organic matter for growth (Ren et al., 2019). Water quality parameters such as pH, temperature, dissolved oxygen, and conductivity significantly influence microbial communities in these ecosystems (Staley et al., 2018; Wang et al., 2018; Li et al., 2021). Rodents dwelling near natural water sources can serve as hosts for environmental bacteria, which adapt to persist in water due to their resilience to biotic and abiotic pressures (Barcina et al., 1997). This is particularly true for water surfaces of rivers and streams, where bacteria thrive even in low-nutrient conditions

(Vital et al., 2010). As a result, the presence and adaptations of bacteria within natural water bodies could lead to transmission to nearby rodent populations, especially those that could be pathogenic and can cause serious illnesses.

Rodents play an important role in the transmission of zoonotic diseases as they are known to act as reservoirs hosts for pathogenic bacteria. These bacteria can persist for long periods of time in their hosts by exploiting the gut environment's resources to promote growth and virulence, resulting in infections that can be transmitted to humans through direct or indirect pathways (Meerburg et al., 2009; Bäumler & Sperandio, 2016). With their adaptability to diverse habitats and proximity to human settlements, wild rodents could pose health concerns, especially in human-altered environments that increase opportunities for disease transmission (Bordes et al., 2015; Dahmana et al., 2020). The utilisation of modern sequencing techniques has facilitated the detection of pathogenic bacteria in wild rodent populations, highlighting the significance of disease transmission prevention.

Recent advancements in next-generation sequencing (NGS) technologies, specifically nanopore sequencing, enable the profiling of bacterial communities in wildlife. This technology enables rapid and precise identification of microbial diversity at the species level. Unlike conventional methods, nanopore sequencing offers extended read lengths, real-time data, and device portability (Chen et al., 2023). Studies have demonstrated its efficacy in characterising rodent gut microbiota and detecting pathogens in rodents, showcasing its potential for concise pathogenic surveys in wild rodent populations near human settlements (Shin et al., 2016; Jahan et al., 2021). Overall, nanopore sequencing emerges as a suitable technology for comprehensive microbial analysis and identifying pathogenic bacteria within rodent populations.

1.2 Problem Statement

In recent years, next-generation sequencing has been successfully employed to distinguish bacterial communities within the gut environment of wild rodents (Weldon et al., 2015; Rosshart et al., 2017; Jahan et al., 2021; Song et al., 2021). Nonetheless, a knowledge gap exists concerning the gut microbiome composition of wild rodents from Borneo due to the limited research using next-generation sequencing. Recent studies that have utilised this technology in Borneo (Hayakawa et al., 2018) and Peninsular Malaysia (Mohd-Yusof et al., 2022; Arumugam et al., 2023) mostly focus on other mammals. The characterisation of the rodents' gut microbiome has the potential to enhance our understanding of the dynamic interactions between gut microbiota and rodent hosts. Research on wild rodents from forests is particularly valuable, given the potential influence of environmental factors on gut microbiome composition. The diversity of the gut microbiome could be influenced by a combination of environmental factors and host genetics (Campbell et al., 2012). Environmental bacteria such as the bacterial communities found in natural water bodies could play a pivotal role in shaping the microbiome of wild rodents (Vital et al., 2010). Therefore, assessing the potential associations between the water quality of natural water sources and the diversity of gut bacteria in rodents will give a better understanding on the environmental pressures that could directly impact the gut microbiome.

Numerous studies have highlighted the presence of pathogenic bacteria in wild rodent populations (Thayaparan et al., 2013; Bilung et al., 2014; Lesley et al., 2016; Pui et al., 2017; Blasdell et al., 2019a; Blasdell et al., 2019b). This includes locations with minimal anthropogenic impact, such as primary forests, where the discovery of native novel pathogenic bacteria is possible (Lau et al., 2020). These findings give rise to significant health concerns due to the synanthropic nature of rodents and the potential for pathogenic

bacteria transmission to adjacent human settlements. Additionally, natural water bodies could play a significant role in the circulation of pathogens among wild rodent populations. A study by Lesley et al. (2016) has suggested the dispersal of antibiotic resistant pathogenic bacteria from a river within a protected area to the surrounding wildlife. By utilising nanopore sequencing, there exists the potential to detect pathogenic bacteria capable of causing zoonotic diseases through characterisation of the wild rodent gut microbiome.

1.3 Objectives

This research aimed to conduct a comprehensive analysis of the gut microbiota of wild rodents inhabiting forests while investigating the associations of water quality and its influence on gut microbial diversity. This was achieved through the following set of objectives:

- i. To determine the gut microbiome composition of wild forest rodents in Sarawak
- ii. To identify the prevalence of potentially pathogenic bacteria within the gut microbiota of the rodent individuals near natural water bodies in Sarawak

1.4 Chapter Summary

The study is composed of five chapters, two of which are research articles written for publication in a scientific journal. The first chapter is the introduction to the general outline of the thesis, followed by the second chapter of a literature review based on related studies. Next, chapters three and four are the aforementioned research articles. Each article includes an introduction, materials and methods, results, discussion and conclusion. Finally, the general discussion and conclusions are written in the last chapter.

CHAPTER 2

LITERATURE REVIEW

2.1 Gut Microbiota of Rodents

The gut microbiota refers to a complex ecosystem of microorganisms residing in the gastrointestinal tract, actively contributing to the maintenance of the host organism's health (Coyte & Rakoff-Nahoum, 2019). These microbes facilitate essential processes, including food digestion, absorption of nutrients, development of colonisation resistance and regulation of the immune system (Kinross et al., 2011; Valdes et al., 2018; Hou et al., 2022). Notably, the composition and diversity of the gut microbiome among rodents could exhibit variation among different individuals as each distinct bacterial community could be shaped by external environmental factors and host genetics (Campbell et al., 2012; Viney, 2019; Wang et al., 2022). Within rodents, the gut microbiota predominantly comprises various phyla, with Firmicutes, Bacteroides, and Proteobacteria prevailing as the most prominent members (Clavel et al., 2016). Certain members of these bacteria within the gut environment could potentially be key components of the gut microbiota in wild rodents.

Members of the gut microbiota that play a pivotal role in the overall health and ecological dynamics of rodents can be considered constituents of the core gut microbiome. These vital microbes influence the essential physiological processes within their host and define their gut microbiome composition (Risely, 2020; Perlman et al., 2022; Sharon et al., 2022). For example, the phyla Bacteroidota and Firmicutes are linked to animals with plant-based diets due to the presence of genes responsible for degradation of polysaccharides (El Kaoutari et al., 2013; Zou et al., 2019). However, identifying specific bacterial taxa as an integral part of the gut microbiome based on their substantial contributions to its ecological

structure presents various challenges (Perlman et al., 2022). Instead, it is favoured to define them by their capability to persist within the gut environment while exhibiting the highest abundance across hosts. Among mammals, the core gut microbiome members were found to be well adapted to the host's environment throughout their lifespan, being resistant to changes across time (Furman et al., 2020). For domesticated mammals such as livestock, the core microbiome was based on a high-occupancy criteria and mostly consists of the Bacteroidota, Firmicutes and Proteobacteria phyla (Perlman et al., 2022). Therefore, the characterisation of the core gut microbiome is possible by determining the concurrent bacterial taxa across rodent individuals.

The gut microbiota structure of wild rodents may undergo alterations based on the maturation of host physiology as an increase of species richness of bacteria can be influenced by the biological age of the host (Kim & Jazwinski, 2018). The ageing process of rodents has been associated with long-term shifts in gut microbiome composition and diversity (Fenn et al., 2023). As rodents age, the bacterial community within the gut environment gradually develops, progressing towards a state of maturity. Moreover, Maurice et al. (2015) have revealed that the reproductive status of wild rodents has been recognised as a factor contributing to changes in the gut microbiota. They have found that reproductively active males (protruding or descended testes) and females (pregnant or perforated vagina) were associated with microbial community structure. The abundance of *Lactobacillus* in particular was higher in reproductively active female rodents than non-active females. However, the specific mechanisms underlying this association are not fully understood and require further investigation.

Certain environmental factors unique to the natural habitat of wild rodents can also contribute to shaping their gut microbiome composition, resulting in distinct microbial

structures. Teng et al. (2022) have discovered that the gut microbiota is significantly more influenced by the specific habitat types of rodents than by variations between species. These variations can be caused by exposure to a diverse range of environmental microbes and the consumption of varied diets based on the conditions of their current habitats (Viney, 2019). Moreover, a study conducted by Maurice et al. (2015) identified short-term shifts in the gut microbial communities of wild rodents. They found that seasonal change in diet corresponded with fluctuations in the abundance of specific bacterial taxa, leading to shifts in the gut microbiome structure. However, the precise ecological influences that can alter the gut microbiota of rodents remain relatively unknown.

As rodents are able to adapt to various types of habitats (Hart, 1971), the environmental pressures experienced by rodent populations could differ depending on their native environments. In peridomestic regions, rodents coexist in close proximity or adjacent to human settlements. This close association provides an opportunity for these rodents to maintain a gut microbiome reflective of their native habitat, while simultaneously engaging in interactions with the surrounding human populations. This includes rodents inhabiting areas such as protected national parks located near human settlements and frequently visited forests that are adjacent to villages. Therefore, wild rodent populations in these localities are suitable for study, not only because of the relatively ease of access to these areas but also due their gut microbiome composition that remains largely unaffected by urbanisation.

Table 2.1: Summary of Different Physiological and Environmental Factors That can Affect the Gut Microbiota of Rodents With Their Respective Citations

Factors	Description	References
Ageing	Gut microbiome diversity and composition shifts toward maturity	Kim & Jazwinski, 2018; Fenn et al., 2023
Reproductive status	<i>Lactobacillus</i> more abundant in sexually active females than in non-active females	Maurice et al., 2015
Habitat types	Variations caused by exposure to different environmental microbes and diet	Viney, 2019; Teng et al., 2022
Seasonal changes	Temporal changes cause fluctuations of food availability	Maurice et al., 2015

2.2 Role of Water Bodies on Bacterial Transmission

Bacteria can persist in various environments, including the natural water bodies of forested areas. Water sources have high potential for bacterial development as they could take advantage of the readily available organic matter for their growth. Several physicochemical properties of water have been thoroughly investigated and have been identified as predominant factors that could shape microbial communities (Li et al., 2021). Some water quality parameters of significance include pH, temperature, dissolved oxygen and conductivity (Staley et al., 2018; Wang et al., 2018). The interactions of these factors within the unique ecosystems of water bodies in forested areas can give rise to distinct bacterial communities.

Assessing the physicochemical parameters of a water body is one of the only ways to determine its level of water quality (Yogendra & Puttaiah, 2008). One of the most widely used parameter is biochemical oxygen demand (BOD) which provides information on the organic load of water. It is defined by the oxygen consumed during the respiratory activity of microbes and serves as a measure of organic pollution in water (Jouanneau et al., 2014). Therefore, a high BOD value is an indicator for polluted water bodies as it would correspond to high microbial load in water. Furthermore, pH is another significant factor in shaping the composition of bacterial communities in water bodies (Lindström et al., 2005). Bacteria exhibit the ability to endure low pH conditions through adaptive measures such as modifying cell membrane composition, altering metabolic pathways and generating alkali (Cotter & Hill, 2003). Consequently, a water body with low pH levels could harbour distinct bacterial communities.

Rodents residing in habitats with natural occurring sources of water have the potential to serve as hosts for the bacteria present in these aquatic environments. Bacteria can adapt to persist within bodies of water until they are able to infect hosts as they are resistant to natural biotic and abiotic pressure present in these environments (Barcina et al., 1997). Hence, natural water reservoirs that serve as drinking sources for wildlife could play an important role in bacterial transmission to rodents. This is especially true for water surface environments as bacteria are able to survive and grow even in these low-nutrient conditions (Vital et al., 2010). This includes shallow waters such as rivers or streams, where rodents could be partially submerged in water. Consequently, this increases the likelihood of circulating pathogenic bacteria among wild rodent populations and the adjacent natural water bodies they inhabit.

For pathogenic bacteria, antibiotic resistant *Listeria monocytogenes* were present in water samples of primary forests from Sarawak (Lesley et al., 2016). It was found that the pathogenic bacteria circulated from the natural water bodies to the surrounding wildlife and the bacteria's ability to adapt to living hosts were believed to have affected their dispersal to wildlife. This is consistent with the findings of Raschle et al. (2021) as they have found that transmission of *L. monocytogenes* to animals are due to the occurrence of the bacteria in water surfaces of riverine ecosystems. There are also reports of some pathogenic bacteria that are adapted to not occur near water sources. A study done by Blasdell et al. (2019a) in Sarawak revealed that the bacteria *Leptospira borgpetersenii* relies on direct transmission and was more likely to persist in rodents at sites without natural water bodies. In contrast, they have also found that the bacteria *Leptospira interrogans* are able to persist in aquatic environment for extended periods. This shows that the presence of bodies of water could influence the occurrence of pathogenic bacteria in rodents due to the different adaptations of the bacteria to its existing environment.

2.3 Pathogenic Bacteria in Rodents

Rodents are known vectors of zoonotic diseases due to their roles as natural reservoirs for microorganisms such as pathogenic bacteria. Pathogenic bacteria are able to thrive in the gut environment by exploiting microbiota-derived nutrients to promote their growth and virulence (Bäumler & Sperandio, 2016). Moreover, rodents are recognised worldwide for transmitting diseases caused by parasites and virus (Mills, 2005; Reperant et al., 2009). Rodents facilitate the transmission of infectious diseases to human populations through both direct means, such as bites and excretions, and indirect pathways involving other animals, including arthropods and livestock (Meerburg et al., 2009). Along with the adaptability of the rodents to occupy diverse range of habitats and their synanthropic nature

(Dahmana et al., 2020), the presence of wild rodents in the vicinity of human settlements raises significant health concerns. According to Bordes et al. (2015), human-altered environments would especially increase the opportunities for disease transmission, primarily due to the heightened and persistent contact between rodent species with humans.

A study done by Blasdell et al. (2019b) in Sarawak revealed multiple species of genus *Bartonella* were detected in wild rodents that were in close proximity with humans. It was found that *Bartonella* species were more prevalent in urban and developing areas when compared with rural areas. This is because the infected individuals were restricted to fewer forest patches in these areas which causes high-density populations of rodents and an increased rate of intraspecies contact. *Bartonella* infections are the principal cause of bartonellosis and have been associated with rodents in recent years (Krügel et al., 2022). The high prevalence of *Bartonella* species due to urbanisation shows that human development would increase the risk of exposure to zoonotic disease transmissions due to the potential rise of human-rodent interactions.

A similar pattern was reported for the genus *Leptospira* as wild mammals were found to be infected with pathogenic bacteria, *Leptospira* species in multiple protected areas in Sarawak (Thayaparan et al., 2013). In Southeast Asia, rodents are important reservoirs for *Leptospira*, and it was believed that they were the main cause of spreading the pathogenic bacteria in these protected areas as the other sampled mammals were not known for being carriers of leptospirosis (Cosson et al., 2014). Furthermore, Blasdell et al. (2019a) have reported on the rise of *Leptospira* species infected rodents as urbanisation increases. They have suggested that the environmental characteristics of urban areas could promote the circulation of the pathogenic bacteria. These findings were consistent with a study by Pui et al. (2017) as they have reported presence of pathogenic *Leptospira* spp. in soil and water

samples of urban areas. The prevalence of pathogenic bacteria in rodent populations could be influenced by their surrounding bacterial community and could be the main cause of transmission to humans, especially in highly urbanised areas.

Furthermore, the study of pathogenic bacteria in rodents inhabiting minimally disturbed areas like primary forests presents an opportunity to highlight novel species. Research done by Lau et al. (2020) uncovered the first occurrences of two *Borrelia* species in Sarawak, introducing new geographic evidence of *Borrelia miyamotoi* within a *Sundamys muelleri* individual from Gunung Gading National Park and detecting *Borrelia yangtzensis* in four *Rattus* spp. from an oil palm plantation. These potentially pathogenic bacteria were believed to circulate between *Ixodes* ticks and wild rodents as the bacteria were detected in both organisms. As these bacteria were detected near human settlements, studying rodents situated in such localities gains significance due to their increased likelihood of disease transmission to humans and their potential role as carriers of novel pathogenic bacteria.

Table 2.2: Summary of Previous Studies in Sarawak That Highlight the Presence of Pathogenic Bacteria in Wild Rodent Populations With Their Respective Citations

Pathogenic Species	Rodent Species	Sample Type	Locality	References
<i>Leptospira</i> spp.	<i>Maxomys</i> spp., <i>Rattus</i> spp., <i>Sundamys muelleri</i> ,	Blood, kidney, liver	Sarikei, Sibu, Kuching	[1]; [2]; [3]; [4]
<i>Escherichia coli</i>	Not stated	Anal swabs	Sibu, Kapit	[5]
<i>Listeria monocytogenes</i>	<i>Maxomys</i> sp.	Anal swabs	Kuching	[6]
<i>Bartonella</i> spp.	<i>Sundamys muelleri</i> , <i>Rattus</i> spp.	Spleen	Kuching	[7]

<i>Borrelia</i> spp.	<i>Sundamys muelleri</i> ,	Spleen,	Lundu,	[8]
	<i>Rattus</i> spp.	ticks	Miri	

[1] Thayaparan et al., 2013; [2] Pui et al., 2017; [3] Suut et al., 2018; [4] Blasdel et al., 2019; [5] Bilung et al., 2014; [6] Lesley et al., 2016; [7] Blasdel et al., 2019b; [8] Lau et al., 2020

2.4 Next-Generation Sequencing.

While the detection of pathogenic bacteria in wildlife has long relied on conventional sequencing methods, recent advancements in next-generation sequencing (NGS) technologies allow for the profiling of bacterial communities. One of the widely used NGS platform is Illumina sequencing by Solexa. Illumina sequencing employs bridge amplification, where DNA fragments (approximately 500 bp) bind to flow cell surfaces and undergo replication, resulting in the formation of clusters comprised of identical fragments (Slatko et al., 2018). The main advantage of this high throughput sequencing technique includes its high accuracy of reads of up to 96 samples per run (Morey et al., 2013). However, the nature of the bridge amplification technique used in Illumina sequencing limits the effective read length of sequences (Chen et al., 2013; Ari & Arikan, 2016).

In a study conducted by Anders et al. (2021), the gut microbial communities of wild rodents were investigated using Illumina sequencing to determine variations in the microbiome across various gut regions. They discovered that the alpha diversity of bacterial communities was not significantly different when comparing different segments of the lower gastrointestinal tract (cecum, colon and rectum). However, there were distinct differences between the microbial communities in the small intestines and the lower gastrointestinal tract, similar to previous studies (Gu et al., 2013; Suzuki & Nachman, 2016). Additionally,

the study highlighted that distinct rodent species possessed a unique microbiome within specific gut sections, particularly noticeable when comparing rodents that are distantly related.

Another example of NGS technology is nanopore sequencing by Oxford Nanopore Technologies (ONT). The core principle of nanopore sequencing involves passing single-stranded DNA or RNA fragments through bioengineered nanopores that generate electric currents, translating into specific nucleotide sequences (Ciuffreda et al., 2021). Nanopore sequencing is uniquely distinguished from other NGS methods due to its capability to yield extended read lengths, provide real-time sequencing data, and offer the added portability of the device itself (Chen et al., 2023). Currently, the basecaller integrated into nanopore sequencing platforms is ONT's Guppy, which has higher accuracy and performance compared to its previous iterations (Wick et al., 2019). The utilisation of nanopore sequencing presents the opportunity to explore complex microbial communities in wild rodents, identifying essential core gut microbiota in the process.

One of the available barcoding kits for nanopore sequencing is the 16S Barcoding Kit by ONT. This kit amplifies the entire ~1500 bp 16S rRNA gene from extracted genomic DNA by utilising the universal primer sets 27F (5'-AGA GTT TGA TCC TGG CTC AG-3') and 1492R (5'-TAC GGY TAC CTT GTT ACG ACT T-3') by Heuer et al. (1997). The 16S gene is widely used as a genetic marker in bacterial identification due to its widespread occurrence in nearly all bacterial species, its constant functionality over time and its substantial base pair size (Clarridge, 2004; Janda & Abbott, 2007). Notably, the current barcoding kits utilising the Kit 9 chemistry are currently being phased out and are being replaced with the Kit 14 chemistry barcoding kits that take advantage of the improved R10.4.1 flow cell.

Shin et al. (2016) conducted a comprehensive investigation into the viability of nanopore sequencing for discerning the gut microbiota composition of rodents. This involved a comparison between the outcomes of full-length reads from nanopore sequencing and the conventional short-read sequencing techniques (Illumina) that have widely been employed in the field. Their study demonstrated that utilising a nanopore sequencer for full-length sequencing enables rapid, precise, and streamlined identification of microbial diversity at the species level. However, others have highlighted several limitations of nanopore sequencing, including high error rates and the requirement for relatively high amounts of nucleic acid material (Wang et al., 2021). Despite these drawbacks, the precise detection of microbial diversity within the gut environment would allow for the complete characterisation of the rodent gut microbiome.

Nanopore sequencing has been successfully employed in detecting pathogenic bacteria in rodent populations near human settlements. A survey was done by Jahan et al. (2021) on the zoonotic bacterial pathogens in peridomestic rodents. Nanopore sequencing was performed on the faecal samples of the wild rodents as transmission would most likely occur through their excrements. They were able to characterise the gut microbiome of numerous rodents and detected multiple foodborne pathogens that could cause serious illnesses in humans. However, some pathogenic bacteria commonly associated with rodents, such as *Bartonella* and *Leptospira*, were not found in their faecal samples as these pathogens might not be transmitted through their faeces. Although the survey was carried out in a different country, the study on the microbiota of wild rodents in peridomestic regions offers invaluable insights into potential findings. Overall, the application of nanopore sequencing emerges as a suitable technique for conducting concise surveys on the pathogenic bacterial prevalence in wild rodent populations.

