A SURVEILLANCE STUDY ON THE SPLEEN MICROBIOME OF WILD RODENTS AND SHREWS CAUGHT IN MARIGAT, BARINGO COUNTY, KENYA

BY

REHEMA MUKUVILANI LIYAI

A THESIS SUBMITED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF SCIENCE IN MOLECULAR MICROBIOLOGY AND BIOTECHNOLOGY

SCHOOL OF PHYSICAL AND BIOLOGICAL SCIENCES

MASENO UNIVERSITY

DECLARATION

I declare that what has been presented herein is my own original work and has not been presented in any other institution or university for the award of degree. This thesis, or any part of it, may not be reproduced without written permission from the author, my supervisors and Maseno University.

Sign……...…… Date……

Rehema M. Liyai
MSc/Sc/0016/2016

Supervisors Declaration

We (the supervisors) confirm that the research findings in this thesis were carried out by the said candidate under our supervision.

Sign……...…… Date……

Prof. David M. Onyango (PhD)
School of Physical and Biological Sciences
Department of Zoology
Maseno University

Sign……...…… Date……

Dr. John N.Waitumbi (PhD)
United States Army Medical Research Directorate,
Kenya / Kenya Medical Research Institute

ACKNOWLEDGMENTS

It is with immense gratitude that I acknowledge the support and help of my advisor Dr. John Waitumbi (PhD). I sincerely do thank you, for the confidence you have instilled in me. Thank you for giving me the opportunity to grow, for your guidance, moral and financial support along the way. Thank you for being my advisor and my mentor. You were always there in times of need and I will always appreciate you.

I would like to extend my gratitude to my Supervisor Prof. David Miruka Onyango (PhD) for your advice, support and input in this project. Thank you for your successful guidance and your patience. Thank you for guiding me step by step throughout my postgraduate journey.

In addition, special thanks go to Dr. Paul Ang'ienda, Dr John Ochieng (CDC),Dr. Cyrus Ayieko –Maseno, Dr. Patrick Onyango- Maseno, and Dr. Benson Nyambega-Maseno for all the great skills they impacted in me through the journey. Not forgetting my colleagues and the entire Walter Reed Laboratory Personnel. My appreciation goes to Beth Mutai, Kimita Gathii, Clement Masakhwe, Rachel Githii, Allan Lemtudo, Arnold Lambisia and Fridah Mwende with whom we worked together closely in the laboratory.

My deepest gratitude goes to my family for their unending love and support throughout my whole life and this journey of Education. I wouldn't be any close to where I am without them. I am indebted to my father Paul Liyai and Mother Susan Luganu.

DEDICATION

I dedicate this work to my family. A special feeling of gratitude to my loving parents, Paul Lukulu Isayi, Susan Luganu and Elizabeth Wagithi Wanja who have been a constant source of support and encouragement through my graduate school and life. My siblings: Melody, Dan Shalom, Dan Calvin, Mitchelle, Lesley and Eliza who have always prayed for me and encouraged me when times got tough.

I also dedicate this dissertation to my large family, my aunties, uncles and grandparents for their great support and prayers during this journey. To all my friends who have supported me throughout the process, I will always appreciate all they have done.

ABSTRACT

There is an observed global increase in emerging infectious diseases, with 13% of over 1400 known human pathogens being classified as emerging or re-emerging. Majority of the emerging pathogens are zoonoses, majority of which have animals (domestic and wildlife) as reservoirs. Of the animals, small mammals represent 40% of mammalian species and because of their widespread distribution and accelerated human encroachment to their habitats; they provide great opportunities for disease transmission to humans. The nomadic pastoralists are most likely at greater risk of encountering zoonoses due to their animal husbandry lifestyle. In Kenya, the prevalence of this tick-borne bacteria has been seen to be on the rise. For instance, *Coxiella burnetti ,Orientiachuto,* Rickettsia *spp*, and *Bartonella spp,* have been detected in vectors, animals and humans in different regions in Kenya. This study investigated the bacterial microbiota of wild caught small mammals of Marigat in Baringo County, Kenya. Communities in this county are nomadic pastoralists. Fifty-four small mammals were trapped from different sites in Marigat. DNA was extracted from the spleen and used to amplify the hyper-variable V3-V4 region of bacterial 16S ribosomal RNA (rRNA). The spleen is a peripheral lymphoid organ in vertebrates that aids in filtering blood. It plays an important role in the modulation of immune responses and hematopoiesis. The spleen can be infected by bacteria during the blood filtration and can therefore be used as an indicator for microorganisms harbored by wild animals. This study assessed bacterial diversity in the spleen of wild caught small mammals. The V3-V4 region demonstrates considerable sequence diversity among different bacteria and can identify all bacteria species to the genus level, with an exception of the closely related enterobacteriaceae. The generated amplicon libraries were sequenced on the Illumina MiSeq. Sequence data were analyzed with Mothur v1.35, queried against the Silver database and visualized on R.For taxonomic classification of the small mammals, cytochrome B (*Cytb*) and cytochrome oxidase subunit 1 (*COI*) genes were amplified and thereafter sequenced using Sanger method on a Genetic Analyzer. CLC main workbench was used to assemble the data into contigs. The sequences were then queried against the reference sequence database. A phylogenetic tree was then inferred using the MEGA software v7. By molecular taxonomy, the small mammals were classified as 41 rodents and 13 shrews. 175,629 sequences were obtained and classed into 196 operational taxonomic units (OTUs), based on unique sequences that mapped to 18 bacteria phyla, with 4 phyla accounting for 97% of the total OTUs. 18 phyla and 196 bacteria genera were detected. Of these phyla, Proteobacteria was the most abundant contributing 64.7% of total contigs. Other phyla included Actinobacteria (18.0%), Firmicutes (6.1%), Chlamydiae (3.8%), Chloroflexi (2.6%) and Bacteroidetes (1.9%) among others. Of the pathogenic bacteria genera, *Bartonella* was the most abundant (41.5%), followed by *Anaplasma* (6.5%), *Methylobacterium* (3.6%), *Delftia* (3.2%), *Coxiella* 2.6%, *Bradyrhizobium* (1.6%) and *Acinetobacter* (1.3%). Other less abundant (<1%) pathogenic bacteria included *Ehrlichia, Rickettsia, Leptospira, Borrelia, Brucella, Chlamydia and Streptococcus*. *Acomys* carried higher bacteria diversity than other small mammals at Shannon diversity index of 3.0 compared