Table 2.3: Summary of Previous Studies That Utilised Next-Generation Sequencing for the Analyses of Gut Microbiome in Rodents With Their Respective Citations

NGS	Gene	Findings	References
Illumina	16S; V3–V4 region;	Microbial communities in the lower gastrointestinal tract were significantly different with the small intestines. Gut microbiome compositions of various rodent species were distinct, especially rodents that are distantly related.	Anders et al., 2021
Illumina and Oxford Nanopore	16S; V3–V4 region (Illumina) 16S; full length region (Oxford Nanopore)	Nanopore sequencing demonstrated rapid, precise and streamlined bacterial identification at the species level. However, limitations such as high error rates and high nucleic acid requirement is present due to the nature of full-length reads obtained from nanopore sequencing.	Shin et al., 2016; Wang et al., 2021
Oxford Nanopore	16S; full length region	Faecal samples of wild peridomestic rodents revealed the presence of various foodborne pathogens with the potential to cause severe illnesses in humans. The findings underscore the efficacy of nanopore sequencing in surveying pathogenic bacterial prevalence in wild rodent populations.	Jahan et al., 2021

CHAPTER 3

GUT MICROBIOME COMPOSITION OF WILD RODENTS IN FORESTED

AREAS OF SARAWAK

3.1 Introduction

The gut microbiota of rodents is a complex ecosystem of microorganisms inhabiting their gastrointestinal tracts. It comprises high bacterial diversity, collectively forming a dynamic and intricate community (Coyte & Rakoff-Nahoum, 2019). The gut microbial population plays a crucial role in maintaining the health of the host by facilitating essential processes such as digestion, nutrient absorption, colonisation resistance and immune system regulation (Kinross et al., 2011; Valdes et al., 2018; Hou et al., 2022). The gut microbiome could also affect the behaviour of wildlife through chemical signalling and effects on the host nervous system (Archie & Tung, 2015). Hence, the symbiotic relationship between these microorganisms and their host allows the formation of a stable bacterial community, comprising of core gut microbiota that contribute to the various physiological processes of wild rodents.

The core gut microbiota of rodents refers to the bacteria that would define their overall microbiome composition (Risely, 2020; Perlman et al., 2022; Sharon et al., 2022). These microbes could be essential in maintaining the overall gut microbial community while ensuring host survival. However, identifying keystone taxa that help shape the ecological structure of the gut microbiome is difficult, as highlighted by Perlman et al. (2022). Instead, a more suitable approach would be to determine reoccurring microbiome members in multiple individual hosts to gain an understanding of their significance. The concurrent bacterial taxa found across different individuals, even in an otherwise variable microbial

community, could presumably be considered members of the core gut microbiome (Turnbaugh et al., 2007). Nevertheless, the wild environment experienced by each rodent individual could differ, and this may affect their core gut microbiome composition.

The gut microbiome composition of wild rodents could vary according to different host individuals and are shaped by a combination of host genetics and environmental factors (Campbell et al., 2012; Viney, 2019, Teng et al., 2022). For instance, the gut microbial communities of wild rodents undergo short-term shifts in response to seasonal changes in their diet (Maurice et al., 2015). Although these shifts may not significantly alter the overall core gut microbiome in the long term, they enable the host rodents to adapt to fluctuations in food resource availability throughout the seasons. Furthermore, the ageing process of rodents were also found to significantly alter the composition and diversity of their gut microbiome in the long-term (Fenn et al., 2023).

Despite these insights, the specific environmental pressures directly influencing gut microbiome diversity in rodents are still not well understood. The gut microbiome of wild rodents from different localities could exhibit variation due to certain environmental pressure like the native vegetation structure (Fan et al., 2022) or the rodents' dietary behaviours (Maurice et al., 2015; Viney et al., 2019). The characterisation of the gut microbiome within wild rodents would contribute to a better understanding on the intricate relationships between the core microbiota and their rodent hosts. Currently, there is also a lack of information on the direct influences of environmental pressure on the gut microbiota of wild rodents. Thus, the main objectives of this study are to determine the gut microbiome composition of rodents inhabiting forested areas. The outcome of this study allows better understanding of the prevailing core microbiome members shared across multiple wild rodent individuals.

3.2 Materials and Methods

3.2.1 Study Sites

The study sites consist of multiple forested areas situated in both protected and non-protected regions of Sarawak, Borneo (Figure 3.1). The chosen locations host a variety of native forest rodent species, adapted to their respective habitats within the study sites. There are two protected areas which are Gunung Gading National Park (N1°41'27.0" E109°50'45.0") and Lambir Hills National Park (N4°11'57.5" E114°02'34.3"). The non-protected areas are Marup Atas Engkilili (N1°07'08.3" E111°38'15.8"), Ulu Poi Kanowit (N1°57'09.9" E112°13'23.0"), and Sungai Sibau Kapit (N2°00'02.7" E112°56'16.4"). Rodent trapping was conducted between October 2021 and March 2023.

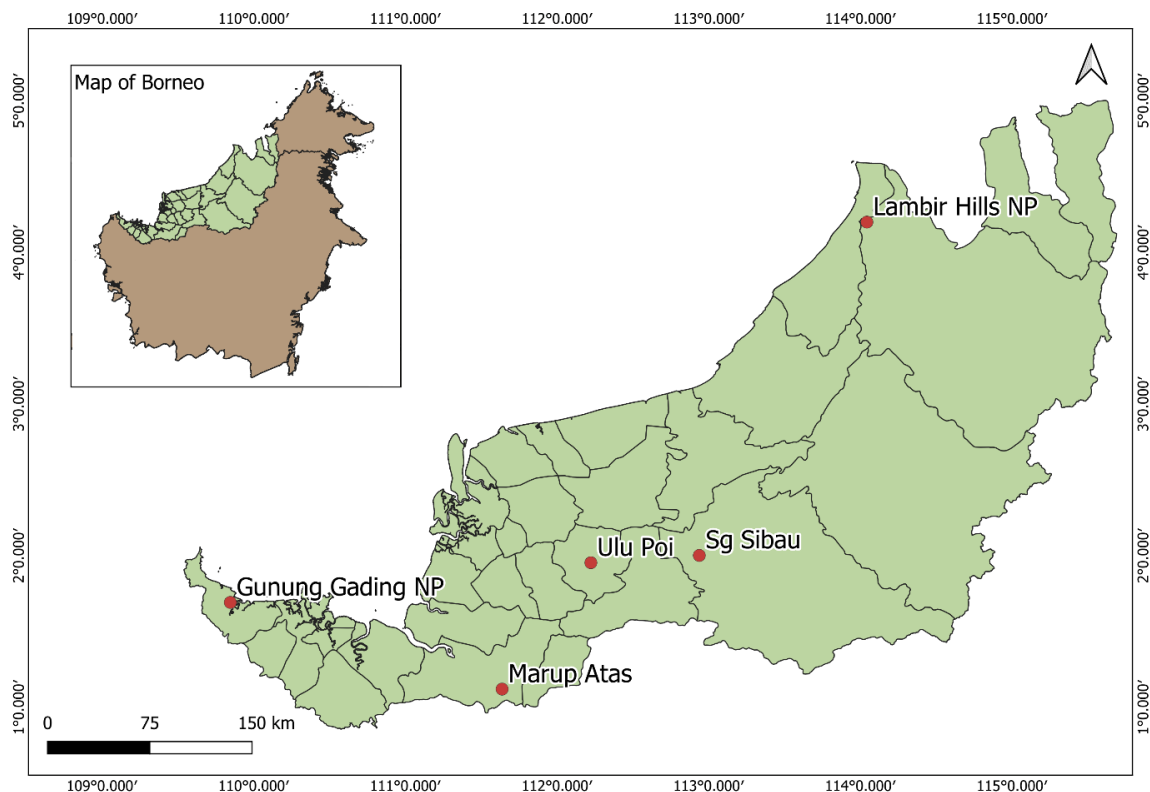


Figure 3.1: Map of Sampling Sites Marked With Red Dots in Sarawak, Borneo

3.2.2 Sampling Methods

Each sampling site consists of five trapping nights, with a total of 50 cage traps and 120 Sherman traps being used for the rodent trapping. The same traps were deployed at each site and sanitised beforehand using 70% ethanol and bleach. The traps were baited with banana covered with peanut butter. The traps were checked from 0900 hours to 1200 hours, as well as from 1400 hours to 1700 hours, with bait replacement occurring every two days. Morphological measurements of the captured rodents, including head body, tail ventral, ear, hindfoot, sex, weight and life stages, were also recorded. Then, the rodents were transferred to separate sterilised breathable containers to avoid cross contamination during faecal collection. The rodents were identified in the field using descriptions and keys from field guides (Payne et al., 1985; Phillipps & Phillipps, 2018). Molecular methods were later used to confirm the species identification due to the presence of several cryptic species. Fresh faecal samples (multiple 1 to 2 cubic mm pellets) of each captured rodent were transferred to sterilised 1.5 ml collection tube containing RNAlater (Ambion, USA) using sterilised forceps. RNAlater (Ambion) serves as a fixative by denaturing proteases and ribonucleases, thereby preserving the integrity of RNA in biological samples for subsequent analysis (Kruse et al., 2017). The collection tubes containing faecal samples were immediately stored in a cooler box. Upon arrival at the laboratory, the sample was promptly transferred to a -20°C freezer for DNA preservation and stability.

3.2.3 DNA Extraction and Molecular Species Identification

A QIAmp PowerFecal Pro DNA Kit (Qiagen, Germany) was used according to the manufacturer's protocol to extract the DNA from collected rodent faecal samples. The extracted DNA was amplified through polymerase chain reaction (PCR), targeting the

cytochrome *b* region of the mtDNA. The primer set used were LGL 765 (5'-GAA AAA CCA YCG TTG TWA TTC AAC T-3') and LGL 766 (5'-GTT TAA TTA GAA TYT YAG CTT TGG G-3') following Bickham et al., 1995. The PCR conditions were: pre-denaturation at 94°C for 7 minutes, 34 cycles of: denaturation at 92°C for 1 minute, annealing at 50°C for 1 minute and extension at 72°C for 1 minute, followed by 72 °C final extension for 7 minutes. The PCR products (~1500 bp) were sent to a commercial sequencing company for purification and sequencing (Apical Scientific, Malaysia).

Sequencher 4.1.4 (Gene Code, USA) was then used to clean the obtained DNA sequences and the NCBI Basic Local Alignment Search Tool (BLAST) was used for species identification based on the lowest E-values of the uploaded sequences. The lower E-values indicate high degree of similarity between sequences defined by many identical residues and few substitutions or gaps (Kerfeld & Scott, 2011). Finally, the DNA sequences, along with sequences from GenBank were analysed in MEGA-X 10.2.4 software to generate a Maximum-Likelihood phylogenetic tree and calculate the pairwise genetic distance between rodent individuals subjected to Kimura two-parameter model (Kimura, 1980). The sequences of the sampled rodents were submitted to the GenBank and Sequence Read Archive databases (Appendix B).

3.2.4 Next-Generation Sequencing

A DeNovix DS-11 Spectrophotometer (Denovix, Delaware, USA) was used for quantification of the extracted faecal DNA as preparation for nanopore sequencing. DNA with a concentration above 20 µg/ml and a purity ratio between 1.8 and 2.0 at 260 nm and 280 nm absorbance are considered to have passed the quality check. A total of 26 rodents were captured throughout the sampling period; however, only 16 faecal samples from the 26

rodents were used in this study due to the lower DNA concentration and purity of the remaining samples. The PCR amplification of 16S rRNA gene (27F and 1482R) were done using 16S Barcoding Kit (SQK-RAB204) from Oxford Nanopore Technologies, United Kingdom (ONT). The quality of PCR products were checked and only passed products (concentration above 20 µg/ml; absorbance ratio between 1.8 and 2.0 at 260/280 ratio) proceeded to sample pooling and library preparation following the manufacturer's protocol (ONT). The Flow Cell Priming Kit (EXP-FLP002) was then used to prime the flow cell (R9.4.1) and load the samples. Lastly, a MinION Mk1C sequencer (ONT) was used to perform nanopore sequencing for around 17 hours for each sequencing run.

3.2.5 Data Analysis

FASTQ files of the 16S rRNA sequences were obtained using Guppy v5.1.12 (ONT), which is the basecaller integrated in the MinKNOW v21.11.6 (ONT) operating system. The EPI2ME (ONT) application was used to perform taxonomic classification of bacterial species according to the National Center for Biotechnology Information (NCBI) taxonomic records. Reads below the threshold of the default quality score of 7 were excluded. The range of the library sizes was between ~25,000 and ~1,300,000 reads. The R package phyloseq v1.42.0 (McMurdie & Holmes, 2013) was used to normalise the dataset to the library with the lowest number of reads for further analysis using the rarefying method. The gut microbiome of rodents was visualised using the MiscMetabar v0.33 package (Taudière, 2023) to produce a Sankey chart. Also, the ggplot2 v3.4.2 package (Wickham, 2011) was used to produce an abundance graph of bacterial genera for each rodent individual. Alpha diversity was estimated using the Shannon index and plots of diversity indices were done using the phyloseq v1.42.0 package (McMurdie & Holmes, 2013).

3.3 Results

Sixteen sampled rodent individuals belonging to six different species were used in this study (Table 3.1). Throughout the study, *Niviventer cremoriventer* (Figure 3.2) was the most sampled rodent species, with a total of five individuals captured, in contrast to *Rattus tanezumi*, (Figure 3.3) which was the least sampled with only one individual recorded. Among the sites, Sungai Sibau Kapit had the highest number of rodent captures, with a total of five individuals. The lowest number of sampled rodents were at Gunung Gading National Park and Lambir Hills National Park which only two individuals each. Furthermore, Sungai Sibau Kapit had the highest diversity of rodents as four species were caught during sampling. In contrast, Gunung Gading National Park recorded only one species. Supplementary data of rodent individuals is available in Appendix A.

Table 3.1: Summary of Rodent Species Trapped in Sarawak

Rodent Species	GGNP*	LHNP*	MAE*	SSK*	UPK*
<i>Maxomys surifer</i>	0	0	0	1	0
<i>Maxomys tajuddini</i>	0	0	2	1	1
<i>Maxomys whiteheadi</i>	2	0	1	0	0
<i>Niviventer cremoriventer</i>	0	1	1	2	1
<i>Rattus tanezumi</i>	0	0	0	0	1
<i>Sundamys muelleri</i>	0	1	0	1	0
Total	2	2	4	5	3

*Gunung Gading National Park (GGNP), *Lambir Hills National Park (LHNP), *Marup Atas Engkilili (MAE), *Sungai Sibau Kapit (SSK), *Sungai Sibau Kapit (SSK) and *Ulu Poi Kanowit (UPK)

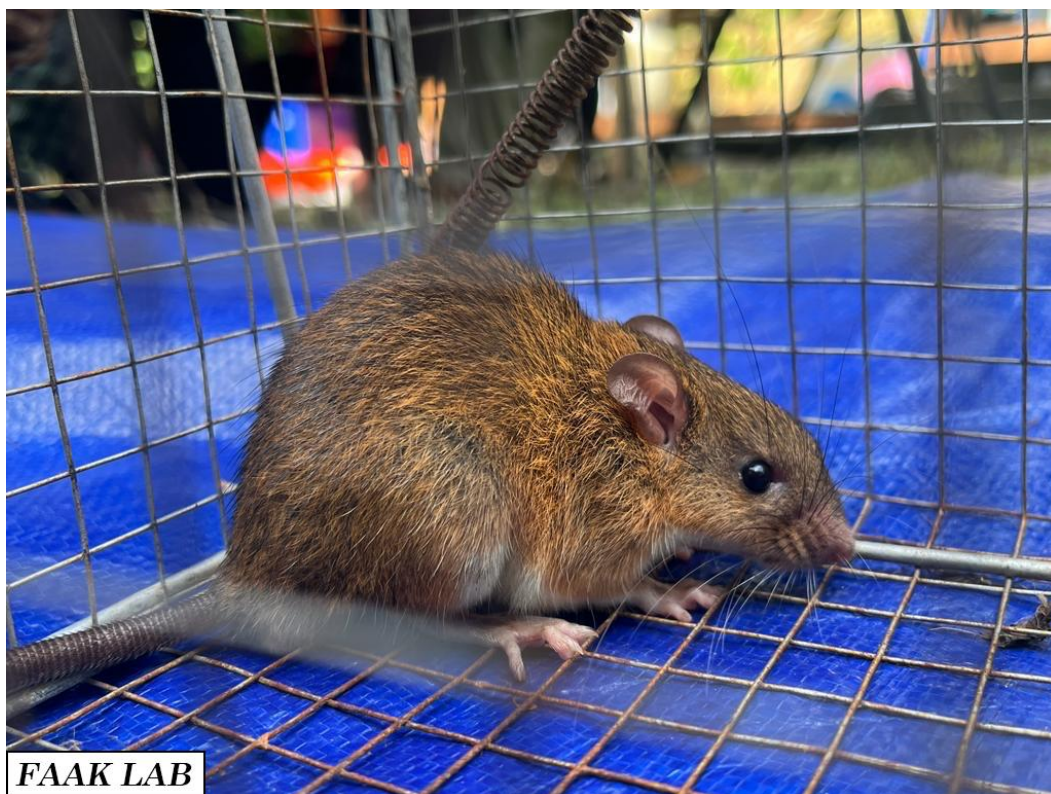


Figure 3.2: Photo of Captured *Niviventer cremoriventer* Individual in a Cage Trap



Figure 3.3: Photo of *Rattus tanezumi* Individual in the Wild

Sequences of sampled rodent individuals (cytochrome *b*) were utilised for genetic distance and phylogenetic analyses (Appendix C). The distance matrix of genetic distances based on the K2P model (Kimura, 1980) shows both intraspecific and interspecific divergence of sampled rodents (Table 3.2). For interspecific variation, the lowest mean genetic distance was between *Maxomys tajuddini* and *M. whiteheadi* (9.7%), while the highest was between *M. tajuddini* and *S. muelleri* (18.0%). Based on the produced Maximum-Likelihood phylogenetic tree (Figure 3.4), the sequences of the sampled rodents grouped with the GenBank sequences with high bootstrap values and confirms the species identification of the rodents.

Table 3.2: Average Percentage of K2P Genetic Distance Values Among Rodent Species Based on Cytochrome *b* Gene Sequences

No.	Rodent Species	1	2	3	4	5	6
1	<i>Maxomys surifer</i>	-					
2	<i>Maxomys tajuddini</i>	14.5	-				
3	<i>Maxomys whiteheadi</i>	12.6	9.7	-			
4	<i>Niviventer cremoriventer</i>	16.6	16.3	15.9	-		
5	<i>Rattus tanezumi</i>	16.5	17.3	16.5	16.2	-	
6	<i>Sundamys muelleri</i>	17.8	18.0	17.0	17.2	15.9	-

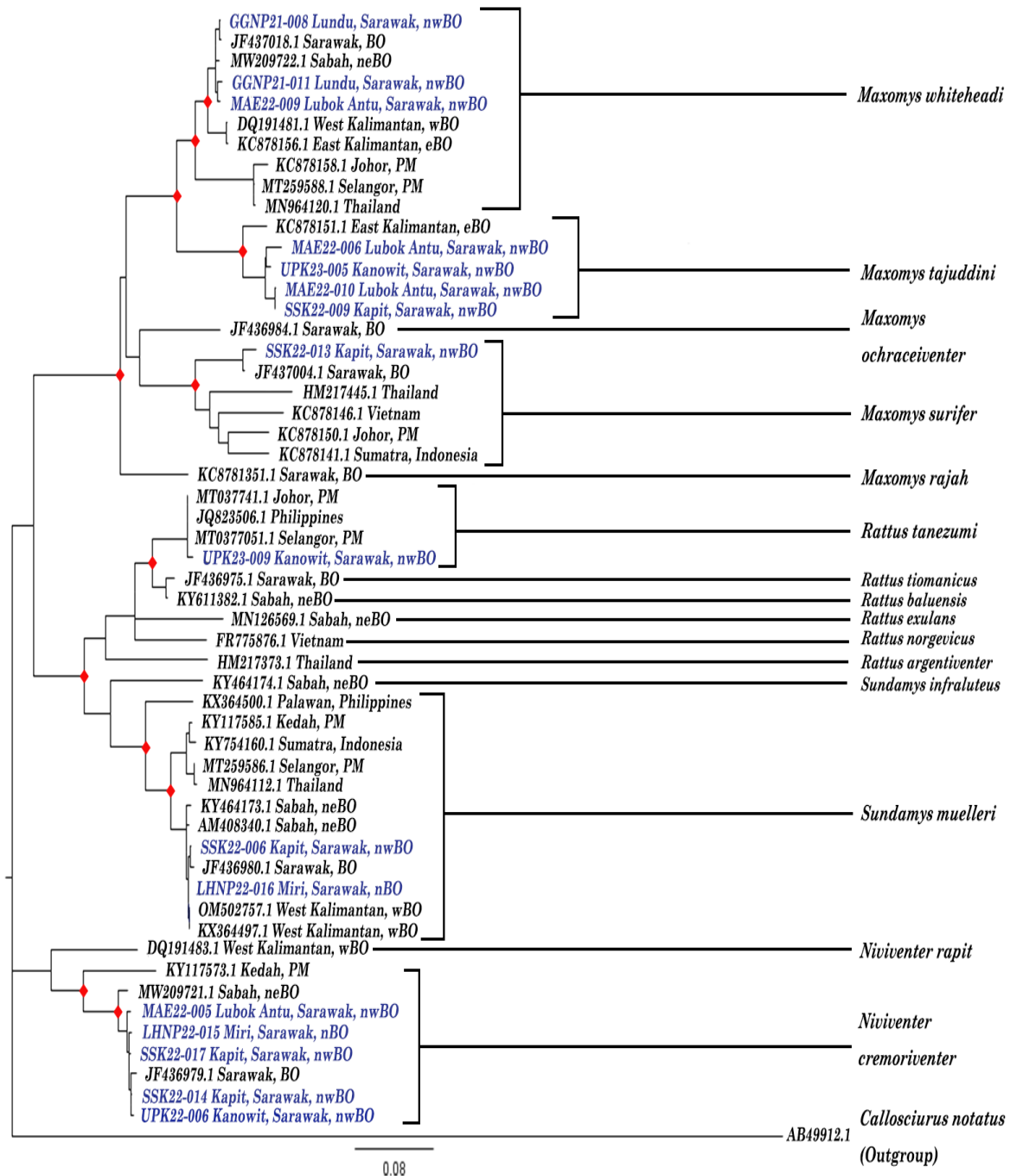


Figure 3.4: Maximum-Likelihood Phylogenetic Tree of Sequenced Rodent Individuals Highlighted in Blue. Red Diamonds Indicate Nodes With Bootstrap Values >90 (PM = Peninsular Malaysia, BO = Borneo, nBO = North Borneo, nwBO = Northwest Borneo, neBO = Northeast Borneo, wBO = West Borneo, eBO = East Borneo)

The analyses of the 16S sequences obtained from nanopore sequencing revealed the detection of several bacterial classifications at different taxonomic levels. The total number of sequences after normalisation represent the combined gut microbiome of all rodent individuals (Figure 3.5). Overall, the most abundant phylum of bacteria across all rodent species is Bacillota (formerly Firmicutes) which represents around 78% of the total abundance. The three most dominant bacterial families are represented by Lachnospiraceae (~18%), Lactobacillaceae (~20%) and Oscillospiraceae (~19%). At the genus level, a total of 1052 bacterial genera were detected (Appendix D), with the most abundant consisting of *Ruminococcus* (12.2%), *Ligilactobacillus* (8.93%), and *Lactobacillus* (7.19%) (Table 3.3). Moreover, the most prominent bacterial species found in the combined gut microbiota are *Ruminococcus albus* (5.82%), *Ligilactobacillus murinus* (4.88%) and *L. animalis* (3.75%). Several potentially pathogenic bacterial genera were also present at low abundances, including *Bacillus* (1.07%), *Bacteroides* (1.06%), *Clostridium* (2.19%), *Escherichia* (1.43%) and *Klebsiella* (1.89%). When comparing each rodent individual, their gut microbiome composition did not adhere to a strict pattern (Figure 3.6). The relative abundance of these predominant genera differs from each rodent individual. In particular, one juvenile rodent individual from Lambir Hills National Park (L1) had genus *Ligilactobacillus* (77.8%) as the overwhelming dominant genus.

Table 3.3: Relative Abundance Table in Percentages of the Top 23 Most Abundant Bacterial Genera (>1% Overall Relative Abundance) According to Rodent Species

Bacterial Genera	<i>Maxomys surifer</i> (n=1)	<i>Maxomys tajuddini</i> (n=4)	<i>Maxomys whiteheadi</i> (n=3)	<i>Niviventer cremoriventer</i> (n=5)	<i>Rattus tanezumi</i> (n=1)	<i>Sundamys muelleri</i> (n=2)	Overall
<i>Ruminococcus</i>	9.69	5.68-26.8	2.59-12.5	0.76-14.2	5.54	22.5-40.4	12.2
<i>Ligilactobacillus</i>	0.26	0.03-0.12	0.06-0.10	0.03-77.8	0.38	0.03-0.09	8.93
<i>Lactobacillus</i>	11.61	3.64-9.17	11.6-25.5	0.03-1.49	4.66	1.00-8.11	7.19
<i>Blautia</i>	5.37	1.91-8.80	0.97-5.49	0.45-11.38	3.62	1.93-3.59	4.07
<i>Lachnoclostridium</i>	5.87	2.91-12.1	1.27-3.97	0.41-7.80	2.96	1.87-2.59	3.68
<i>Flintibacter</i>	3.12	1.82-5.28	0.66-2.76	0.24-2.80	6.22	3.24-4.80	2.80
<i>Limosilactobacillus</i>	1.59	0.75-5.46	2.10-4.04	0.03-4.33	6.12	1.11-3.42	2.62
<i>Flavonifractor</i>	2.18	2.40-3.48	1.31-3.00	0.22-4.00	4.82	2.43-3.79	2.52
<i>Clostridium</i>	3.57	1.09-3.05	1.09-6.40	0.34-3.03	1.90	1.11-2.11	2.19
<i>Duncaniella</i>	1.90	0.64-2.13	2.45-3.81	0.39-5.74	4.77	0.56-0.91	2.04
<i>Klebsiella</i>	0.06	0.38-3.35	1.00-7.38	0.07-11.9	0.09	0.06	1.89
<i>Anaerostipes</i>	0.96	1.41-4.21	0.55-1.34	0.26-6.39	1.55	1.37-2.31	1.72
<i>Muribaculum</i>	1.61	0.52-1.89	2.17-3.32	0.34-4.53	3.30	0.33-0.54	1.70
<i>Prevotella</i>	0.68	0.37-2.80	0.83-4.52	0.06-2.80	3.38	1.54-2.72	1.59
<i>Romboutsia</i>	4.85	0.02-4.06	0.02-11.1	0.01-1.65	0.13	0.03-0.14	1.49
<i>Lacrimispora</i>	1.03	0.92-3.13	0.20-1.45	0.40-2.82	1.53	1.10-3.40	1.48
<i>Escherichia</i>	0.31	0.29-3.13	0.84-1.71	0.13-3.73	0.18	0.16-0.27	1.43
<i>Intestinimonas</i>	2.26	1.01-2.23	0.52-2.35	0.10-1.44	1.85	1.82-1.86	1.35
<i>Lachnospira</i>	0.82	0.21-3.42	0.07-0.29	0.06-3.98	0.77	0.63-1.39	1.16
<i>Phascolarctobacterium</i>	0.74	0.00-3.59	0.00-5.61	0.00-0.17	0.00	0.69-2.25	1.08
<i>Bacillus</i>	1.06	0.79-1.28	0.82-1.08	0.98-1.55	1.11	0.61-1.03	1.07
<i>Bacteroides</i>	1.92	0.25-2.63	0.87-3.62	0.28-0.94	1.02	0.38-0.61	1.06
<i>Anaerobutyricum</i>	0.52	0.53-3.47	0.14-0.87	0.12-2.69	1.18	0.64-1.26	1.00

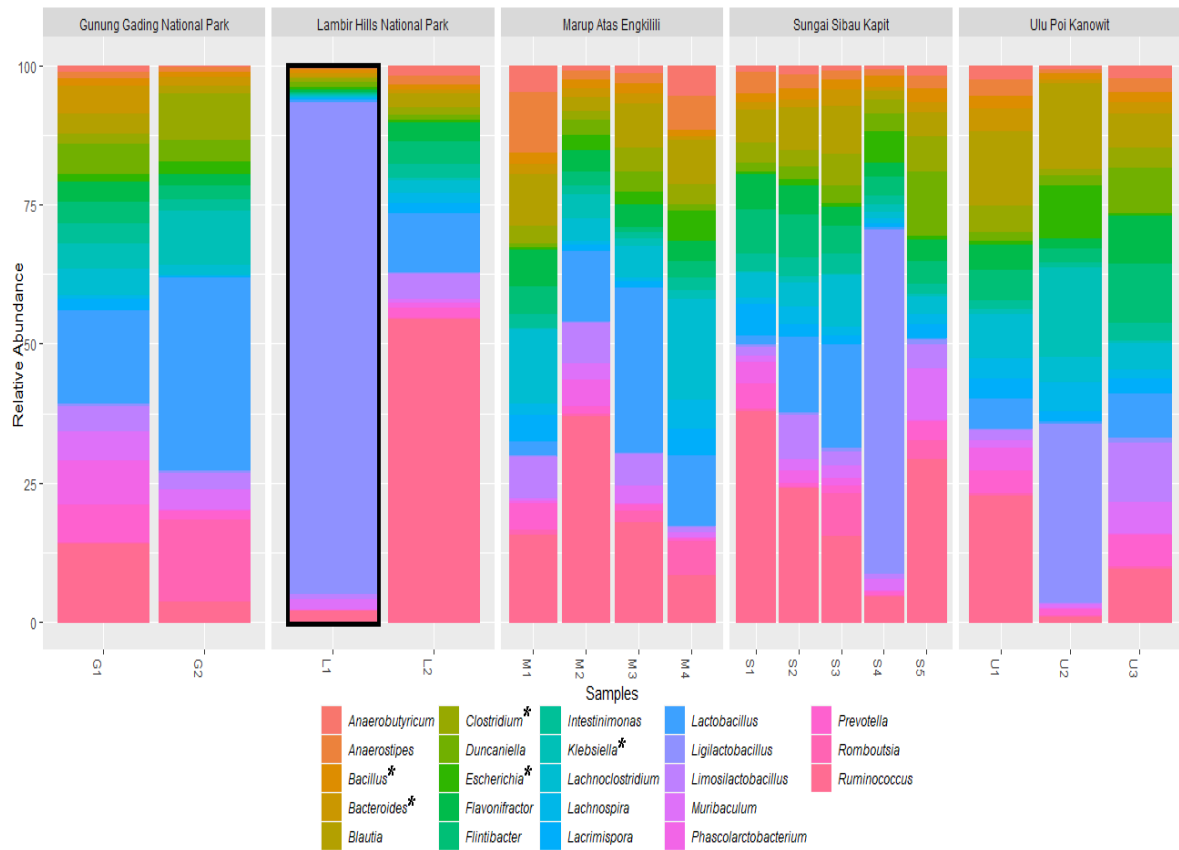


Figure 3.6: Abundance Bar Plots of the Top 23 Most Abundant Bacterial Genera From Rodent Individuals Across Different Localities. The Top 23 Genera Were Selected as They Represent >1% of the Relative Abundance of the Total Bacterial Abundance. Potentially Pathogenic Bacterial Genera Marked With '*'. Juvenile *N. cremoriventer* Individual Outlined in Black. Abbreviations of Samples Follow Appendix A

The estimated alpha diversity, based on the Shannon diversity index values, displayed variation among the gut microbiomes of all rodent individuals. The diversity indices range from 2.2 in Lambir Hills National Park to 5.0 in Sungai Sibau Kapit (Figure 3.7). Non-parametric Kruskal-Wallis test was used to assess the differences between alpha diversity according to each study site and revealed these differences were not significant ($p < 0.05$).

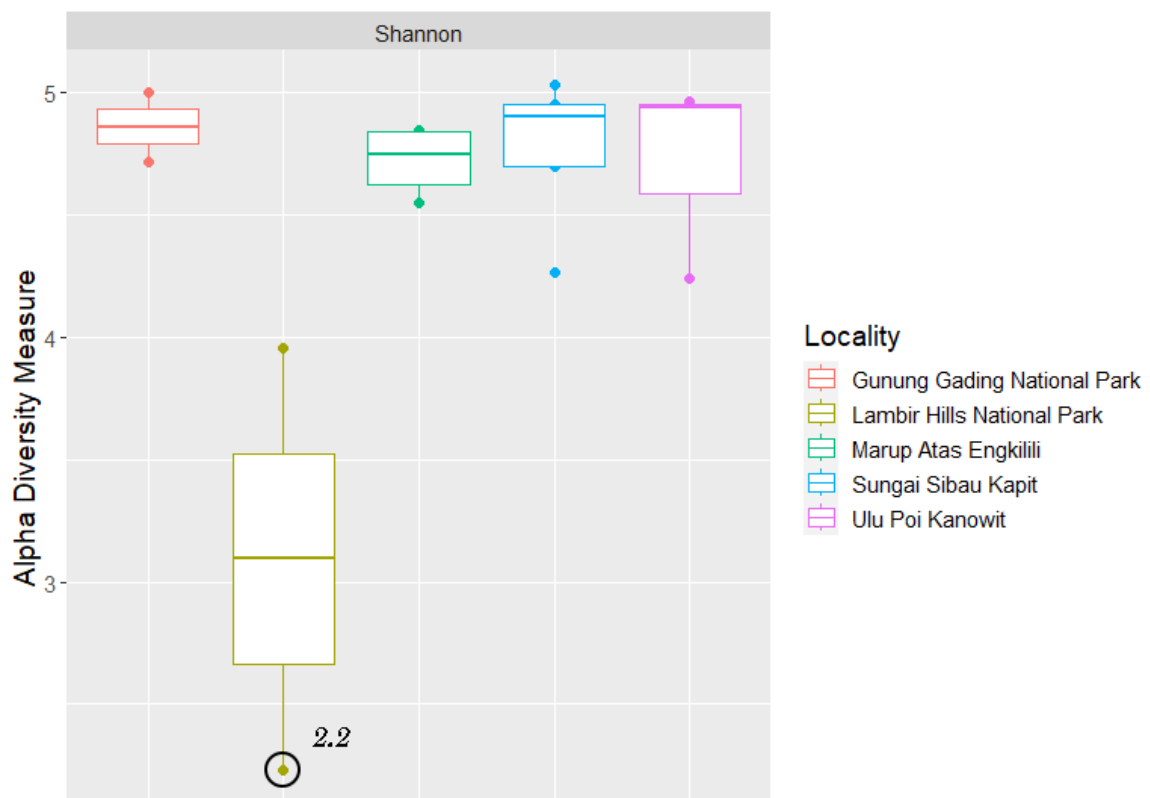


Figure 3.7: Box Plot of Shannon Diversity Index of the Gut Microbiome of Rodent Individuals Across Different Study Sites. Juvenile *N. cremoriventer* Individual is Labelled and Circled in Black

3.4 Discussion

The gut microbiomes of various rodent individuals in this study exhibit patterns of variation, suggesting the presence of distinct bacterial communities. Upon comparison with other rodent microbiota studies, the consistent prevalence of the Bacillota phylum emerges as a notable finding within the microbiome compositions of wild rodents (Weldon et al., 2015; Jahan et al., 2021; Teng et al., 2022). Consistent with our findings, these investigations have identified Bacillota, formerly recognised as Firmicutes, as one of the most abundant bacterial phyla within the gut microbiota of wild rodents. These microbes are commonly found within the mammalian gut microbiota, contributing to the regulation of the host's immune system and the development of the gut epithelium (Leser & Mølbak, 2009). Therefore, members of the phylum Bacillota could be considered as members of the core gut microbiome of rodents, due to their prevalence across the gut microbiota of multiple rodent individuals and their significant involvement in physiological processes.

Microbial diversity analysis revealed the most abundant families are Oscillospiraceae, Lachnospiraceae and Lactobacillaceae. The Oscillospiraceae family is also recognised as Ruminococcaceae, which is considered as a heterotypic synonym to Oscillospiraceae (Euzéby, 2010). Members of this bacterial family, alongside Lachnospiraceae, are common in the gut environment and plays a vital role in digestion and nutrient absorption. They were found to be specialised for degrading complex plant material through enzymatic activities (Biddle et al., 2013). The genus *Ruminococcus* (family Oscillospiraceae) is the most abundant genus observed across all rodent individuals in this study. Studies have shown that *Ruminococcus* is a constituent of the rodent core gut bacteria based on their significant abundance (Wang et al., 2019). The high prevalence of this core

bacteria observed in wild rodents could have potential importance to their feeding behaviours (Kohl et al., 2022).

Lachnospiraceae is also a prominent bacterial family and was represented by multiple genera at varying abundances. Among these, the genus *Blautia* was observed as the most highly detected genus. Liu et al. (2021) have highlighted the nutritional benefits of some *Blautia* species for their probiotic properties in mammals. However, they have also stated that further research is required due to the lack of comprehensive review concerning this genus. Although it was observed at relatively low abundance, another noteworthy bacterial genus within the same Lachnospiraceae family is *Anaerobutyricum*. Similar to *Blautia*, certain strains of *Anaerobutyricum* could potentially serve as a probiotic and recent studies have undergone preclinical testing in mice models for medical use (Wortelboer et al., 2022). The presence of various Lachnospiraceae members within the gut microbiota suggests their pivotal role as a component of the bacterial community, and further research on the significance of this family could contribute to a better understanding on beneficial gut bacteria.

Another significant bacterial family is Lactobacillaceae which recorded the highest abundance for the overall combined gut microbiome of wild rodents. This family represents the lactic acid bacteria that are responsible for metabolising lactose within the gut environment (König & Fröhlich, 2017). This includes the genus *Lactobacillus* which were found to be significantly higher in reproductively active wild female rodents (Maurice et al., 2015). Notably, the gut microbiota of a juvenile *Niviventer cremoriventer* individual from Lambir Hills National Park was observed to be predominantly *Ligilactobacillus*, another genus from the family Lactobacillaceae that was formerly known as the *Lactobacillus salivarius* group (Zheng et al., 2020). It is possible that *Ligilactobacillus* could share similar

properties with *Lactobacillus* as lactic acid bacteria. The relative abundance of *Ligilactobacillus* in the other adult *N. cremoriventer* individuals were noticeably lower (0.03% to 39.8%) than the juvenile individual (77.8%). The prevalence of this genus in the gut of the young rodent could imply its significance during infancy, potentially contributing to the functional maturation of its intestinal microbiome (Schaedler et al., 1965; Peña et al., 2004).

Several potentially pathogenic bacteria were detected at low abundance in multiple rodent individuals from various localities, including *Bacteroides* (1.06%), *Clostridium* (2.19%), *Klebsiella* (1.89%), *Escherichia* (1.43%), *Bacillus* (1.07%) and *Bacteroides* (1.06%). Along with other genera, these bacteria represent approximately 11% of the combined rodent gut microbiota. Similarly, Jahan et al. (2021) reported the detection of pathogenic bacteria in peridomestic rodents, which represents around 15% of the overall gut microbiome. While the combined abundance of these pathogens is lower than that of the presumed core microbiome constituents, their prevalence in all wild rodent individuals indicates the potential existence of a commensal relationship between them. These bacteria may possess the capability to coexist within the established bacterial community in the rodent gut microbiota, capitalising on the nutrient-rich environment derived from the gut microbiota (Bäumler & Sperandio, 2016). However, the introduction of pathogenic bacteria into the gut could lead to the acquisition of resistance or the facilitation of infection, driven by the alteration of the existing intestinal environment (Bäumler & Sperandio, 2016). Nevertheless, rodents are recognised as reservoir hosts for pathogenic bacteria, and conducting specific analyses that emphasise the prevalence of pathogenic bacteria occurring at low abundances could lead to a better understanding of their transmission dynamics.

However, it is apparent that the host rodent species plays a substantial role in shaping their gut microbiota. For *Maxomys* rodents, *Lactobacillus* is one of the most dominant taxa with a relative abundance ranging from 3.64% to 25.50%. In contrast, for other rodent genera (*Niviventer*, *Sundamys* and *Rattus*), the relative abundance of *Lactobacillus* falls between 0.03% to 8.11%. The monophyletic relationship among *Maxomys* individuals implies that the phylogenetic similarity of rodent species may reflect their gut microbiome composition. According to Wang et al. (2022), microbiota within a species tends to exhibit greater similarity to that of conspecifics in different locations than to individuals from closely related species residing in the same geographical area. Nevertheless, due to the higher number of rodent individuals within genus *Maxomys*, this assumption needs to be treated with caution. On the other hand, the life stage of the rodents seems to affect their gut microbiota based on the observed low alpha diversity exhibited by the juvenile *N. cremoriventer* individual from Lambir Hills National Park. As highlighted by Fenn et al. (2023), microbiome alpha diversity increases with age for wild rodents as there is a positive shift in species richness. This pattern might arise due to the incomplete development of the bacterial community within the gut environment of young rodents.

The characterisation of the combined core gut microbiota in this study may not accurately reflect the microbiome composition of distinct rodent species. Given the non-invasive nature of faecal sampling, it is recommended to obtain a larger sample size that encompasses a diverse range of wild rodent species and accounts for different reproductive stages. This approach would yield a more comprehensive understanding on the dynamics of the host-specific core gut microbiome, especially with an increased number of representatives for each species. Furthermore, including more localities with distinct habitat types and incorporating additional environmental data into the characterisation of the

rodents' gut microbiome would provide insights into the constituents that influence the structure of the bacterial community (Lobato-Bailón et al., 2023).

3.5 Conclusion

This study effectively characterises the gut microbiome of multiple wild rodents from forested areas in Sarawak, Borneo and a total of 1052 bacterial genera were detected in the gut microbiota from 16 individuals of rodents belonging to six different species. There are three bacterial families of importance (Lachnospiraceae, Lactobacillaceae, Oscillospiraceae), all falling under the phylum Bacillota. Members of these families have potential to be constituents of the core rodent gut microbiome due to their prevalence throughout various rodent individuals. Furthermore, it appears that the gut microbiome composition of wild rodents is influenced by the host rodent species and their age, as evidenced by the abundance patterns of certain taxa.

3.6 Recommendation

Based on the outcome of this study, the gut microbiome compositions of specific rodent species were not accurately represented due to the low number of samples caught for each species. Expanding the sampling efforts to include a larger number of samples would substantially enhance the depth and accuracy of the research. Additionally, while this study indicates that certain environmental influences shape the gut microbiota of wild rodents, it overlooks potential interactions between the rodents and their surrounding habitat. Further investigation into additional environmental determinants could reveal the intricate relationships between gut bacterial communities and their host organisms. This includes exploring the rodents' dietary preferences, which could reveal the significance of the native vegetation in their habitat.

CHAPTER 4

PREVALENCE OF POTENTIALLY PATHOGENIC BACTERIA IN WILD

RODENTS NEAR NATURAL WATER BODIES FROM SARAWAK

4.1 Introduction

Rodents are widely considered as the reservoir hosts for zoonotic diseases as they can be carriers of different types of pathogenic bacterial species (Dahmana et al., 2020). Pathogenic bacteria that can cause zoonotic diseases, are able to thrive in the mammalian gut due to several factors such as the presence of suitable nutritional environments, host immune system interactions and niche availability (Bäumler & Sperandio, 2016; Greene et al., 2020; Kubasova et al., 2021). While there have not been reports of recent disease outbreaks attributed to rodents in Sarawak, they were regarded as the principal reservoirs for leptospirosis infections in Peninsular Malaysia in recent years (Benacer et al., 2016). They are known to be transmitters of diseases due to their synanthropic nature as they can co-exist near human settlements and adapt to changing activities (Kucheruk, 1965). This proximity increases the chance of interactions between rodents and humans, creating opportunities for disease transmission. Rodents can transmit these pathogens directly to humans through bites, excrements such as saliva and faeces, or through indirect routes which involves other animals such as arthropods or livestock that comes into contact with them (Meerburg et al., 2009).

Wild forest rodents could exhibit a diverse gut microbiota due to their varied diets and exposure to a wider range of environmental microbes based on their habitat (Viney, 2019). This diverse natural environment may also serve as a source for the discovery of novel pathogenic bacteria that have yet to be identified. Notably, Lau et al. (2020) presented new

geographical evidence of novel pathogenic bacteria in rodent populations from forested areas in Sarawak. Their findings highlight the significance of the boundaries between forest fringes and human settlements, as they can act as the interfaces where infected rodent individuals come into contact with humans that potentially leads to the transmission of zoonotic diseases (Bordes et al., 2015; Dahmana et al., 2020).

While specific environmental features that directly influence the gut microbiome of wild rodents remain largely unknown, naturally occurring water bodies could be one of the contributors in shaping their gut microbiota. This can be attributed to the presence of organic matter and sediments within these water bodies, which provide favourable conditions for bacterial growth (Vital et al., 2010; Mwanamoki et al., 2014). The gut of wild rodents could act as a reservoir for a wide range of bacteria, and the diverse microbial populations in water sources may contribute to the diversity of the rodent's gut microbiota. According to reports, antimicrobial-resistant *Listeria monocytogenes* bacteria were detected in water samples, with indications that it may have spread to the surrounding wildlife, including a rodent individual (Lesley et al., 2016). The interaction between water bodies and the wildlife highlights the potential relationship between environmental microbiota and the host, further emphasising the significance of water sources in the diversification of the gut microbiota in wild rodents.

While recent studies focus on conventional sequencing methods to detect specific strains of pathogenic bacteria (Bilung et al., 2014; Lesley et al., 2016; Pui et al., 2017), recent advancements in next-generation sequencing (NGS) technologies have revolutionised the study and characterisation of complex microbiomes (Jahan et al., 2021). Nanopore sequencing is a portable, real-time NGS technology that produces long DNA sequence reads, enabling rapid and accurate identification of bacterial species without the need for extensive

culturing methods (Borges et al., 2023). The utilisation of nanopore sequencing enables the comprehensive study of complete microbial communities in organisms like rodents, thereby facilitating the identification of certain pathogenic bacteria that may otherwise go undetected using traditional methods (Ciuffreda et al., 2021). Furthermore, the characterisation of the rodent gut microbiota could provide insights into their potential role in the transmission of pathogenic bacteria.

Understanding the gut microbiome of wild rodents is essential for assessing its implications for the transmission of pathogenic bacteria to nearby human settlements. Additionally, the intricate ecosystem within the rodent gut microbiome could be influenced by their surrounding environment (Amato et al., 2019; Viney, 2019; Wang et al., 2022). Natural water bodies, in particular, could play a significant role in shaping this bacterial diversity due to their potential to harbour diverse microbial communities (Barcina et al., 1997; Vital et al., 2010). Therefore, the objective of this study is to determine the prevalence of potentially pathogenic bacteria found in rodents trapped from forested areas. This study provides valuable insights into the potential transmission of zoonotic diseases, particularly those caused by novel pathogenic bacteria.

4.2 Materials and Methods

4.2.1 Study Sites

The study sites consist of several forested areas throughout Sarawak that have not undergone much anthropogenic pressure to maximise the possibilities to recover novel bacteria that are not previously found in urban environments. At the same time, the study sites are in close proximity to human settlements and there were reports of pathogenic bacteria detection in rodent populations from some of these areas. These sites include both

protected areas, such as Gunung Gading National Park and Lambir Hills National Park, as well as non-protected areas including Marup Atas Engkilili, Sungai Sibau Kapit and Ulu Poi Kanowit (Figure 4.1). The habitat types and water bodies for all study sites were taken into consideration and described (Table 4.1).

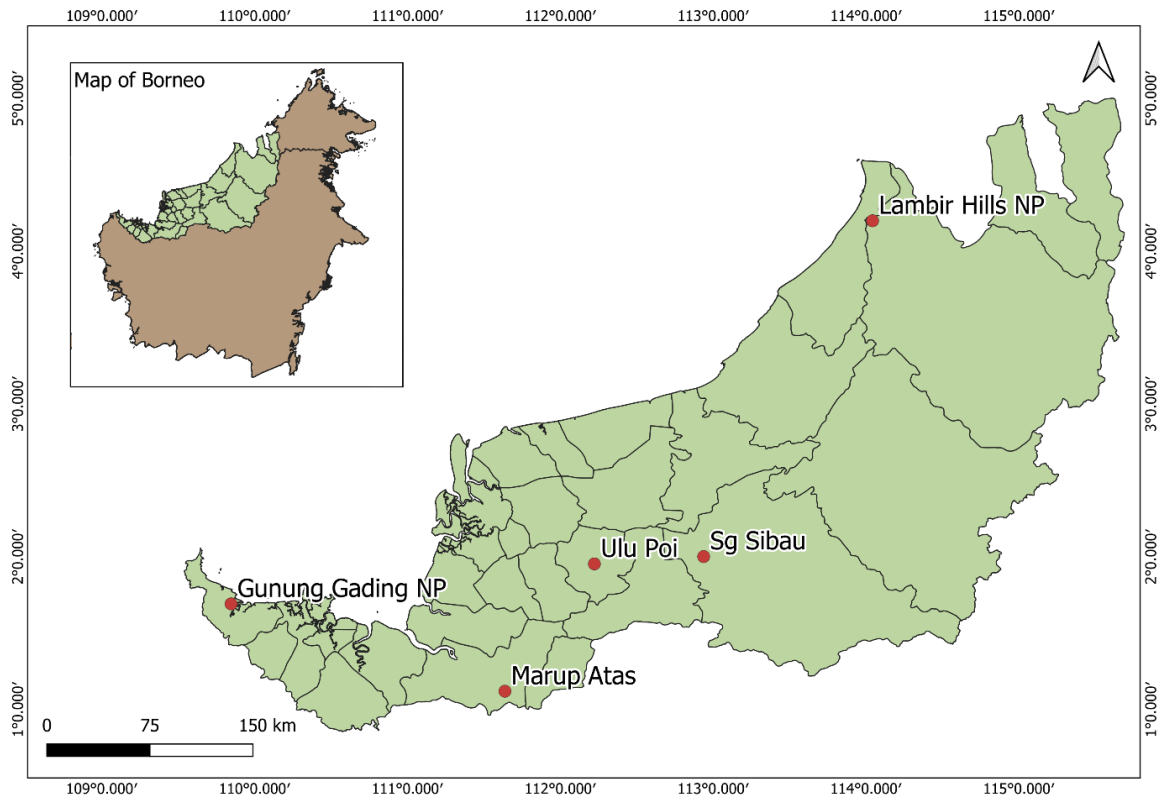


Figure 4.1: Map of Sampling Sites Indicated by Red Dots in Sarawak, Borneo

Table 4.1: Environmental Information and Description of Sampling Sites in Sarawak

Locality	Gunung Gading National Park	Marup Atas Engkilili	Lambir Hills National Park	Sungai Sibau Kapit	Ulu Poi Kanowit
District	Lundu	Lubok Antu	Miri	Kapit	Kanowit
Protected Area	Yes	No	Yes	No	No

Habitat type	Forested area at the outskirts of township	Forested area with nearby human settlement	Forested area away from human settlement	Forested area with nearby human settlement	Forested area with nearby human settlement
Water body description	Flowing river near HQ of park; used for recreational activities	Flowing river at the outskirts of village; used by locals as water source	Flowing river at main waterfall trail; used for recreational activities	Flowing river at the outskirts of village; used by locals as water source	Fast flowing river adjacent to village; used by locals for transportation

4.2.2 Sampling Methods

A total of 120 Sherman traps and 50 cage traps were used in rodent trapping. The traps were sanitised with 70% ethanol and bleach beforehand and were baited with banana fruits covered with peanut butter and placed near natural water bodies. They were placed at least 5 metres away from natural water bodies to increase the likelihood that captured rodents would have interacted with the surrounding water. The traps were checked from 0900 hours to 1200 hours and were re-baited every two days. To reduce cross contamination of samples, captured rodents were transferred to separate sterilised containers for faecal sample collection (multiple 1 to 2 cubic mm pellets). Morphological measurements (sex, weight, ear, hindfoot, tail ventral and head body) of the captured rodents were also taken and were then identified based on descriptions and keys from field guides (Payne et al., 1985; Phillipps & Phillipps, 2018). The faecal samples from the captured individuals were preserved in RNAlater (Ambion) for the gut microbiome analysis.

4.2.3 Water Quality Sampling

Water quality parameters were obtained from the water bodies at each sampling site using a YSI Professional Pro Plus multiparameter metre (YSI Ohio, USA). The natural water bodies from all sampling sites were identified as flowing rivers. Before water collection, 300 ml glass bottles were washed with distilled water and subsequently rinsed three times using water from the specific river intended for collection. Measurements were done by immersing the probe in the collected water for a few minutes until a stable reading was obtained. The parameters assessed include dissolved oxygen (DO), biochemical oxygen demand (BOD), total dissolved solids (TDS), salinity, conductivity and pH which have been identified as predominant factors that can potentially shape microbial communities in water bodies (Staley et al., 2018; Wang et al., 2018; Li et al., 2021). Standard methods by Rice et al. (2012) were adapted for the collection and analysis of these water quality parameters. Measurements were conducted at three distinct points within each water body to evaluate potential variability in readings, and the average of the replicate readings was obtained. The different points were selected based on accessibility and are of equal distances between each other whenever possible. The probe was also rinsed with distilled water before each subsequent measurement to ensure the metre was not affected by previous readings.

4.2.4 DNA Extraction and Next-Generation Sequencing

DNA was extracted from the collected faecal samples using QIAamp PowerFecal Pro DNA Kit (Qiagen, Hilden, Germany) according to manufacturer's protocol. Quantification of the extracted faecal DNA was then carried out using DeNovix DS-11 Spectrophotometer (Denovix, Delaware, USA). The 16S rRNA gene was amplified by polymerase chain reaction (PCR) using the universal primers, 27F and 1482R included in

the 16S Barcoding Kit (SQK-RAB204) from Oxford Nanopore Technologies (ONT, Oxford, United Kingdom). The same barcoding kit was used for library preparation and the flow cell (R9.4.1) was loaded and primed using the Flow Cell Priming Kit (EXP-FLP002). Nanopore sequencing was then carried out using a MinION Mk1C sequencer (ONT) with each sequencing run lasting around 17 hours.

4.2.5 Data Analysis

Guppy 5.1.12 (ONT), which is the basecaller integrated in the MinKNOW 21.11.6 (ONT) operating system, was used to obtain basecalled FASTQ files of the 16S rRNA sequences. Taxonomic classification was done up to the species level according to the National Center for Biotechnology Information (NCBI) taxonomic records using the application EPI2ME (ONT). The data was normalised by rarefying method available in the R package phyloseq v1.42.0 (McMurdie & Holmes, 2013) and was filtered to only include potentially pathogenic bacterial genera. The ggplot2 v3.4.2 package (Wickham, 2011) was then used to produce an abundance graph of bacterial genera according to different localities and the shared species between sites was visualised using a Venn diagram that was produced using the MicEco v0.9.19 (Russel, 2021) package. Non-metric multidimensional scaling (NMDS) based on the Jaccard distance metric was also done using the phyloseq v1.42.0 package (McMurdie & Holmes, 2013). Lastly, Pearson correlation was used to assess the relationship between the diversity of the gut microbiome and the abundance of pathogenic bacteria with the various water quality parameters. Additionally, correlograms based on the correlation analysis were generated using the corrplot v0.92 package (Wei et al., 2017).

4.3 Results

Analyses of the gut microbiome composition in trapped rodents revealed several potentially pathogenic bacterial genera (Table 4.2). This includes genus, *Clostridium*, *Klebsiella*, *Escherichia*, *Bacillus* and *Bacteroides*, all of which were present in all sampling sites but at varying abundances. These bacteria were detected at relatively low abundances compared to the other gut bacteria, with only four out of the 23 bacterial genera (representing more than 1% of the overall relative abundance) can be considered potentially pathogenic. Furthermore, additional potentially pathogenic bacterial genera are present at even lower abundances and could only be observed after filtering out the other gut bacteria from the dataset.

Table 4.2: Abundance Table of the Top 23 Most Abundant Bacterial Genera (>1% Overall Relative Abundance) According to Different Localities. Potentially Pathogenic Bacterial Genera Marked With ‘*’

Bacterial Genera	Gunung Gading National Park	Lambir Hills National Park	Marup Atas Engkilili	Sungai Sibau Kapit	Ulu Poi Kanowit
<i>Ruminococcus</i>	3331	11412	14590	17587	5675
<i>Ligilactobacillus</i>	43	20944	46	10949	6449
<i>Lactobacillus</i>	9978	2203	10657	5877	2239
<i>Blautia</i>	907	639	4943	4636	6406
<i>Lachnoclostridium</i>	1249	615	7208	3725	3054
<i>Flintibacter</i>	1266	935	1957	4721	3158
<i>Limosilactobacillus</i>	1462	1046	3922	2884	1975
<i>Flavonifractor</i>	1004	715	3278	3434	2406
<i>Clostridium*</i>	2015	391	2233	3281	1500

<i>Duncaniella</i>	1757	413	1508	3241	1871
<i>Klebsiella</i> *	2825	54	1434	483	3337
<i>Anaerostipes</i>	379	439	3589	1849	1130
<i>Muribaculum</i>	1673	558	1324	2498	1258
<i>Prevotella</i>	1488	429	1354	1624	1946
<i>Romboutsia</i>	2989	9	1540	1819	56
<i>Lacrimispora</i>	447	404	2078	2074	1388
<i>Escherichia</i> *	688	208	1904	1344	2021
<i>Intestinimonas</i>	990	528	1182	2170	934
<i>Lachnospira</i>	93	390	1345	1322	1847
<i>Phascolarctobacterium</i>	1518	185	1012	1225	710
<i>Bacillus</i> *	474	427	1107	1685	898
<i>Bacteroides</i> *	1206	178	925	1197	1057
<i>Anaerobutyricum</i>	226	370	2033	870	817

The filtered pathogenic bacteria constitute approximately 11% of the overall bacterial abundance (around 47000 reads). Variations in the abundance of pathogenic bacterial genera were observed among different localities (Figure 4.2; Appendix E). In Gunung Gading National Park (GGNP) and Ulu Poi Kanowit (UPK), the genus *Klebsiella* exhibited the highest abundance with 2825 and 3337 reads, respectively. Conversely, in Marup Atas Engkilili (MAE) and Sungai Sibau Kapit (SSK), *Clostridium* emerged as the most abundant taxon with 2233 and 3281 reads, respectively. This genus is also consistently ranked among the top three dominant genera across all study sites. For Lambir Hills National Park (LHNP), *Helicobacter* was found to be the most abundant genus with 707 reads. While *Helicobacter*

is also dominant in SSK (2027 reads), it is present at very low abundance in other localities. Furthermore, the taxa consolidated into ‘Others’ encompasses 39 different potentially pathogenic bacterial genera (Appendix E). The bacterial genera under ‘Others’ were detected at very low overall relative abundance, ranging from 0.968% (*Vibrio*) to 0.004% (*Nocardia*).

A significant overlap of bacterial species was observed, with 191 species shared among them (Figure 4.3). UPK exhibited the highest diversity of bacterial species, with a recorded count of 537 identified species. In contrast, only 277 species of bacteria were detected in LHNP, the lowest out of all the localities. A similar pattern can be seen when observing the number of species exclusively found in each study site. The highest number of unique species of bacteria was in UPK, with a total count of 70 species. Conversely, LHNP contained the lowest number of exclusive bacterial species, with only 10 identified species. The NMDS ordination plot (Figure 4.4) revealed that there is an overlap in gut microbiome composition of rodents across all localities as the plots are clustered close together.

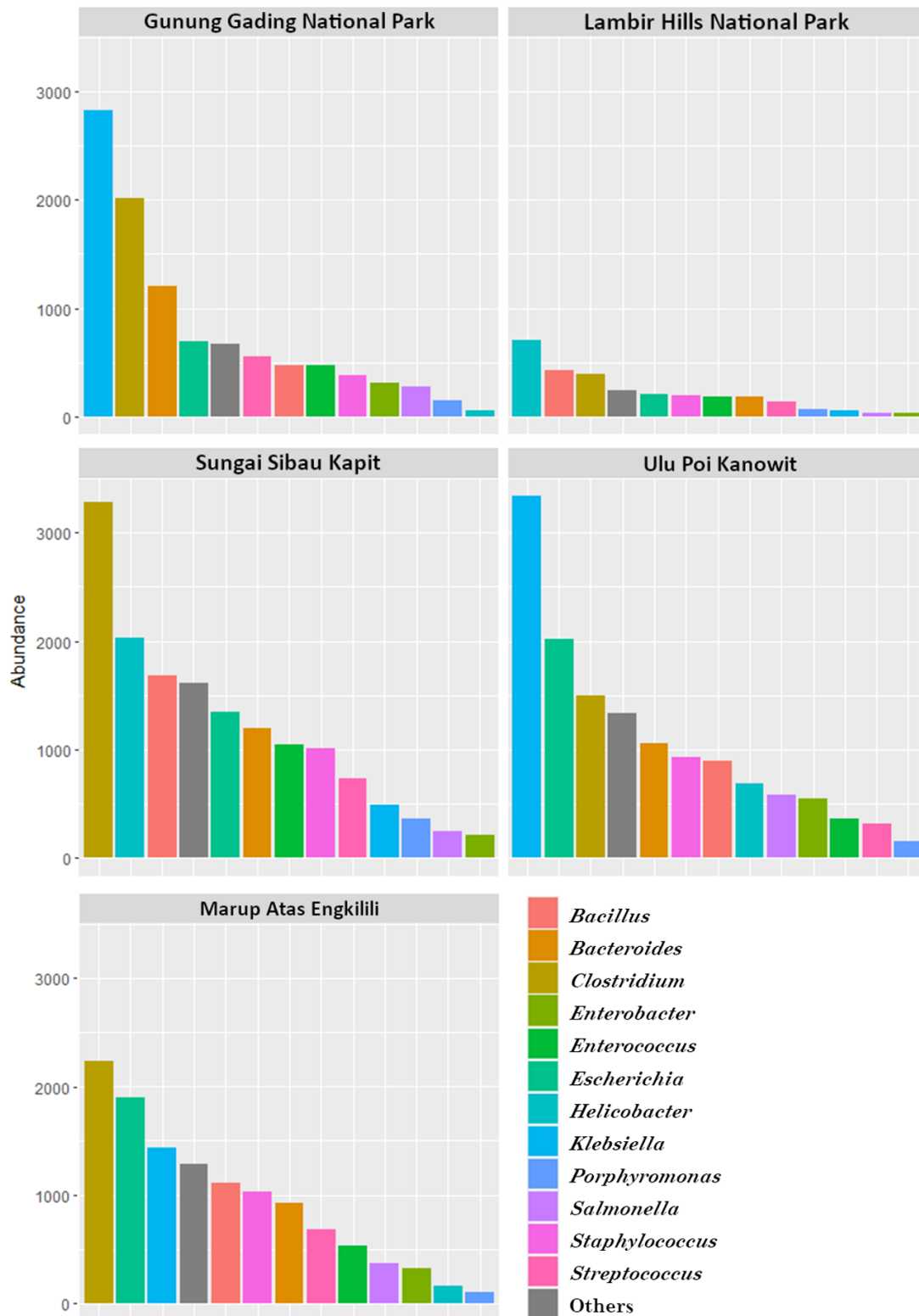


Figure 4.2: Abundance Bar Plots of the Top 12 Most Abundant Potentially Pathogenic Genera Across Different Localities. The Selected Genera Represented >1% of the Relative Abundance of the Total Number of Pathogenic Bacteria While the Remaining Genera Below the Threshold Were Consolidated and Labelled as ‘Others’

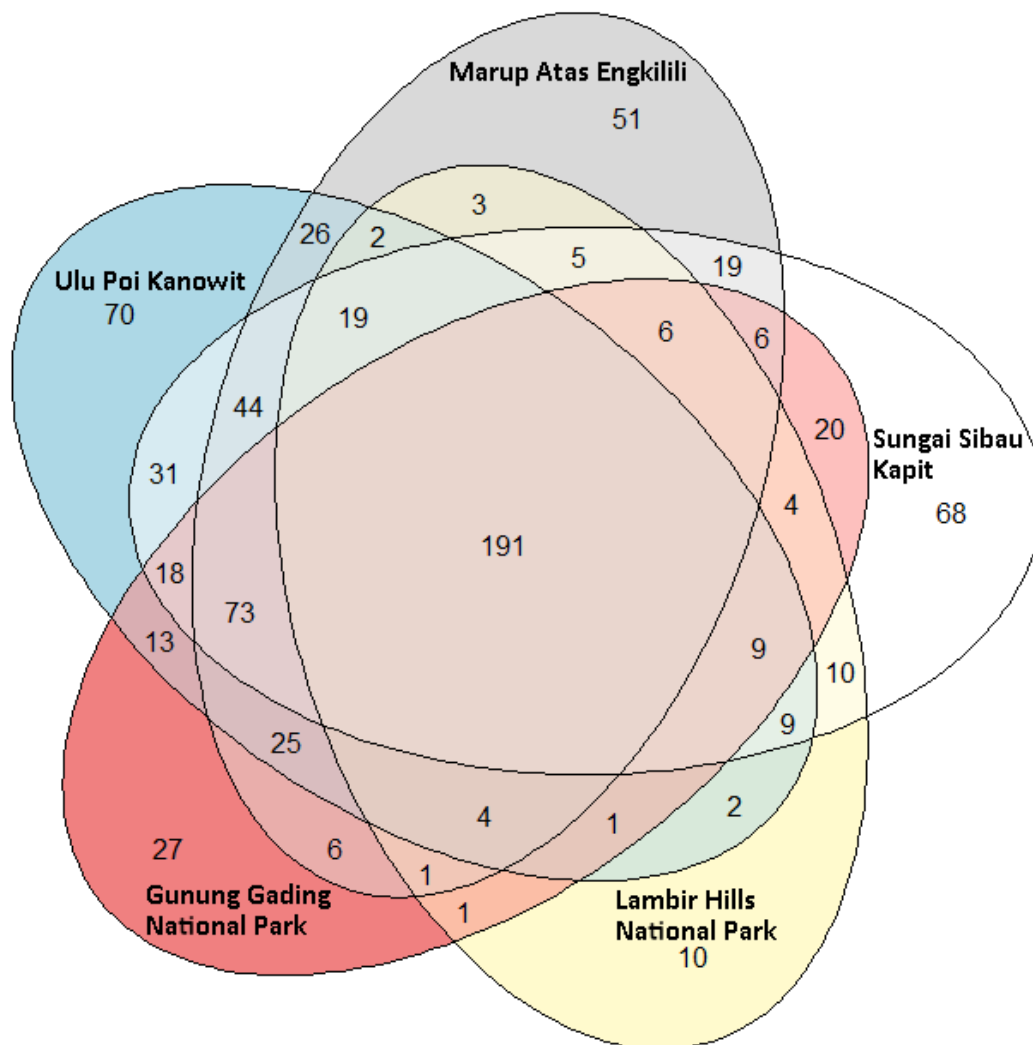


Figure 4.3: A Venn Diagram Showing the Number of Bacterial Species Being Detected Between Five Localities of Study Sites

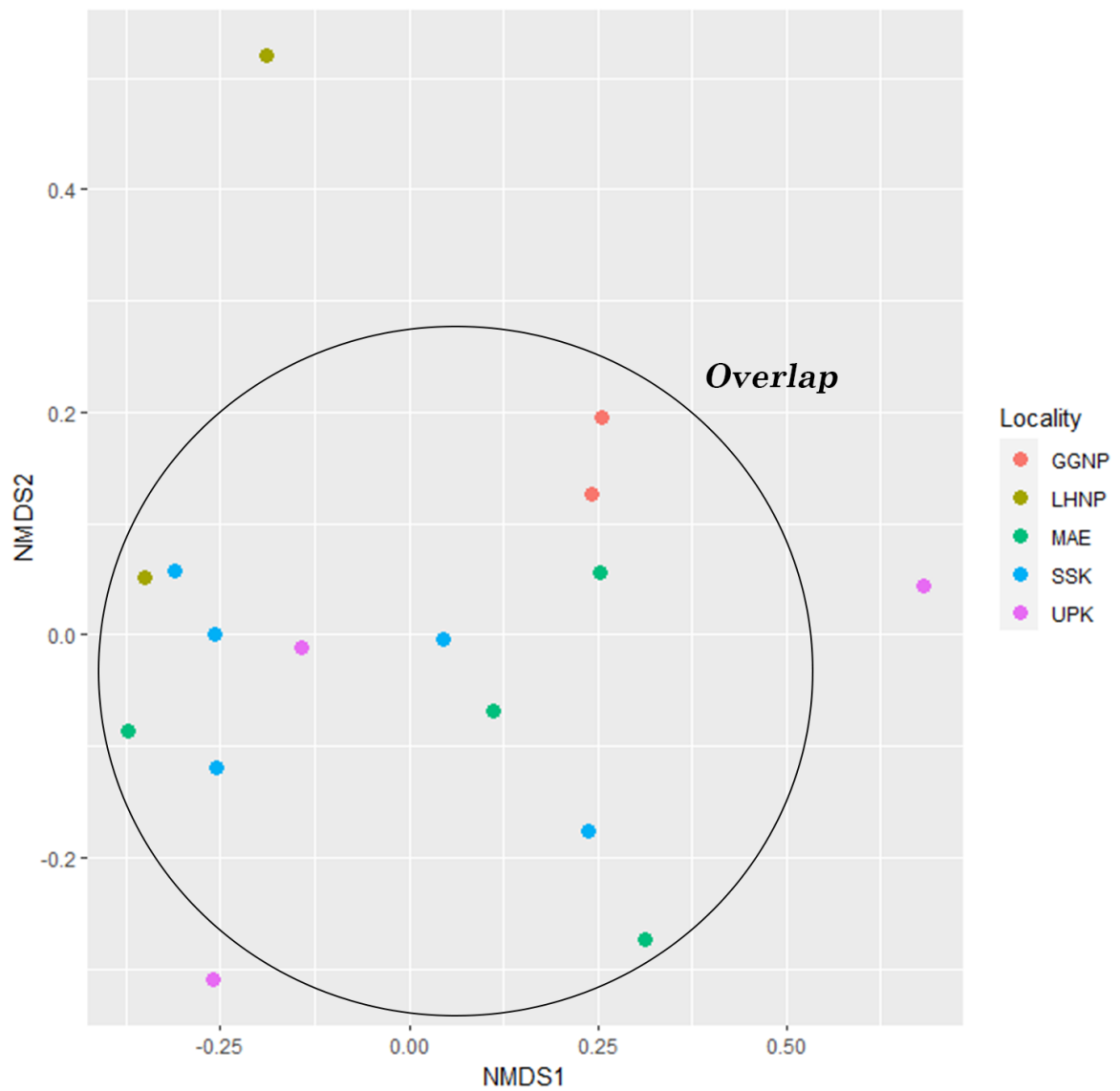


Figure 4.4: A Non-Metric Multidimensional Scaling (NMDS) Ordination Plot Using the Jaccard Index of all Rodent Individuals Colour Coded to Their Locality

The water quality assessment revealed that none of the sampled water bodies can be considered polluted based on the low obtained BOD values (Table 4.3). However, the water body in LHNP is noticeably more acidic as it has the lowest pH value (5.84). Additionally, water quality parameters from natural water bodies across all localities were correlated with the alpha diversity of rodent individuals found in those same areas. There were no significant correlations (Appendix F) observed between the Shannon diversity index and the tested water quality parameters (DO, BOD, TDS, salinity, conductivity, pH). Moreover, almost all tested parameters did not significantly correlate with pathogenic bacterial abundance, with the exception of BOD, which displayed a positive correlation with the genus *Enterococcus* (Figure 4.5).

Table 4.3: Water Quality Parameters of Natural Water Bodies of Sampling Sites in Sarawak. Correlation Coefficients Between Shannon Diversity Indices and Water Quality Parameters are Shown in the Last Row

Locality	Dissolved Oxygen (%)	Biochemical Oxygen Demand (mg/L)	Total Dissolved Solids (mg/L)	Salinity (ppt)	Conductivity (ms/cm)	pH
GGNP*	48.4	0.46	68.25	0.05	0.114	7.81
LHNP*	41.2	0.37	89.70	0.06	0.146	5.84
MAE*	46.7	0.90	42.25	0.03	0.069	7.11
SSK*	52.5	1.71	107.90	0.08	0.177	7.37
UPK*	71.7	0.02	21.45	0.01	0.032	6.50
Correlation Coefficient	0.58	0.25	-0.37	-0.28	-0.36	0.77

*Gunung Gading National Park (GGNP), *Lambir Hills National Park (LHNP), *Marup Atas Engkilili (MAE), *Sungai Sibau Kapit (SSK), *Sungai Sibau Kapit (SSK) and *Ulu Poi Kanowit (UPK)

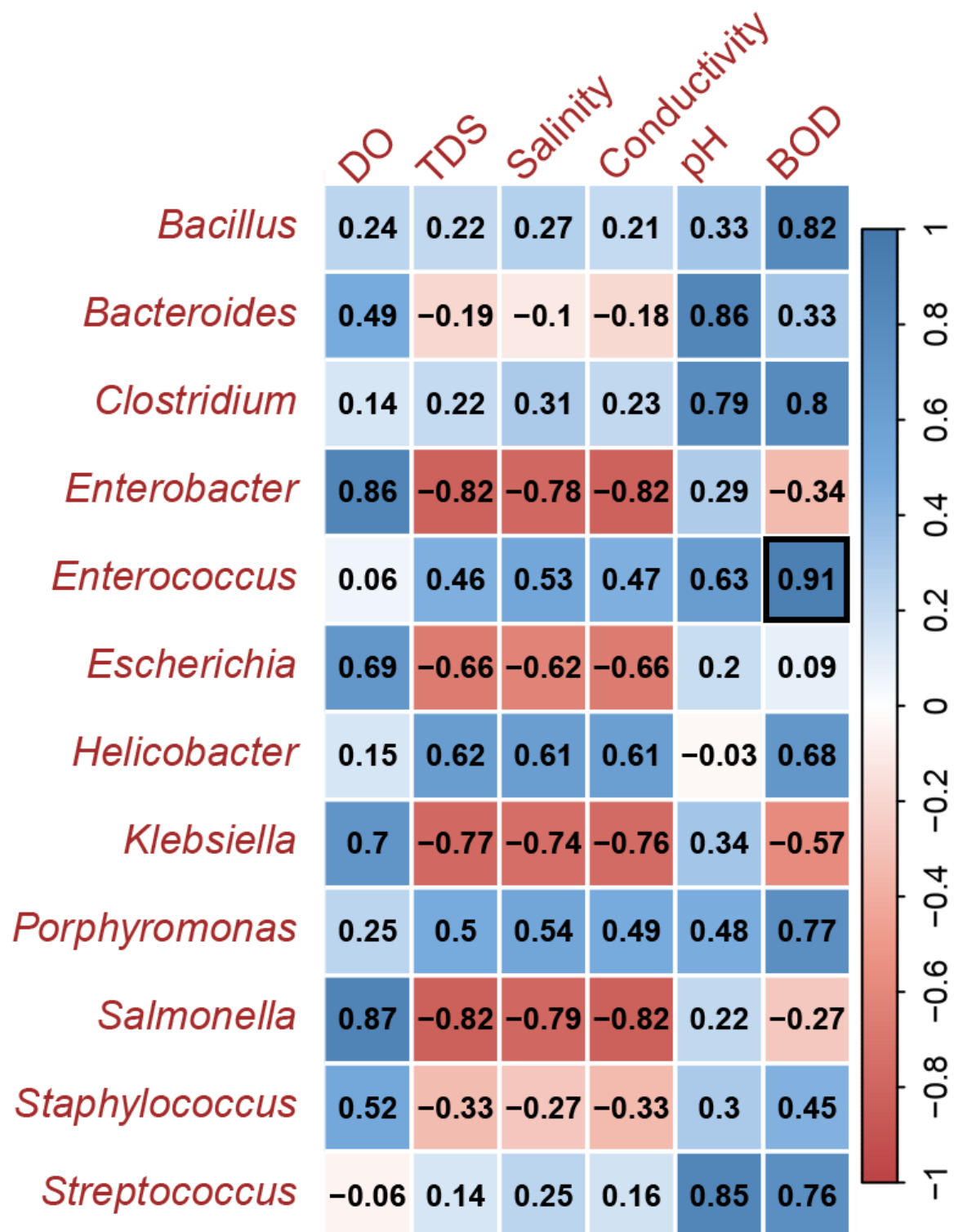


Figure 4.5: Correlogram Showing Correlations Between Water Quality Parameters and Abundance of Pathogenic Bacteria. Positive Correlations are Shown in Blue and Negative Correlations are Displayed in Red With the Intensity of the Colour Being Proportional to Correlation Strength. Values Marked With Black Outline are Significant at $p < 0.05$

4.4 Discussion

Based on this study, the gut microbiome composition of wild rodents is found to be diverse with potentially pathogenic bacterial species, congruent to another study by Jahan et al. (2021). Although faecal sampling may not capture the entire diversity of the gut microbiome, previous research by Ingala et al. (2018) demonstrated that faecal samples from wild mammals exhibit substantial species richness, indicating their relevance in providing insights into the overall microbiota composition. The variations in abundance patterns observed among bacterial genera across different localities indicate that the gut microbiota of rodents is influenced by the surrounding environment, similar to Niu et al. (2023) and Scholier et al. (2024). Given that the captured rodents were predominantly from forested areas, it is reasonable to hypothesise that their microbiomes reflect the presence of novel pathogenic bacteria that naturally occur in their respective localities.

Among the identified bacterial genera, *Clostridium* was found to be one of the most abundant taxa across all study sites. Notably, the pathogenic species *Clostridium perfringens* was detected in every rodent individual, and it is known to be prevalent in Malaysian rivers (Lee et al., 2012). This is consistent with our findings as the rodents were captured near natural water bodies. Previous studies have indicated that pathogenic bacteria can originate from natural water sources and disperse to the surrounding wildlife (Lesley et al., 2016; Raschle et al., 2021). Although *C. perfringens* is primarily associated with food contamination and bowel infections, infected rodent populations could potentially serve as reservoirs, posing a risk of contact with nearby human settlements.

Furthermore, the genus *Klebsiella* was abundantly found in Gunung Gading National Park, Marup Atas Engkilili and Ulu Poi Kanowit. *Klebsiella* species are recognised as

opportunistic pathogens that primarily target individuals with compromised immune systems, causing infections such as pneumonia (Podschun & Ullmann, 1998). Among the various *Klebsiella* species detected in the rodent individuals, *Klebsiella pneumoniae* emerged as the most prevalent in this study. In Sarawak, this pathogenic bacterial species was documented in natural water bodies utilised for recreational purposes (Lihan et al., 2017). Another notable species found was *Klebsiella variicola*, a bacterial species that has a wide range of distribution in a natural environment, which has raised concerns in public health due to its potential to acquire virulence and antimicrobial genes (Rodríguez-Medina et al., 2019). The presence of opportunistic pathogens in wild rodent near human settlements can pose a significant public health risk, considering their inherit ability to develop antimicrobial resistance.

Another noteworthy genus is *Helicobacter*, which is one of the predominant bacteria observed in rodents from Lambir Hills National Park and Sungai Sibau Kapit. While certain members within this genus, like *Helicobacter pylori*, were detected and have a notable history of infection in Malaysia (Tay et al., 2009), the seroprevalence of *H. pylori* has been globally declining, a trend that is also evident in Borneo (Mori et al., 2022). *H. pylori* infections can cause chronic stomach inflammation and may lead to cancer development (Sepulveda, 2013). However, the most abundant *Helicobacter* species found in all rodent individuals in this study was *Helicobacter apodemus*. This potentially pathogenic species was described in the past decade and is known to inhabit the gastrointestinal tract of rodents (Jeon et al., 2015). Despite limited knowledge regarding its pathogenesis, the higher relative abundance of *H. apodemus* compared to the other more well-known species, *H. pylori*, indicates that it likely plays a significant role in the rodent gut microbiota. The transmission of bacteria that are relatively unknown in their pathogenicity but belongs to a genus with

known pathogenic members can lead to emerging infections without established control measures (Nii-Trebi, 2017).

Our findings revealed an overlap in distribution of numerous bacterial species across all localities, suggesting their potential as core microbiota in wild rodents and their crucial role in specific niches within the gut microbiome. Some of the commonly shared species includes *Escherichia coli*, *Klebsiella pneumoniae* and *Staphylococcus aureus*. Typically, the core gut microbiota of rodents consists of bacteria that are involved in metabolism of nutrients and immune system regulation (Wang et al., 2019). The pathogenic bacterial species could have formed a commensal relationship with the core gut microbiota, shaping a complex microbial community. Understanding the dynamics between these relationships could be essential for comprehending the adaptations of certain pathogenic bacterial species and the susceptibility of hosts to infections (Bäumler & Sperandio, 2016).

A diverse range of bacterial species also exhibited site-specific distribution, with Ulu Poi Kanowit and Sungai Sibau Kapit standing out as the study sites harbouring the highest number of bacterial species that are exclusive to that specific area. In contrast, Gunung Gading National Park and Lambir Hills National Park displayed relatively lower numbers of exclusive species. One possible explanation for this discrepancy is that both Gunung Gading National Park and Lambir Hills National Park are designated as protected areas, where environmental conditions may differ from other sites, potentially limiting the diversity of bacterial species. There could be differences in the rodents' diet, habitat type and social structure, all of which could cause variation in the gut microbiome (Amato et al., 2019; Viney, 2019; Raulo et al., 2021). Moreover, the lower number of captured rodents in these areas might have contributed to the reduced species count. However, it is noteworthy that despite the overall lower species diversity, Gunung Gading National Park displayed a

notably higher abundance of pathogenic bacteria compared to Lambir Hills National Park. The difference in pathogenic bacterial load between the two sites may be because one of the captured rodents in Lambir Hills National Park was juvenile. The relative abundance of pathogenic bacteria in its gut microbiome was the lowest (4.5%) out of all rodent individuals. The gut microbiome may not have fully developed to support a diverse range of pathogenic bacteria as ageing in wild mammals is associated with an increase in gut bacteria richness (Fenn et al., 2023).

Several water quality parameters that could potentially influence bacterial abundance were assessed for this study. This includes the parameters pH, dissolved oxygen and conductivity which have been identified as predominant factors that could shape the microbial communities in water bodies (Staley et al., 2018; Wang et al., 2018; Li et al., 2021). However, investigations into water quality parameters did not establish significant correlations with the diversity of the rodents' gut microbiota. While the assessed water quality parameters are known to influence microbial communities in water (Staley et al., 2018; Wang et al., 2018; Li et al., 2021), insufficient sample size could have led to a lack of significant correlations. Additionally, due to the dynamic nature of water quality parameters in flowing rivers (Edwards & Thornes, 1973; Vega et al., 1998; Ouyang et al., 2006; Fiedler et al., 2018), their variability may not be suitable for correlations with the gut microbiome diversity of rodents. These fluctuations can diverge significantly from stable characteristics found in stagnant water systems.

Additionally, most of the tested parameters did not significantly correlate with the abundance of pathogenic bacteria. While the gut microbiome of rodent communities may vary across different localities, not much is known about the environmental factors that could have influenced the diversity of pathogenic bacteria. Although it is possible to hypothesise

that the bacterial communities present in the water surface of surrounding natural water bodies could shape the gut microbiota of rodents, *Enterococcus* was the only genus that had significant correlation with water quality. The abundance of the genus was seen to have positively correlated with biochemical oxygen demand. This finding is reasonable as *Enterococcus* is frequently detected in contaminated water surfaces and are used as an indicator of faecal contamination in freshwater (Cho et al., 2020). There were no significant correlations between the other bacterial genera and any of the water quality parameters. Similarly, this could be due to all rodent samplings were conducted near rivers, where the continuous flowing water might had prevented the water surface from maintaining a stable bacterial community. Therefore, it is possible that other stable environmental factors, such as the vegetation could have more of an influence on the gut microbiome of the rodent populations in forested areas (Fan et al., 2022).

While the use of next-generation sequencing has allowed the characterisation of the gut microbiome of rodent populations, some challenges were faced due to the disparity in library sizes between samples. Due to the nature of high throughput sequencing, the number of samples and sequencing depth may differ across samples, causing uneven datasets. To address this issue, the rarefying method was used to eliminate bias associated with differences in sample size, ensuring uniformity across all samples. In this study, the dataset was subsampled to a common threshold, until each sample reached the minimum number of reads. Although this may have resulted in the underestimation of rare taxa and affected the overall diversity, a large, normalised library size was retained to minimise the introduction of artificial variation (Cameron et al., 2021). Artificial variation can occur when a significant portion of the dataset is omitted during subsampling, a scenario often observed in cases of excessive normalisation resulting in small library sizes (McMurdie & Holmes, 2014).

Additionally, delving into other environmental factors could reveal associations with the pathogenic bacteria in the wild rodent gut microbiome. Recent studies have shown that feeding behaviours and the surrounding vegetation structure can influence the gut microbiota of mammals as their dietary preferences can promote susceptibility to pathogens (Fan et al., 2022; Kohl et al., 2022; Neumann et al., 2022). Lastly, the characterisation of the juvenile rodent's gut microbiota highlights the importance of understanding the prevalence of pathogenic bacteria at different stages of their life cycles. Further research is needed to explore the relationship between different reproductive stages of rodents and their microbiome communities.

4.5 Conclusion

This study successfully highlights the presence of potentially pathogenic bacterial genera in the gut microbiota of wild rodents from forested areas (~11%), with variations in abundance and distribution patterns across different localities. A total of 774 potentially pathogenic species belonging to 51 bacterial genera were detected. Potential pathogenic species such as *Clostridium perfringens*, *Klebsiella pneumoniae* and *Helicobacter pylori* can cause zoonotic diseases and pose as public health risks to the nearby human settlements. Also, an overlap in bacterial species was observed when comparing the gut microbiome of rodents from different study sites. The occurrence of shared and site-specific bacterial species highlights the potential relationships between core microbiota and their environmental influences. Further research is needed to understand the dynamics between pathogenic bacteria and the core gut microbiota, as well as the environmental factors shaping the diversity of these bacteria in rodent populations.

4.6 Recommendation

Utilising environmental DNA sequencing would greatly improve this study as it presents an opportunity to gain understanding regarding the existence of pathogenic bacteria within the environment. This approach facilitates the identification of potential bacterial circulation between rodent populations and their surroundings. These findings would strengthen the assumption that environmental factors directly impact the presence of potentially pathogenic bacteria in wild rodents. Also, antimicrobial resistance detection methods can significantly enhance the relevance of studies on pathogenic bacteria. These techniques enable the identification of specific genes or mutations responsible for resistance mechanisms.

CHAPTER 5

GENERAL DISCUSSION

5.1 Discussion

5.1.1 Gut Microbiome Composition

The rodents' gut microbiome exhibits variation, indicating distinct bacterial communities. Factors such as dietary preferences, host genetics, and environmental exposures contribute to this heterogeneity. Phylum Bacillota dominates these microbiomes, contributing to the regulation of the host's immune system and the development of the gut epithelium. Oscillospiraceae, also known as Ruminococcaceae, is one of the most abundant bacterial families in the gut microbiota of rodents, playing a vital role in digestion and nutrient absorption. The genus *Ruminococcus* is a potential constituent of the rodent core gut bacteria due to its prevalence across all rodent individuals. Furthermore, Lachnospiraceae is a prominent bacterial family with multiple genera at varying abundances, and their presence within the gut microbiota suggests their pivotal role as a component of the bacterial community. Another notable family is Lactobacillaceae, which represents the lactic acid bacteria responsible for metabolising lactose within the gut environment. The prevalence of *Lactobacillus* in the gut of young rodents may indicate its significance during infancy, potentially contributing to the digestion of maternal milk.

Several pathogenic bacteria were detected at low abundance in multiple rodent individuals from various localities as they represent around 11% of the collective gut microbiome. This is comparable to the findings of a survey by Jahan et al. (2021), where the relative abundance of pathogenic bacteria was approximately 15% of overall rodent gut microbiota. While their combined abundance is lower than that of the presumed core

microbiome constituents, their prevalence in all wild rodent individuals indicates the potential existence of a commensal relationship between them. However, the interactions between specific environmental factors and the composition of the wild rodent gut microbiome remain largely unknown in this study. The gut microbiota of rodents could potentially be specific to their hosts due to the potential impact of phylogenetic similarity among rodent species.

5.1.2 Prevalence of Pathogenic Bacteria

The variations in abundance patterns observed among pathogenic bacterial genera across different localities indicate that the gut microbiota of rodents is influenced by the surrounding environment. Among the identified bacterial genera, *Clostridium* was found to be one of the most abundant taxa across all study sites. Previous studies have indicated that pathogenic bacteria can originate from natural water sources and disperse to the surrounding wildlife. Infected rodent populations could potentially serve as reservoirs, posing a risk of contact with nearby human settlements. Also, the genus *Klebsiella* was abundantly found in Gunung Gading National Park, Marup Atas Engkilili, and Ulu Poi Kanowit. *Klebsiella* species are recognised as opportunistic pathogens that primarily target individuals with compromised immune systems. Another noteworthy genus is *Helicobacter*, which is one of the predominant bacteria observed in rodents from Lambir Hills National Park and Sungai Sibau Kapit. One of the most abundant *Helicobacter* species was found to be *Helicobacter apodemus*, a potentially pathogenic species known to inhabit the gastrointestinal tract of rodents.

The findings revealed an overlap in distribution of numerous bacterial species across all localities, suggesting their potential as core microbiota in wild rodents and their crucial

role in specific niches within the gut microbiome. Conversely, a diverse range of pathogenic bacteria exhibited site-exclusive distribution, possibly due to differences in rodents' diet, habitat type, and social structure. While the gut microbiome of wild rodents may vary across different localities, not much is known about the environmental factors that could have influenced the diversity of pathogenic bacteria.

Several water quality parameters were assessed for this study but most of the tested parameters did not significantly correlate with the gut microbial diversity and abundance of pathogenic bacteria. The instability of bacterial communities in naturally occurring water bodies could have hindered direct transmission to nearby rodent populations. Factors such as rainfall or variations in river flow could have potentially caused significant impact on the stability of bacterial communities. *Enterococcus* was the only genus that had a significant correlation with water quality, as the abundance of the genus was seen to have positively correlated with biochemical oxygen demand. Shifting focus towards dietary behaviours and vegetation composition of the natural habitat would help determine the environmental factors that contribute to variations in the gut microbial communities within wild rodents.

CHAPTER 6

CONCLUSION AND RECOMMENDATION

6.1 Conclusion

In this study, the gut microbiome compositions of numerous wild rodents are successfully characterised in this study. A total of 1052 bacterial genera were detected using nanopore sequencing, most of which represented the phylum Bacillota. Several bacterial genera could potentially be considered integral members of the core gut microbiome for wild rodents, given their prevalence across the gut microbiota of multiple rodent individuals. Phylogenetic analyses revealed that individuals that were closely related show patterns of similar gut microbiome composition. However, the alpha diversity for the gut microbiome within each rodent individual is similar, apart from a juvenile individual showing lower diversity. The gut microbial diversity also did not correlate with any of the tested water quality parameters, implying that the natural water bodies situated near the wild rodents did not significantly influence their gut microbiome composition.

Pathogenic bacteria have also been detected in rodent individuals across all localities, exhibiting varying levels of abundance. These microbes represent approximately 11% of the total number of sequenced bacteria, comprising 774 distinct bacterial species from 51 potentially pathogenic genera. Among these, prominent genera such as *Clostridium*, *Klebsiella* and *Helicobacter* display notably high abundances. Generally, the pathogenic bacteria abundance in wild rodents did not correlate with the quality of nearby natural water bodies, although *Enterococcus* did significantly correlate positively with biochemical oxygen demand. Each sampling site was situated near or adjacent to human settlements, thus highlighting the potential risk of zoonotic disease transmission.

6.2 Recommendation

A more robust sampling design and increased sampling efforts could significantly improve this study. The characterisation of the rodent gut microbiota does not accurately represent specific rodent species due to the low number of samples for each species. Having more representatives for each rodent species, including both juveniles and adults, would result in much more comprehensive research that effectively describes the gut microbiome of wild rodents. Furthermore, collecting samples throughout the year could detect short-term shifts in the gut microbiota composition caused by temporal changes. These recommendations are plausible due to the non-invasive nature of faecal sampling, allowing a higher number of collected samples without affecting the wild rodent populations themselves.

Additionally, delving into other environmental factors could reveal associations with the wild rodent gut microbiome. Some potential constituents of the core gut microbiota are responsible for the hosts' physicochemical processes and could be influenced by the diet preferences of the rodents. Therefore, exploring their feeding behaviours and the vegetation structure of their native habitat could provide more information on the relationships between the gut microbiome and the surrounding environment. Taking advantage of environmental DNA sequencing could also yield insights into these interactions. This method would enable the detection of the potential circulation of bacteria between the rodent populations and the environment. This would further strengthen the assumptions that environmental pressure would directly influence the gut microbiome composition of wild rodents.

REFERENCES

- Amato, K. R., Martinez-Mota, R., Righini, N., Raguet-Schofield, M., Corcione, F. P., Marini, E. E., Humphrey, G., Gogul, G., Gaffney, J., Lovelace, E., Williams, L., Luong, A., Dominguez-Bello, M. G., Stumpf, R., White, B., Nelson, K., Knight, R., & Leigh, S. (2019). Phylogenetic and ecological factors impact the gut microbiota of two Neotropical primate species. *Journal of Animal Ecology*, 88(3), 347-358.
- Anders, J. L., Moustafa, M. A. M., Mohamed, W. M. A., Hayakawa, T., Nakao, R., & Koizumi, I. (2021). Comparing the gut microbiome along the gastrointestinal tract of three sympatric species of wild rodents. *Scientific Reports*, 11(1), 19929.
- Archie, E. A., & Tung, J. (2015). Social behavior and the microbiome. *Current Opinion in Behavioral Sciences*, 6, 28-34.
- Ari, Ş., & Arikian, M. (2016). Next-generation sequencing: Advantages, disadvantages, and future. In *Plant Omics: Trends and Applications*, pp. 109-135. Cham, Switzerland: Springer.
- Arumugam, R., Ravichandran, P., Yeap, S. K., Sharma, R. S. K., Zulkifly, S. B., Yawah, D., & Annavi, G. (2023). Application of high-throughput sequencing (HTS) to enhance the well-being of an endangered species (Malayan Tapir): Characterization of gut microbiome using MG-RAST. In *Metagenomic Data Analysis*, pp. 175-194. New York, NY: Springer US.
- Barcina, I., Lebaron, P., & Vives-Rego, J. (1997). Survival of allochthonous bacteria in aquatic systems: A biological approach. *FEMS Microbiology Ecology*, 23(1), 1-9.

- Bäumler, A. J., & Sperandio, V. (2016). Interactions between the microbiota and pathogenic bacteria in the gut. *Nature*, 535(7610), 85-93.
- Benacer, D., Thong, K. L., Verasahib, K. B., Galloway, R. L., Hartskeerl, R. A., Lewis, J. W., & Zain, S. N. M. (2016). Human leptospirosis in Malaysia: Reviewing the challenges after 8 decades (1925-2012). *Asia Pacific Journal of Public Health*, 28(4), 290-302.
- Bickham, J. W., Wood, C. C., & Patton, J. C. (1995). Biogeographic implications of cytochrome *b* sequences and allozymes in sockeye (*Oncorhynchus nerka*). *Journal of Heredity*, 86(2), 140-144.
- Biddle, A., Stewart, L., Blanchard, J., & Leschine, S. (2013). Untangling the genetic basis of fibrolytic specialization by Lachnospiraceae and Ruminococcaceae in diverse gut communities. *Diversity*, 5(3), 627-640.
- Bilung, L. M., Ling, K. K., Apun, K., Abdullah, M. T., Rahman, M. A., Ming, C. Y., Benjamin, A., Chong, Y. L., & Hashim, H. F. (2014). Occurrence of *Escherichia coli* in wildlife from different habitats of Sarawak, Malaysia. *Borneo Journal of Resource Science and Technology*, 4(1), 19-27.
- Blasdell, K. R., Morand, S., Perera, D., & Firth, C. (2019a). Association of rodent-borne *Leptospira* spp. with urban environments in Malaysian Borneo. *PLoS Neglected Tropical Diseases*, 13(2), e0007141.
- Blasdell, K. R., Perera, D., & Firth, C. (2019b). High prevalence of rodent-borne *Bartonella* spp. in urbanizing environments in Sarawak, Malaysian Borneo. *The American Journal of Tropical Medicine and Hygiene*, 100(3), 506.

- Bordes, F., Blasdell, K., & Morand, S. (2015). Transmission ecology of rodent-borne diseases: New frontiers. *Integrative Zoology*, 10(5), 424-435.
- Borges, A. S. G., Basu, M., Brinks, E., Bang, C., Cho, G. S., Baines, J. F., Franke, A., & Franz, C. M. A. P. (2023). Fast identification method for screening bacteria from faecal samples using Oxford Nanopore Technologies MinION Sequencing. *Current Microbiology*, 80(3), 101.
- Cameron, E. S., Schmidt, P. J., Tremblay, B. J. M., Emelko, M. B., & Müller, K. M. (2021). Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. *Scientific Reports*, 11(1), 22302.
- Campbell, J. H., Foster, C. M., Vishnivetskaya, T., Campbell, A. G., Yang, Z. K., Wymore, A., Palumbo, A. V., Chesler, E. J., & Podar, M. (2012). Host genetic and environmental effects on mouse intestinal microbiota. *The ISME Journal*, 6(11), 2033-2044.
- Chen, F., Dong, M., Ge, M., Zhu, L., Ren, L., Liu, G., & Mu, R. (2013). The history and advances of reversible terminators used in new generations of sequencing technology. *Genomics, Proteomics, & Bioinformatics*, 11(1), 34–40.
- Chen, P., Sun, Z., Wang, J., Liu, X., Bai, Y., Chen, J., Liu, A., Qiao, F., Chen, Y., Yuan, C., Sha, J., Zhang, J., Xu, L., & Li, J. (2023). Portable nanopore-sequencing technology: Trends in development and applications. *Frontiers in Microbiology*, 14, 1043967.
- Cho, S., Jackson, C. R., & Frye, J. G. (2020). The prevalence and antimicrobial resistance phenotypes of *Salmonella*, *Escherichia coli* and *Enterococcus* sp. in surface water. *Letters in Applied Microbiology*, 71(1), 3-25.

- Ciuffreda, L., Rodríguez-Pérez, H., & Flores, C. (2021). Nanopore sequencing and its application to the study of microbial communities. *Computational and Structural Biotechnology Journal*, 19, 1497-1511.
- 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.
- Clavel, T., Lagkouvardos, I., Blaut, M., & Stecher, B. (2016). The mouse gut microbiome revisited: From complex diversity to model ecosystems. *International Journal of Medical Microbiology*, 306(5), 316-327.
- Cosson, J. F., Picardeau, M., Mielcarek, M., Tatard, C., Chaval, Y., Suputtamongkol, Y., Buchy, P., Jittapalapong, S., Herbreteau, V., & Morand, S. (2014). Epidemiology of *Leptospira* transmitted by rodents in Southeast Asia. *PLoS Neglected Tropical Diseases*, 8(6), e2902.
- Cotter, P. D., & Hill, C. (2003). Surviving the acid test: Responses of gram-positive bacteria to low pH. *Microbiology and Molecular Biology Reviews*, 67(3), 429-453.
- Coyte, K. Z., & Rakoff-Nahoum, S. (2019). Understanding competition and cooperation within the mammalian gut microbiome. *Current Biology*, 29(11), R538-R544.
- Dahmana, H., Granjon, L., Diagne, C., Davoust, B., Fenollar, F., & Mediannikov, O. (2020). Rodents as hosts of pathogens and related zoonotic disease risk. *Pathogens*, 9(3), 202.
- Edwards, A. M. C., & Thornes, J. B. (1973). Annual cycle in river water quality: A time series approach. *Water Resources Research*, 9(5), 1286-1295.

- El Kaoutari, A., Armougom, F., Gordon, J. I., Raoult, D., & Henrissat, B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nature Reviews Microbiology*, *11*, 497–504.
- Euzéby, J. (2010). List of new names and new combinations previously effectively, but not validly, published. *International Journal of Systematic and Evolutionary Microbiology*, *60*(3), 469–472.
- Fan, C., Zhang, L., Jia, S., Tang, X., Fu, H., Li, W., Liu, C., Zhang, H., Cheng, Q., & Zhang, Y. (2022). Seasonal variations in the composition and functional profiles of gut microbiota reflect dietary changes in plateau pikas. *Integrative Zoology*, *17*(3), 379–395.
- Fenn, J., Taylor, C., Goertz, S., Wanelik, K. M., Paterson, S., Begon, M., Jackson, J., & Bradley, J. (2023). Discrete patterns of microbiome variability across timescales in a wild rodent population. *BMC Microbiology*, *23*(1), 1–17.
- Fiedler, C. J., Schönher, C., Proksch, P., Kerschbaumer, D. J., Mayr, E., Zunabovic-Pichler, M., Domig, K. J., & Perfler, R. (2018). Assessment of microbial community dynamics in river bank filtrate using high-throughput sequencing and flow cytometry. *Frontiers in Microbiology*, *9*, 2887.
- Furman, O., Shenhav, L., Sasson, G., Kokou, F., Honig, H., Jacoby, S., Hertz, T., Cordero, O. X., Halperin, E., & Mizrahi, I. (2020). Stochasticity constrained by deterministic effects of diet and age drive rumen microbiome assembly dynamics. *Nature Communications*, *11*(1), 1904.

- Greene, L. K., Williams, C. V., Junge, R. E., Mahefarisoa, K. L., Rajaonarivelo, T., Rakotondrainibe, H., O'Connell, T. M., & Drea, C. M. (2020). A role for gut microbiota in host niche differentiation. *The ISME Journal*, 14(7), 1675-1687.
- Gu, S., Chen, D., Zhang, J. N., Lv, X., Wang, K., Duan, L. P., Nie, Y., & Wu, X. L. (2013). Bacterial community mapping of the mouse gastrointestinal tract. *PloS ONE*, 8(10), e74957.
- Hart, J. S. (1971). Rodents. In *Comparative Physiology of Thermoregulation, Volume II Mammals* (Whittow, G. C.), pp. 1-149. New York, NY: Academic Press.
- Hayakawa, T., Nathan, S. K., Stark, D. J., Saldivar, D. A. R., Sipangkui, R., Goossens, B., Tuuga, A., Clauss, M., Sawada, A., Fukuda, S., Imai, H., & Matsuda, I. (2018). First report of foregut microbial community in proboscis monkeys: Are diverse forests a reservoir for diverse microbiomes? *Environmental Microbiology Reports*, 10(6), 655-662.
- Heuer, H., Krsek, M., Baker, P., Smalla, K., & Wellington, E. (1997). Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. *Applied and Environmental Microbiology*, 63(8), 3233-3241.
- Hou, K., Wu, Z. X., Chen, X. Y., Wang, J. Q., Zhang, D., Xiao, C., Zhu, D., Koya, J. B., Wei, L., Li, J., & Chen, Z. S. (2022). Microbiota in health and diseases. *Signal Transduction and Targeted Therapy*, 7(1), 135.
- Ingala, M. R., Simmons, N. B., Wultsch, C., Krampis, K., Speer, K. A., & Perkins, S. L. (2018). Comparing microbiome sampling methods in a wild mammal: Fecal and

- intestinal samples record different signals of host ecology, evolution. *Frontiers in Microbiology*, 9, 803.
- Jahan, N. A., Lindsey, L. L., Kipp, E. J., Reinschmidt, A., Heins, B. J., Runck, A. M., & Larsen, P. A. (2021). Nanopore-based surveillance of zoonotic bacterial pathogens in farm dwelling peridomestic rodents. *Pathogens*, 10(9), 1183.
- Janda, J. M., & Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. *Journal of Clinical Microbiology*, 45(9), 2761-2764.
- Jeon, W. J., Dong, H. J., Shin, J. H., Kim, I. Y., Ho, H., Oh, S. H., Yoon, Y. M., Choi, Y. K., Suh, J. G., Nam, K. H., & Seong, J. K. (2015). *Helicobacter apodemus* sp. nov., a new *Helicobacter* species identified from the gastrointestinal tract of striped field mice in Korea. *Journal of Veterinary Science*, 16(4), 475-481.
- Jouanneau, S., Recoules, L., Durand, M. J., Boukabache, A., Picot, V., Primault, Y., Lakel, A., Sengelin, M., Barillon, B., & Thouand, G. (2014). Methods for assessing biochemical oxygen demand (BOD): A review. *Water Research*, 49, 62-82.
- Kerfeld, C. A., & Scott, K. M. (2011). Using BLAST to teach “E-value-tionary” concepts. *PLoS Biology*, 9(2), e1001014.
- Kim, S., & Jazwinski, S. M., (2018). The gut microbiota and healthy aging: A mini-review. *Gerontology*, 65(6), 513-520.
- Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111-120.

- Kinross, J. M., Darzi, A. W., & Nicholson, J. K. (2011). Gut microbiome-host interactions in health and disease. *Genome Medicine*, 3, 1-12.
- Kohl, K. D., Dieppa-Colón, E., Goyco-Blas, J., Peralta-Martínez, K., Scafidi, L., Shah, S., Zawacki, E., Barts, N., Ahn, Y., Hedayati, S., Secor, S. M., & Rowe, M. P. (2022). Gut microbial ecology of five species of sympatric desert rodents in relation to herbivorous and insectivorous feeding strategies. *Integrative and Comparative Biology*, 62(2), 237-251.
- König, H., & Fröhlich, J. (2017). Lactic acid bacteria. In *Biology of Microorganisms on Grape, in Must and in Wine* (Konig, H., Uden, G., & Frohlich, J.), pp. 3-41. Cham, Switzerland: Springer.
- Krügel, M., Król, N., Kempf, V. A., Pfeffer, M., & Obiegala, A. (2022). Emerging rodent associated *Bartonella*: A threat for human health? *Parasites & Vectors*, 15(1), 113.
- Kruse, C. P., Basu, P., Luesse, D. R., & Wyatt, S. E. (2017). Transcriptome and proteome responses in RNAlater preserved tissue of *Arabidopsis thaliana*. *PloS ONE*, 12(4), e0175943.
- Kubasova, T., Seidlerova, Z., & Rychlik, I. (2021). Ecological adaptations of gut microbiota members and their consequences for use as a new generation of probiotics. *Frontiers in Microbiology*, 11, 553.
- Kucheruk, V. V. (1965). Synanthropic rodents and their significance in the transmission of infections. In *Theoretical Questions of Natural Foci of Diseases* (Rosicky, B. & Heyberger, K.), pp. 353-366. Prague: Czechoslovak Academy of Sciences.

- Lau, A. C., Qiu, Y., Moustafa, M. A. M., Nakao, R., Shimozuru, M., Onuma, M., Mohd Azlan, J., & Tsubota, T. (2020). Detection of *Borrelia burgdorferi* Ssensu Lato and relapsing fever *Borrelia* in feeding *Ixodes* ticks and rodents in Sarawak, Malaysia: New geographical records of *Borrelia yangtzensis* and *Borrelia miyamotoi*. *Pathogens*, 9(10), 846.
- Lee, F. C. H., Hakim, S. L., Kamaluddin, M. A., & Thong, K. L. (2012). *Clostridium perfringens* and sulphite reducing clostridia densities in selected tropical Malaysian rivers. *The Southeast Asian Journal of Tropical Medicine and Public Health*, 43(1), 129-135.
- Leser, T. D., & Mølbak, L. (2009). Better living through microbial action: The benefits of the mammalian gastrointestinal microbiota on the host. *Environmental Microbiology*, 11(9), 2194-2206.
- Lesley, M. B., Velnetti, L., Fazira, A. A., Kasing, A., Samuel, L., Micky, V., & Awang, A. S. A. H. (2016). Detection and antibiotic susceptibility profiles of *Listeria monocytogenes* in wildlife and water samples in Kubah National Park, Sarawak, Malaysia. *International Food Research Journal*, 23(1), 360.
- Li, K., Hu, J., Li, T., Liu, F., Tao, J., Liu, J., Zhang, Z., Luo, X., Li, L., Deng, Y., & Che, R. (2021). Microbial abundance and diversity investigations along rivers: Current knowledge and future directions. *Wiley Interdisciplinary Reviews: Water*, 8(5), e1547.
- Lihan, S., Tian, P. K., Chiew, T. S., Ching, C. L., Shahbudin, A., Hussain, H., & Mohd-Azlan, J. (2017). The distribution and characteristics of bacteria in recreational river water

- of a community resort in Baram, Sarawak, Malaysian Borneo. *International Food Research Journal*, 24(5), 2238-2245.
- Lindström, E. S., Kamst-Van Agterveld, M. P., & Zwart, G. (2005). Distribution of typical freshwater bacterial groups is associated with pH, temperature, and lake water retention time. *Applied and Environmental Microbiology*, 71(12), 8201-8206.
- Liu, X., Mao, B., Gu, J., Wu, J., Cui, S., Wang, G., Zhao, J., Zhang, J., & Chen, W. (2021). *Blautia*—a new functional genus with potential probiotic properties? *Gut Microbes*, 13(1), 1875796.
- Lobato-Bailón, L., García-Ulloa, M., Santos, A., Guixé, D., Camprodon, J., Florensa-Rius, X., Molleda, R., Manzano, R., Ribas, M. P., Espunyes, J., Dias-Alves, A., Marco, I., Migura-Garcia, L., Martínez-Urtaza, J., & Cabezón, O. (2023). The fecal bacterial microbiome of the Kuhl's pipistrelle bat (*Pipistrellus kuhlii*) reflects landscape anthropogenic pressure. *Animal Microbiome*, 5(1), 1-14.
- Maurice, C. F., Knowles, S. C., Ladau, J., Pollard, K. S., Fenton, A., Pedersen, A. B., & Turnbaugh, P. J. (2015). Marked seasonal variation in the wild mouse gut microbiota. *The ISME Journal*, 9(11), 2423-2434.
- McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*, 8(4), e61217.
- McMurdie, P. J., & Holmes, S. (2014). Waste not, want not: Why rarefying microbiome data is inadmissible. *PLoS Computational Biology*, 10(4), e1003531.
- Meerburg, B. G., Singleton, G. R., & Kijlstra, A. (2009). Rodent-borne diseases and their risks for public health. *Critical Reviews in Microbiology*, 35(3), 221-270.

- Mills, J. N. (2005). Regulation of rodent-borne viruses in the natural host: Implications for human disease. *Arch Virol Suppl*, (19), 45-57.
- Mohd-Yusof, N. S., Abdul-Latiff, M. A. B., Badrulisham, A. S., Othman, N., Yaakop, S., Md-Nor, S., & Md-Zain, B. M. (2022). First report on metabarcoding analysis of gut microbiome in Island Flying Fox (*Pteropus hypomelanus*) in island populations of Malaysia. *Biodiversity Data Journal*, 10, e69631.
- Morey, M., Fernández-Marmiesse, A., Castiñeiras, D., Fraga, J. M., Couce, M. L., & Cocho, J. A. (2013). A glimpse into past, present, and future DNA sequencing. *Molecular Genetics and Metabolism*, 110(1-2), 3-24.
- Mori, D., John, J. L., Sabri, S. I. B., Shaharom, S. M. B., Iha, H., Yamaoka, Y., Matsumoto, T., & Ahmed, K. (2022). Seroepidemiological survey of the prevalence of *Helicobacter pylori* infection in Sabah, Malaysia. *IJID Regions*, 2, 126-129.
- Mwanamoki, P. M., Devarajan, N., Thevenon, F., Atibu, E. K., Tshibanda, J. B., Ngelinkoto, P., Mpiana, P. T., Prabakar, K., Mubedi, J. I., Kabele, C. G., Wildi, W., & Poté, J. (2014). Assessment of pathogenic bacteria in water and sediment from a water reservoir under tropical conditions (Lake Ma Vallée), Kinshasa Democratic Republic of Congo. *Environmental Monitoring and Assessment*, 186(10), 6821-6830.
- Neumann, M., Steimle, A., Grant, E. T., Wolter, M., Parrish, A., Willieme, S., Brenner, D., Martens, E. C., & Desai, M. S. (2021). Deprivation of dietary fiber in specific pathogen-free mice promotes susceptibility to the intestinal mucosal pathogen *Citrobacter rodentium*. *Gut Microbes*, 13(1), 1966263.
- Nii-Trebi, N. I. (2017). Emerging and neglected infectious diseases: Insights, advances, and challenges. *BioMed Research International*, 2017, 5245021.

- Niu, L. N., Zhang, G. N., Xuan, D. D., Lin, C., Lu, Z., Cao, P. P., Chen S. W., Zhang, Y., Cui, X. J., & Hu, S. K. (2023). Comparative analysis of the gut microbiota of wild adult rats from nine district areas in Hainan, China. *World Journal of Gastroenterology*, 29(22), 3469.
- Ouyang, Y., Nkedi-Kizza, P., Wu, Q. T., Shinde, D., & Huang, C. H. (2006). Assessment of seasonal variations in surface water quality. *Water Research*, 40(20), 3800-3810.
- Payne, J., Francis, C. M., & Phillipps, K. (1985). *A field guide to the mammals of Borneo*, pp. 250-260, Kota Kinabalu: The Sabah Society.
- Peña, J. A., Li, S. Y., Wilson, P. H., Thibodeau, S. A., Szary, A. J., & Versalovic, J. (2004). Genotypic and phenotypic studies of murine intestinal lactobacilli: Species differences in mice with and without colitis. *Applied and Environmental Microbiology*, 70(1), 558-568.
- Perlman, D., Martínez-Álvaro, M., Moraïs, S., Altshuler, I., Hagen, L. H., Jami, E., Roehe, R., Pope, P. B., & Mizrahi, I. (2022). Concepts and consequences of a core gut microbiota for animal growth and development. *Annual Review of Animal Biosciences*, 10, 177-201.
- Phillipps, Q., & Phillipps, K. (2018). *Phillipps field guide to the mammals of Borneo*, 2nd ed., pp. 226-243, Oxford, UK: John Beaufoy Publishing.
- Podschun, R., & Ullmann, U. (1998). *Klebsiella* spp. as nosocomial pathogens: Epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clinical Microbiology Reviews*, 11(4), 589-603.

- 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, 3760674.
- Raschle, S., Stephan, R., Stevens, M. J., Cernela, N., Zurfluh, K., Muchaamba, F., & Nüesch Inderbilen, M. (2021). Environmental dissemination of pathogenic *Listeria monocytogenes* in flowing surface waters in Switzerland. *Scientific Reports*, 11(1), 1-11.
- Raulo, A., Allen, B. E., Troitsky, T., Husby, A., Firth, J. A., Coulson, T., & Knowles, S. C. (2021). Social networks strongly predict the gut microbiota of wild mice. *The ISME Journal*, 15(9), 2601-2613.
- Ren, Z., Qu, X., Peng, W., Yu, Y., & Zhang, M. (2019). Nutrients drive the structures of bacterial communities in sediments and surface waters in the river-lake system of Poyang Lake. *Water*, 11(5), 930.
- Reperant, L. A., Hegglin, D., Tanner, I., Fischer, C., & Deplazes, P. (2009). Rodents as shared indicators for zoonotic parasites of carnivores in urban environments. *Parasitology*, 136(3), 329-337.
- Rice, E. W., Baird, R. B., & Eaton, A. D. (2012). *Standard Methods for the Examination of Water and Wastewater*, 23rd ed., pp. 48-62, Washington, DC: American Public Health Association.
- Risely, A. (2020). Applying the core microbiome to understand host-microbe systems. *Journal of Animal Ecology*, 89(7), 1549-1558.

- Rodríguez-Medina, N., Barrios-Camacho, H., Duran-Bedolla, J., & Garza-Ramos, U. (2019). *Klebsiella variicola*: An emerging pathogen in humans. *Emerging Microbes & Infections*, 8(1), 973-988.
- Rosshart, S. P., Vassallo, B. G., Angeletti, D., Hutchinson, D. S., Morgan, A. P., Takeda, K., Hickman, H. D., McCulloch, J. A., Badger, J. H., Ajami, N. J., Trinchieri, G., de Villena, F. P., Yewdell, J. W., & Rehermann, B. (2017). Wild mouse gut microbiota promotes host fitness and improves disease resistance. *Cell*, 171(5), 1015-1028.
- Russel, J. (2021). *MicEco v0.9.19: Various functions for analysis for microbial community data*. Available at: <https://github.com/Russel88/MicEco>. [Accessed on 14 January 2024]
- Schaedler, R. W., Dubos, R., & Costello, R. (1965). The development of the bacterial flora in the gastrointestinal tract of mice. *The Journal of Experimental Medicine*, 122(1), 59-66.
- Scholier, T., Lavrinienko, A., Kallio, E. R., Watts, P. C., & Mappes, T. (2024). Effects of past and present habitat on the gut microbiota of a wild rodent. *Proceedings of the Royal Society B*, 291(2016), 20232531.
- Sepulveda, A. R. (2013). Helicobacter, inflammation, and gastric cancer. *Current Pathobiology Reports*, 1, 9-18.
- Sharon, I., Quijada, N. M., Pasolli, E., Fabbrini, M., Vitali, F., Agamennone, V., Dötsch, A., Selberherr, E., Grau, J. H., Meixner, M., Liere, K., Ercolini, D., de Filippo, C., Caderni, G., Brigidi, P., & Turrone, S. (2022). The core human microbiome: Does it exist and how can we find it? A critical review of the concept. *Nutrients*, 14(14), 2872.

- Shin, J., Lee, S., Go, M. J., Lee, S. Y., Kim, S. C., Lee, C. H., & Cho, B. K. (2016). Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing. *Scientific Reports*, 6(1), 29681.
- Slatko, B. E., Gardner, A. F., & Ausubel, F. M. (2018). Overview of next-generation sequencing technologies. *Current Protocols in Molecular Biology*, 122(1), e59.
- Song, H., Kim, J., Guk, J. H., Kim, W. H., Nam, H., Suh, J. G., Seong, J. K., & Cho, S. (2021). Metagenomic analysis of the gut microbiota of wild mice, a newly identified reservoir of *Campylobacter*. *Frontiers in Cellular and Infection Microbiology*, 10, 596149.
- Staley, C., Pongsilp, N., Nimnoi, P., Kaiser, T., & Sadowsky, M. J. (2018). Influence of physicochemical factors on bacterial communities along the lower Mekong River assessed by Illumina next-generation sequencing. *Water Air and Soil Pollution*, 229(10), 321.
- Suut, L., Mazlan, A., Arif, M. T., Katip, T., Nor Aliza, A. R., & Haironi, Y. (2018). Serovar diversity of *Leptospira* sp. infecting wild rodents in Sarawak, Malaysia. *Tropical Biomedicine*, 35(1), 252-258.
- Suzuki, T. A., & Nachman, M. W. (2016). Spatial heterogeneity of gut microbial composition along the gastrointestinal tract in natural populations of house mice. *PloS ONE*, 11(9), e0163720.
- Taudière, A. (2023). *MiscMetabar v0.33: Miscellaneous functions for metabarcoding analysis*. Available at: <https://github.com/adrientaudiere/MiscMetabar>. [Accessed on 14 January 2024]

- Tay, C. Y., Mitchell, H., Dong, Q., Goh, K. L., Dawes, I. W., & Lan, R. (2009). Population structure of *Helicobacter pylori* among ethnic groups in Malaysia: Recent acquisition of the bacterium by the Malay population. *BMC Microbiology*, 9, 1-13.
- Teng, Y., Yang, X., Li, G., Zhu, Y., & Zhang, Z. (2022). Habitats show more impacts than host species in shaping gut microbiota of sympatric rodent species in a fragmented forest. *Frontiers in Microbiology*, 13, 811990.
- Thayaparan, S., Robertson, I. A. N., Amraan, F., Su'ut, L., & Abdullah, M. T. (2013). Serological prevalence of leptospiral infection in wildlife in Sarawak, Malaysia. *Borneo Journal of Resource Science and Technology*, 2(2), 71-74.
- Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., & Gordon, J. I. (2007). The human microbiome project. *Nature*, 449(7164), 804-810.
- Valdes, A. M., Walter, J., Segal, E., & Spector, T. D. (2018). Role of the gut microbiota in nutrition and health. *British Medical Journal* 2018, 361, 36-44.
- Vega, M., Pardo, R., Barrado, E., & Debán, L. (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research*, 32(12), 3581-3592.
- Viney, M. (2019). The gut microbiota of wild rodents: Challenges and opportunities. *Laboratory Animals*, 53(3), 252-258.
- Vital, M., Stucki, D., Egli, T., & Hammes, F. (2010). Evaluating the growth potential of pathogenic bacteria in water. *Applied and Environmental Microbiology*, 76(19), 6477-6484.

- Wang, J., Lang, T., Shen, J., Dai, J., Tian, L., & Wang, X. (2019). Core gut bacteria analysis of healthy mice. *Frontiers in Microbiology*, *10*, 887.
- Wang, L., Zhang, J., Li, H. L., Yang, H., Peng, C., Peng, Z. S., & Lu, L. (2018). Shift in the microbial community composition of surface water and sediment along an urban river. *Science of the Total Environment*, *627*, 600–612.
- Wang, Y., Zhao, Y., Bolas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology*, *39*(11), 1348-1365.
- Wang, Z., Zhang, C., Li, G., & Yi, X. (2022). The influence of species identity and geographic locations on gut microbiota of small rodents. *Frontiers in Microbiology*, *13*, 983660.
- Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y., & Zemla, J. (2017). Package ‘corrplot’. *Statistician*, *56*(316), e24.
- Weldon, L., Abolins, S., Lenzi, L., Bourne, C., Riley, E. M., & Viney, M. (2015). The gut microbiota of wild mice. *PLoS ONE*, *10*(8), e0134643.
- Wick, R. R., Judd, L. M., & Holt, K. E. (2019). Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biology*, *20*, 1-10.
- Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(2), 180-185.
- Wortelboer, K., Koopen, A. M., Herrema, H., de Vos, W. M., Nieuwdorp, M., & Kemper, E. M. (2022). From fecal microbiota transplantation toward next-generation beneficial

microbes: The case of *Anaerobutyricum soehngenii*. *Frontiers in Medicine*, 9, 1077275.

Yogendra, K., & Puttaiah, E. T. (2008). Determination of water quality index and suitability of an urban waterbody in Shimoga Town, Karnataka. In *Proceedings of Taal2007: The 12th World Lake Conference*, pp. 342-346.

Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M., Harris, H. M., Mattarelli, P., O'toole, P. W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G. E., Gänzle, M. G., & Lebeer, S. (2020). A taxonomic note on the genus *Lactobacillus*: Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. *International Journal of Systematic and Evolutionary Microbiology*, 70(4), 2782-2858.

Zou, Y., Xue, W., Luo, G., Deng, Z., Qin, P., Guo, R., Sun, H., Xia, Y., Liang, S., Dai, Y., Wan, D., Jiang, R., Su, L., Feng, Q., Jie, Z., Guo, T., Xia, Z., Liu, C., Yu, J., Lin, Y., Tang, S., Huo, G., Xu, X., Hou, Y., Liu, X., Wang, J., Yang, H., Kristiansen, K., Li, J., Jia, H., & Xiao, L. (2019). 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nature Biotechnology*, 37, 179-185.

APPENDICES

Appendix A: Sample Information of Rodent Individuals

Sample ID	Field ID	Species	Sex	Age	Sampling Date	Locality
G1	GGNP21-008	<i>Maxomys whiteheadi</i>	M	A	22/10/2021	GGNP
G2	GGNP21-011	<i>Maxomys whiteheadi</i>	F	A	22/10/2021	GGNP
M1	MAE22-005	<i>Niviventer cremoriventer</i>	M	A	20/05/2022	MAE
M2	MAE22-006	<i>Maxomys tajuddini</i>	M	A	20/05/2022	MAE
M3	MAE22-009	<i>Maxomys whiteheadi</i>	M	A	20/05/2022	MAE
M4	MAE22-010	<i>Maxomys tajuddini</i>	M	A	20/05/2022	MAE
L1	LHNP22-015	<i>Niviventer cremoriventer</i>	M	J	03/07/2022	LHNP
L2	LHNP22-016	<i>Sundamys muelleri</i>	M	A	03/07/2022	LHNP
S1	SSK22-006	<i>Sundamys muelleri</i>	M	A	15/07/2022	SSK
S2	SSK22-009	<i>Maxomys tajuddini</i>	M	A	16/07/2022	SSK
S3	SSK22-013	<i>Maxomys surifer</i>	M	A	16/07/2022	SSK
S4	SSK22-014	<i>Niviventer cremoriventer</i>	F	A	16/07/2022	SSK
S5	SSK22-017	<i>Niviventer cremoriventer</i>	F	A	17/07/2022	SSK
U1	UPK23-005	<i>Maxomys tajuddini</i>	M	A	12/03/2023	UPK
U2	UPK23-006	<i>Niviventer cremoriventer</i>	F	A	12/03/2023	UPK
U3	UPK23-009	<i>Rattus tanezumi</i>	M	A	12/03/2023	UPK

(M = Male, F = Female, A= Adult, = J = Juvenile, GGNP = Gunung Gading National Park, MAE = Marup Atas Engkilili, LHNP = Lambir Hills National Park, SSK = Sungai Sibau Kapit, UPK = Ulu Poi Kanowit)

Appendix B: Genbank and BioSample Accession Numbers for Sampled Rodents

Sample ID	GenBank	BioSample
G1	PP054540	SAMN37668508
G2	PP054541	SAMN37668509
M1	PP054544	SAMN37668512
M2	PP054545	SAMN37668513
M3	PP054546	SAMN37668514
M4	PP054547	SAMN37668515
L1	PP054542	SAMN37668510
L2	PP054543	SAMN37668511
S1	PP054548	SAMN37668516
S2	PP054549	SAMN37668517
S3	PP054550	SAMN37668518
S4	PP054551	SAMN37668519
S5	PP054552	SAMN37668520
U1	PP054553	SAMN37668521
U2	PP054554	SAMN37668522
U3	PP054555	SAMN37668523

Appendix C: Multiple Sequence Alignment of Sampled Rodent Sequences

GGNP21-008	TCACTCCTTT	ATTGACCTAC	CTACTCCTTC	CAATATCTCA	TCATGATGAA	ACTTTGGCTC	60
GGNP21-011							60
MAE22-005	...G...C	..C.....	..G.C..A..	...C..T...		..T.....	60
MAE22-006		T...	..C..C..A..	...C..T...		...C.....	60
MAE22-009							60
MAE22-010		T...	..C..C..A..	...C..T...		...C.....	60
LHNP22-015	...G...C	..C.....	..G.C..A..	...C..T...		..T.....	60
LHNP22-016	T...	TT...	..CG.C..A..	...C.....			60
SSK22-006	T...	TT...	..CG.C..A..	...C.....			60
SSK22-009		T...	..C..C..A..	...C..T...		...C.....	60
SSK22-013	C.....C	..C..T...	..C..C..A..	...C.....		A..	60
SSK22-014	...G...C	..C.....	..G.C..A..	...C..T...		..T.....	60
SSK22-017	...G.T..C	..C.....	..G.C..A..	...C..T...		..T.....	60
UPK23-005		T...	..C..C..A..	...C..T...		...C.....	60
UPK23-006	...G...C	..C.....	..G.C..A..	...C..T...		..T.....	60
UPK23-009	C	..C.....T	..CG.C..A..	...C.....		..T.....	60
GGNP21-008	CCTCCTAGGA	GTATGCCTAA	TTATCCAAAT	CGTAACAGGA	CTATTTCTAG	CAATACATTA	120
GGNP21-011							120
MAE22-005	..AT...C	C.....T	..C..T...	..A.C.....	C...	C..	120
MAE22-006	T..T.....	A.....	...T...	TA.....	C...		120
MAE22-009							120
MAE22-010	T.....		...T...	TA.....	C...		120
LHNP22-015	..AT...C	C.....T	..C..T...	..A.C.....	C...	C..	120
LHNP22-016	..A.....	C	..AG.....	..A.C....C		C..	120
SSK22-006	..A.....	C	..AG.....	..A.C....C		C..	120
SSK22-009	T.....	T	...T...	TA.....	C...		120
SSK22-013		T	...T...	..A.T....C		C..	120
SSK22-014	..AT...C	C.....T	..C..T...	..A.C.....	C...	C..	120
SSK22-017	..AT...C	C.....T	..C..T...	..A.C.....	C...	C..	120
UPK23-005	T..T.....		...T...	TA.....	C...		120
UPK23-006	..AT...C	C.....T	..C..T...	..A.C.....	C...	C..	120
UPK23-009	TT.A.....	C	..AG.....	TA.C....C	C...	C..	120
GGNP21-008	CACATCCGAC	ACTATAACAG	CATTCTCATC	AGTTACCCAT	ATCTGCCGAG	ATGTAAACTA	180
GGNP21-011							180
MAE22-005	T.....	..A.....	...T...	..A....C		..C.....	180
MAE22-006	T.....	..C.....	...T...	..A.....	..T.....		180
MAE22-009							180
MAE22-010	T.....	..C.....	...T...	..A.....	..T.....		180
LHNP22-015	T.....	..A.....	...T...	..A....C		..C.....	180
LHNP22-016	T			..C..G...	..T.....	..C.....	180
SSK22-006	T			..C..G...	..T.....	..C.....	180
SSK22-009	T.....	..C.....	...T...	..A.....	..T.....		180
SSK22-013	T.....			..A..A...		..C.....	180
SSK22-014	T.....	..A.....	...T...	..A....C		..C.....	180
SSK22-017	T.....	..A.....	...T...	..A....C		..C.....	180
UPK23-005	T.....	..C.....	...T...	..A.....	..T.....		180
UPK23-006	T.....	..A.....	...T...	..A....C		..C.....	180
UPK23-009		...T.....		C		..C.....	180
GGNP21-008	CGGCTGACTA	ATTGATATA	TACATGCAAA	CGGCGCCTCC	ATATTCTTTA	TCTGCCTATT	240
GGNP21-011				T.....			240
MAE22-005	...T.....			T..A.....	T...	..T.....	240
MAE22-006	T.....	CT	...C.....	T...			240
MAE22-009				T.....			240
MAE22-010	T.....	CT	...C.....	T...	T...		240
LHNP22-015				T..A.....	T...	..T.....	240
LHNP22-016	T..	..C.....CT	C..	T..A..T..A	T..C	T...	240
SSK22-006	T..	..C.....CT	C..	T..A..T..A	T..C	T...	240
SSK22-009	T.....	CT	...C.....	T...	T...		240
SSK22-013		..C.....		T..G..T...		T...	240
SSK22-014	...T.....			T..A.....	T...	..T.....	240
SSK22-017				T..A.....	T...	..T.....	240
UPK23-005	T.....	CT	...C.....	T...			240
UPK23-006	...T.....			T..A.....	T...	..T.....	240
UPK23-009		..C.....CT	...C..C..	..A..T..A			240

GGNP21-008	CCTGCATGTA	GGACGAGGAA	TATATTATGG	ATCTTATACT	TTCATAGAAA	CATGAAACAT	300
GGNP21-011							300
MAE22-005	...C.....		...C..C..	A	...C.....		300
MAE22-006	...A..C...			...A...G.C			300
MAE22-009							300
MAE22-010	...A..C...		...C..	...A...G.C			300
LHNP22-015	...T.....	G.	...C..C..	A	...C.....		300
LHNP22-016	...T..C...	G.	...C..C..	...C..C..A	...C.....		300
SSK22-006	...C..C...	G.	...C..C..	...C..C..A	...C.....		300
SSK22-009	...A..C...		...C..	...A...G.C			300
SSK22-013	TT.A..C...	C.	...C..	...C..C		...C.....	300
SSK22-014	...T.....		...C..C..	A	...C.....		300
SSK22-017	...T.....		...C..C..	A	...C.....		300
UPK23-005	...A..C...		...C..	...A...G.C			300
UPK23-006	...T.....		...C..C..	A	...C.....		300
UPK23-009	...T.....	...G....G.	...C..C..	...C.....	...C.....		300
GGNP21-008	TGGAGTCATT	CTACTATTTG	CAGTAATAGC	AACTGCATTC	ATAGGCTATG	TACTTCCATG	360
GGNP21-011		C.					360
MAE22-005	C.....A..C	..T.....C.	...C.....	...C.....T	A....		360
MAE22-006	C.....G.C	G....		T		...C.....	360
MAE22-009							360
MAE22-010	C.....G.C	G....		T		...C.....	360
LHNP22-015	C..G..A...	..T.....C.	...C.....	...C.....T	A....		360
LHNP22-016	A...	C.	..ACT....	...C.....	T....		360
SSK22-006	A...	C.	..ACT....	...C.....	T....		360
SSK22-009	C.....G.C	G....		T		...C.....	360
SSK22-013	A..C	..C..T..C.	...C.....		C.		360
SSK22-014	C..G..A...	..T.....C.	...C.....	...C.....T	A....		360
SSK22-017	C..G..A...	..T.....C.	...C.....	...C.....T	A....		360
UPK23-005	C.....G.C	G....		T		...C.....	360
UPK23-006	C..G..A...	..T.....C.	...C.....	...C.....T	A....		360
UPK23-009	AG.C		...C.....			...C.....	360
GGNP21-008	AGGACAAATA	TCCTTCTGAG	GAGCAACAGT	AATTACAAAT	CTTTTATCAG	CCATCCCATA	420
GGNP21-011							420
MAE22-005			..C..C....	...C.....C	..CC.....	T..	420
MAE22-006	G.....		..G.....		..CC.....		420
MAE22-009							420
MAE22-010	G.....				..CC.....		420
LHNP22-015			..C..C....	...C.....C	..AC.....	T..	420
LHNP22-016	...C.....		..T..C....	T..C	..AC.....	..T..T....	420
SSK22-006	...C.....		..T..C....	T..C	..AC.....	..T..T....	420
SSK22-009	G.....				..CC.....		420
SSK22-013			..C.....			...T....	420
SSK22-014			..C..C....	...C.....C	..CC.....	T..	420
SSK22-017			..C..C....	...C.....C	..CC.....	T..	420
UPK23-005			..G.....		..CC.....		420
UPK23-006			..C..C....	...C.....C	..CC.....	T..	420
UPK23-009		..A.....	..G..C....	C	..A.....	C..	420
GGNP21-008	TATCGGAACT	ACCCTAGTCG	AATGAATCTG	AGGAGGCTTC	TCAGTAGACA	AAGCTACTCT	480
GGNP21-011						C..	480
MAE22-005	C..T..G...	...T...A.		...C.....		...A...T.	480
MAE22-006	C..T.....	...T.....			T.	...C.....	480
MAE22-009						C..	480
MAE22-010	C..T.....	...T.....				..G..C....	480
LHNP22-015	C..T..G...	...T...A.		...C.....		...A...T.	480
LHNP22-016	C.....	...T...A.		T...	T.	...A..C..	480
SSK22-006	C.....	...T...A.		T...	T.	...A..C..	480
SSK22-009	C..T.....	...T.....				...C.....	480
SSK22-013	C..T.....A			...T.....		T.	480
SSK22-014	C..T..G...	...T...A.	...G.....	...C.....		...A...T.	480
SSK22-017	C..T..G...	...T...A.		...C.....		...A...T.	480
UPK23-005	C..T.....	...T.....			T.	...C.....	480
UPK23-006	C..T..G...	...T...A.		...C.....		...A...T.	480
UPK23-009	...T..C..C			T...	...A.....	...A..C..	480

GGNP21-008	AACACGCTTT	TTCGCATTCC	ACTTTATCT	CCCATTTCATT	ATTGCAGCCC	TTGCAATCGT	540
GGNP21-011						T..	540
MAE22-005	T..	T.	.T.....	T..C		.A.T...T..	540
MAE22-006	T..		...C..C..	T..C		.C.....T..	540
MAE22-009					..C.....	T..	540
MAE22-010	T..		...C..C..	T..C		.C.....T..	540
LHNP22-015	T..	T.	.T.....	T..C		.A.T...T..	540
LHNP22-016		T.	...C..C..	T..	T...	T..	540
SSK22-006		T.	...C..C..	T..	T...	T..	540
SSK22-009	T..		...C..C..	T..C		.C.....T..	540
SSK22-013	A..	T.	...C..C..	T..C		.C.....T..	540
SSK22-014	T..	T.	.T.....	T..C		.A.T...T..	540
SSK22-017	T..	T.	.T.....	T..C		.A.T...T..	540
UPK23-005			...C..C..	T..C		.C.....T..	540
UPK23-006	T..	T.	.T.....	A.....T..C		.A.T...T..	540
UPK23-009		.C.....	...C..C..	C	.C..C....	C....	540
GGNP21-008	CCACTTATTA	TTTCTTCACG	AAACAGGCTC	AAACAATCCA	ACAGGCCTAG	ACTCTGACGC	600
GGNP21-011							600
MAE22-005	...C.....	.C..C..T.	...T..A..	C..C	A...A		600
MAE22-006	..T.....	T.		...T..C..	C...A...		600
MAE22-009		C..					600
MAE22-010		T.		C..	A...		600
LHNP22-015	...C.....	.C..C..T.	...T..A..	C..C	A...A		600
LHNP22-016	...C.TC.C		A..	C..C	AT...	CA...	600
SSK22-006	...C.TC.C		A..	C..C	AT...	.T..CA...	600
SSK22-009		T.		C..	A...		600
SSK22-013	...C.T...	.C.....		C.....C..C	T..A	A....	600
SSK22-014	...C.....	.C..C..T.	...T..A..	C..C	A...A		600
SSK22-017	...C.....	.C..C..T.	...T..A..	C..C	A...A		600
UPK23-005	..T.....	T.		C..	A...		600
UPK23-006	...C.....	.C..C..T.	...T..A..	C..C	A...A	.C.....	600
UPK23-009	A..TC.CC.C	.C..C....	A..	C..C	G...A	T....A....	600
GGNP21-008	AGACAAAATC	CCCTTCCACC	CCTACTACAC	AATTAAAGAT	CTCCTAGGAA	TTTTTATCT	660
GGNP21-011		.T.....	T..				660
MAE22-005	C.....	.A..T...		T.....C	G	.A.....T.	660
MAE22-006	.A.....	T..T.	.A..T..T.	...C.....	GG	.A..C....	660
MAE22-009		.T.....					660
MAE22-010		T..T.	.A..T..T.		GG	C....	660
LHNP22-015	C.....	.A..T...		T.....C	G	.A.....T.	660
LHNP22-016	G...	.A..T...	.A.....	...C.....C	G	.A.....	660
SSK22-006	G...	.A..T...	.A.....	...C.....C	G	.A.....	660
SSK22-009		T..T.	.A..T..T.		GG	C....	660
SSK22-013	G...		.T.....	C..C.....C		C..C..	660
SSK22-014	C.....	.A..T...		T.....C	G	.A.....T.	660
SSK22-017	C.....	.A..T...		T.....C	G	.A.....T.	660
UPK23-005		T..T.	.G..T..T.		GG	C....	660
UPK23-006	C.....	.A..T...		T.....C	G	.A.....T.	660
UPK23-009	T	.A..T..T.	.A.....	C	.A..T...G	.A..C...A..	660
GGNP21-008	CATCCTCTCC	CTAATAATCT	TAGTCCTATT	CTTTCAGAC	CTCCTAGGAG	ACCCAGACAA	720
GGNP21-011				...C.....			720
MAE22-005	AC.A..A..A		T.G.	...C.....			720
MAE22-006	T..	A.....TC	...T.....	...C.....			720
MAE22-009							720
MAE22-010		.T.....C	...T.....	...C.....			720
LHNP22-015	AC.A..A..A		T.G.	...C.....			720
LHNP22-016	AC.TT.A.T.	TC	...A.....	C..		T..	720
SSK22-006	AC.TT.A.T.	C	...A.....	...C.....		T..	720
SSK22-009		.T.....C	...T.....	...C.....			720
SSK22-013	T.....	CT.			.A.....		720
SSK22-014	AC.A..A..A		T.G.	...C.....			720
SSK22-017	AC.A..A..A		T.G.	...C.....			720
UPK23-005	T..	.T.....C	...T.....	...C.....			720
UPK23-006	AC.A..A..A		T.G.	...C.....	C		720
UPK23-009	AC....A.TT	CTC	...AT....	...C.....	.A.....		720

GGNP21-008	CTATACACCA	GCAAACCCCC	TCAATACACC	ACCCCATATT	AAACCAGAAT	GATATTTCTT	780
GGNP21-011		T....	.T.....				780
MAE22-005	...CT....C	..C.....	.A..C..C..	...A..C..C			780
MAE22-006			...C.....	C...			780
MAE22-009		T....	.T.....				780
MAE22-010	...C.....		...C.....				780
LHNP22-015	...CT....C	..C.....	.A..C..C..	...A..C..C			780
LHNP22-016	C	..T....AT	.A..C..C..	...A.....			780
SSK22-006	C	..T....AT	.A..C..C..	...A.....			780
SSK22-009	...C.....		...C.....				780
SSK22-013	T..C		.A..C.....	C		...C.....	780
SSK22-014	...CT....C	..C.....	.A..C..C..	...A..C..C		.G.....	780
SSK22-017	...CT....C	..C.....	.A..C..C..	...A..C..C			780
UPK23-005	...C.....		...C.....				780
UPK23-006	...CT....C	..C.....	.A..C..C..	...A..C..C		.G.....	780
UPK23-009	T..C.....T	..T..T..A.	.A.....C..	...A.....	.G.....	.G.....	780
GGNP21-008	ATTTGCTTAC	GCCATCCTAC	GCTCTATCCC	AAACAAACTA	GGAGGAGTCG	TAGCCCTAGT	840
GGNP21-011		T....					840
MAE22-005	...C..A...	T..C.	...C.....	C.....	G.		840
MAE22-006	C...	T....	...C.....	T..	T.		840
MAE22-009							840
MAE22-010	C...	T....	...C.....	T..	T.		840
LHNP22-015	...C..A...	T..C.	...C.....	C.....	A.		840
LHNP22-016	C...	T....	T..	C.....	A.		840
SSK22-006	C...	T....	T..	C.....	A.		840
SSK22-009	C...	T....	...C.....	T..	T.		840
SSK22-013	...C..C..T	..T..T....		T.....T..	.G....A.		840
SSK22-014	...C..A...	T..C.	...C.....	C.....	A.		840
SSK22-017	...C..A...	T..C.	...C.....	C.....	G.		840
UPK23-005	C...	T....	...C.....		T.	A.	840
UPK23-006	...C..A...	T..C.	...C.....	C.....	A.		840
UPK23-009	C...	..T..T....	.T..C.....	C..T.....	G..A.		840
GGNP21-008	CCTTTCAATC	CTCGTCTAG	CTTTCTTACC	ATTCTCTCCAC	ACCTCAAAAC	AACGAAGCCT	900
GGNP21-011							900
MAE22-005	A..A..T...	T.AA..T...	.C...C...	A..T	..T..T...	...C...T.	900
MAE22-006	...C.....	..T.....	...C.....	...T..T...		T..	900
MAE22-009							900
MAE22-010	...C.....		...C.....	...T..T...		T..	900
LHNP22-015	A..A..T...	T.AA..T...	.C...C...	...T..A..T	..T..T...	...C..TT.	900
LHNP22-016	..A.....T	T.AA.....	.C...C...	.C.A..A...		...C...T.	900
SSK22-006	..A.....T	T.AA.....	.C...C...	.C.A..A...		...C...T.	900
SSK22-009	...C.....		...C.....	...T..T...		T..	900
SSK22-013	.T.A..T..T	..A.....	.C..TC...	T..T	.A.....	T..	900
SSK22-014	A..A..T...	T.AA..T...	.C...C...	...T..A..T	..T..T...	...C...T.	900
SSK22-017	A..A..T...	T.AA..T...	.C...C...	A..T	..T..T...	...C...T.	900
UPK23-005	...C.....	..T.....	...C.....	...T..T...		T..	900
UPK23-006	A..A..T...	T.AA..T...	.C...C...	...T..A..T	..T..T...	...C...T.	900
UPK23-009	..A.....T	..AA.....	.C..TC...	A..T	..T.....	...C...T.	900
GGNP21-008	TACCTTCCGC	CCAATCACCC	AAGCCCTATA	TTGAATTCTA	GTAGCAAATC	TTCTTATCCT	960
GGNP21-011							960
MAE22-005	A..T.....		.GA.....	C.....T	T..C.	.A.....T.	960
MAE22-006	...T.....	T....	...TT....	C.....C...	C.	...G....	960
MAE22-009						T..	960
MAE22-010	...T.....	T....	...TT....	C.....C...	C.	...G....	960
LHNP22-015	A..T.....		.GA.....	C.....T	T..C.	.A.....T.	960
LHNP22-016	A..A..T...	T....	..ATT....	C	.C...C..C.	...CG.T..	960
SSK22-006	A..A..T...	T....	..ATT....	C.....C...	.C...C..C.	...CG.T..	960
SSK22-009	...T.....	T....	...TT....	C.....C...	C.	...G....	960
SSK22-013	C.....T		..AT...C.	C.....C...	T....	.C.....	960
SSK22-014	A..T.....		.GA..T...	C.....T	T..C.	.A.....T.	960
SSK22-017	A..T.....		.GA.....	C.....T	T..C.	.A.....T.	960
UPK23-005	...T.....	T....	...TT....	C.....C...	C.	...G....	960
UPK23-006	A..T.....		.GA.....	C.....T	T..C.	.A.....T.	960
UPK23-009	A..A.....			C.....C...	C....	.CT.C..TT.	960

GGNP21-008	AACATGAATC	GGAGGACAAC	CAGTTGAACA	CCCATTTCATT	ATCATCGGAC	AACTAGCATC	1020
GGNP21-011							1020
MAE22-005		C.....	C.....	T.....T...	..T.....C.	T..	1020
MAE22-006			A.....				1020
MAE22-009							1020
MAE22-010			A.....		..T.....		1020
LHNP22-015			C.....	T.....T...	..T.....C.	T..	1020
LHNP22-016		C.....	A.....	T...	T.		1020
SSK22-006		C.....	A.....	T..C	T.		1020
SSK22-009			A.....		..T.....		1020
SSK22-013	T	C.....	A.....		T..G.		1020
SSK22-014		C.....	C.....	T.....T...	..T.....C.	T..	1020
SSK22-017		C.....	C.....	T.....T...	..T.....C.	T..	1020
UPK23-005			A.....		..T.....		1020
UPK23-006		C.....	C.....	T.....T...	..T.....C.	T..	1020
UPK23-009		G.	A.....	T.....T...	T..C.	..G.....	1020
GGNP21-008	CATCAGCTAC	TTCTCAATTA	TTCTCATCCT	CATACCCATT	TCAGGAGTAA	TCGAAGACAA	1080
GGNP21-011							1080
MAE22-005	...T..T...		AG....	AGCC	..C...A.T.		1080
MAE22-006	T...	C.	T.	A.....T...			1080
MAE22-009							1080
MAE22-010	T...	C.	T.	A.....T...			1080
LHNP22-015	...T..T...		AG....	AGCA	..C...A.T.		1080
LHNP22-016	T.....T...		..C..T.....	A.....A..C	..T...A.T.	..T.....	1080
SSK22-006	T.....T...		..C..T.....	A.....A..C	..T...A.T.	..T.....	1080
SSK22-009	T...	C.	T.	A.....T...			1080
SSK22-013			..C..T.....		A...	..T.....	1080
SSK22-014	...T..T...		AG....	AGCC	..C...A.T.		1080
SSK22-017	...T..T...		AG....	AGCC	..C...A.T.		1080
UPK23-005	T...		T.	A.....T...			1080
UPK23-006	...T..T...		AG....	AGCC	..C...A.T.		1080
UPK23-009			..C..T..T..	A.....A..C	..C...A.T.		1080

Appendix D: Abundance Table of Bacterial Genera of >1% Relative Abundance

Genus \ Sample	G1	G2	M1	M2	M3	M4	L1	L2
<i>Anaerobutyricum</i>	187	39	723	143	234	933	32	338
<i>Anaerostipes</i>	232	147	1718	379	360	1132	71	368
<i>Bacillus</i>	220	254	306	297	291	213	263	164
<i>Bacteroides</i>	973	233	252	250	355	68	75	103
<i>Blautia</i>	646	261	1503	514	1476	1450	121	518
<i>Clostridium</i>	293	1722	473	292	818	650	91	300
<i>Duncaniella</i>	1025	732	105	573	659	171	263	150
<i>Escherichia</i>	227	461	34	484	365	1021	135	73
<i>Flavonifractor</i>	651	353	1077	748	808	645	60	655
<i>Flintibacter</i>	742	524	754	489	178	536	64	871
<i>Intestinimonas</i>	633	357	387	273	140	382	27	501
<i>Klebsiella</i>	840	1985	19	901	268	246	37	17
<i>Lachnoclostridium</i>	908	341	2098	782	1067	3261	111	504
<i>Lachnospira</i>	74	19	291	56	79	919	16	374
<i>Lacrimispora</i>	393	54	759	247	231	841	107	297
<i>Lactobacillus</i>	3116	6862	401	2468	5499	2289	20	2183
<i>Ligilactobacillus</i>	16	27	9	14	16	7	20936	8
<i>Limosilactobacillus</i>	897	565	1166	1468	1086	202	125	921
<i>Muribaculum</i>	892	781	92	509	583	140	468	90
<i>Phascolarctobacterium</i>	1510	8	45	965	1	1	0	185
<i>Prevotella</i>	1217	271	752	279	223	100	15	414
<i>Romboutsia</i>	6	2983	98	6	344	1092	2	7
<i>Ruminococcus</i>	2634	697	2485	7214	3362	1529	548	10864

Genus	Sample	S1	S2	S3	S4	S5	U1	U2	U3
<i>Anaerobutyricum</i>		171	269	141	78	211	412	88	317
<i>Anaerostipes</i>		622	482	259	193	293	561	151	418
<i>Bacillus</i>		277	345	286	418	359	328	270	300
<i>Bacteroides</i>		165	216	517	85	214	708	74	275
<i>Blautia</i>		965	1415	1445	267	544	2369	3063	974
<i>Clostridium</i>		568	516	961	421	815	820	168	512
<i>Duncaniella</i>		244	384	510	559	1544	249	338	1284
<i>Escherichia</i>		43	165	84	1004	48	77	1896	48
<i>Flavonifractor</i>		1020	936	586	403	489	802	308	1296
<i>Flintibacter</i>		1291	1420	840	608	562	957	528	1673
<i>Intestinimonas</i>		489	601	609	240	231	309	128	497
<i>Klebsiella</i>		17	147	15	279	25	103	3211	23
<i>Lachnoclostridium</i>		724	799	1580	180	442	1425	832	797
<i>Lachnospira</i>		170	563	221	148	220	568	1072	207
<i>Lacrimispora</i>		914	395	276	149	340	624	353	411
<i>Lactobacillus</i>		268	2454	3123	23	9	979	7	1253
<i>Ligilactobacillus</i>		23	33	70	10698	125	21	6326	102
<i>Limosilactobacillus</i>		298	1459	429	152	546	322	7	1646
<i>Muribaculum</i>		145	312	434	387	1220	207	162	889
<i>Phascolarctobacterium</i>		606	420	198	0	1	710	0	0
<i>Prevotella</i>		731	132	182	138	441	753	283	910
<i>Romboutsia</i>		37	23	1306	10	443	12	9	35
<i>Ruminococcus</i>		6045	4294	2608	820	3820	3979	205	1491

Appendix E: Abundance Table of Pathogenic Bacterial Genera

Genus	Sample	G1	G2	M1	M2	M3	M4	L1	L2
<i>Clostridium</i>		293	1722	473	292	818	650	91	300
<i>Klebsiella</i>		840	1985	19	901	268	246	37	17
<i>Escherichia</i>		227	461	34	484	365	1021	135	73
<i>Bacillus</i>		220	254	306	297	291	213	263	164
<i>Bacteroides</i>		973	233	252	250	355	68	75	103
<i>Helicobacter</i>		60	1	3	126	25	2	205	502
<i>Staphylococcus</i>		255	123	333	161	242	299	63	128
<i>Enterococcus</i>		323	145	115	103	193	117	99	82
<i>Streptococcus</i>		406	153	206	131	226	117	48	91
<i>Salmonella</i>		101	172	3	126	51	187	23	10
<i>Enterobacter</i>		127	189	3	105	101	113	20	7
<i>Porphyromonas</i>		96	52	15	33	43	9	15	52
<i>Vibrio</i>		21	79	30	26	26	14	7	6
<i>Citrobacter</i>		33	12	0	50	128	42	0	1
<i>Aeromonas</i>		35	49	19	23	26	28	1	10
<i>Pseudomonas</i>		28	13	36	17	23	31	10	8
<i>Serratia</i>		24	38	2	24	38	26	4	1
<i>Proteus</i>		20	34	7	26	9	9	3	2
<i>Ureaplasma</i>		13	23	13	20	2	3	6	3
<i>Corynebacterium</i>		7	14	24	5	3	3	26	29
<i>Elizabethkingia</i>		10	5	20	6	10	32	2	5
<i>Burkholderia</i>		7	0	16	8	21	12	4	6
<i>Yersinia</i>		16	27	0	10	21	9	1	4
<i>Listeria</i>		17	10	18	5	19	17	9	3
<i>Actinomyces</i>		1	2	45	4	2	3	0	17

<i>Stenotrophomonas</i>	3	5	9	3	12	4	3	3
<i>Treponema</i>	13	1	1	9	0	7	0	3
<i>Brucella</i>	4	8	12	13	4	1	4	8
<i>Mycoplasma</i>	5	4	4	1	2	4	0	0
<i>Providencia</i>	3	2	0	5	2	2	1	0
<i>Morganella</i>	7	4	0	1	13	4	1	0
<i>Neisseria</i>	6	2	3	5	1	9	0	2
<i>Fusobacterium</i>	17	8	0	1	11	5	1	5
<i>Campylobacter</i>	1	0	14	3	2	1	4	7
<i>Francisella</i>	6	4	6	3	1	3	3	2
<i>Bordetella</i>	3	3	4	2	2	3	1	1
<i>Moraxella</i>	6	0	10	1	2	11	8	3
<i>Haemophilus</i>	1	0	0	3	2	5	1	0
<i>Pasteurella</i>	1	3	0	4	2	4	1	3
<i>Shigella</i>	2	1	0	2	2	2	0	0
<i>Mannheimia</i>	1	0	5	0	1	0	0	1
<i>Legionella</i>	0	0	0	2	0	3	0	0
<i>Ehrlichia</i>	1	0	0	0	0	0	0	0
<i>Gardnerella</i>	0	0	2	0	1	0	0	1
<i>Mycobacterium</i>	0	1	1	2	0	0	0	1
<i>Rickettsia</i>	1	0	0	2	0	3	0	0
<i>Chlamydia</i>	0	0	1	0	0	1	0	0
<i>Coxiella</i>	0	0	0	0	0	0	2	0
<i>Bartonella</i>	1	0	0	0	1	0	0	0
<i>Leptospira</i>	0	0	0	0	0	0	0	0
<i>Nocardia</i>	0	0	0	0	0	1	0	0

Genus	Sample	S1	S2	S3	S4	S5	U1	U2	U3
<i>Clostridium</i>		568	516	961	421	815	820	168	512
<i>Klebsiella</i>		17	147	15	279	25	103	3211	23
<i>Escherichia</i>		43	165	84	1004	48	77	1896	48
<i>Bacillus</i>		277	345	286	418	359	328	270	300
<i>Bacteroides</i>		165	216	517	85	214	708	74	275
<i>Helicobacter</i>		1204	52	5	63	703	74	29	578
<i>Staphylococcus</i>		257	183	128	207	234	289	263	380
<i>Enterococcus</i>		80	525	98	228	115	101	167	88
<i>Streptococcus</i>		324	115	88	121	85	98	73	149
<i>Salmonella</i>		10	51	8	167	8	22	541	14
<i>Enterobacter</i>		3	105	3	97	5	22	517	8
<i>Porphyromonas</i>		80	33	62	83	107	39	23	92
<i>Vibrio</i>		29	61	14	34	76	19	63	17
<i>Citrobacter</i>		0	46	0	24	1	2	105	3
<i>Aeromonas</i>		13	23	9	22	31	23	73	20
<i>Pseudomonas</i>		31	22	21	28	24	34	19	19
<i>Serratia</i>		1	5	1	31	0	7	149	1
<i>Proteus</i>		1	59	4	23	12	24	109	6
<i>Ureaplasma</i>		13	35	15	17	75	1	32	17
<i>Corynebacterium</i>		29	18	31	35	20	5	2	26
<i>Elizabethkingia</i>		17	15	9	6	21	48	9	14
<i>Burkholderia</i>		11	28	15	14	17	15	7	11
<i>Yersinia</i>		0	20	1	10	1	5	51	1
<i>Listeria</i>		10	13	6	11	8	8	6	12
<i>Actinomyces</i>		7	4	7	9	4	5	14	7
<i>Stenotrophomonas</i>		7	17	9	19	20	4	3	9
<i>Treponema</i>		19	2	5	6	11	2	1	40
<i>Brucella</i>		3	9	11	2	10	5	3	11
<i>Mycoplasma</i>		4	14	5	21	13	7	1	6
<i>Providencia</i>		0	4	0	3	2	18	47	2
<i>Morganella</i>		0	1	0	2	0	14	37	3

<i>Neisseria</i>	6	10	4	6	13	9	2	8
<i>Fusobacterium</i>	4	4	5	4	4	10	0	1
<i>Campylobacter</i>	8	5	0	1	15	6	1	1
<i>Francisella</i>	2	1	4	4	3	5	11	5
<i>Bordetella</i>	4	2	1	10	3	3	2	8
<i>Moraxella</i>	1	0	2	0	1	0	1	0
<i>Haemophilus</i>	1	13	0	2	2	3	4	5
<i>Pasteurella</i>	2	7	0	0	1	1	3	1
<i>Shigella</i>	0	1	0	8	0	0	7	0
<i>Mannheimia</i>	5	2	0	0	4	1	0	0
<i>Legionella</i>	0	0	0	3	2	1	2	3
<i>Ehrlichia</i>	0	2	0	1	1	9	0	0
<i>Gardnerella</i>	0	0	1	0	0	5	2	1
<i>Mycobacterium</i>	0	0	0	0	0	6	1	1
<i>Rickettsia</i>	2	0	2	0	0	0	1	0
<i>Chlamydia</i>	0	0	0	1	4	1	0	0
<i>Coxiella</i>	0	1	0	1	0	2	0	2
<i>Bartonella</i>	0	1	0	0	0	0	0	1
<i>Leptospira</i>	0	0	1	0	0	0	0	2
<i>Nocardia</i>	0	1	0	0	0	0	0	0

Appendix F: Correlogram Showing Correlations Between Bacterial Diversity and Water Quality Parameters. Negative Correlations are Shown in Red and Positive Correlations are Displayed in Blue. The Intensity of the Colour is Proportional to Correlation Strength. Values Labelled With ‘*’ are Significant at $p < 0.05$

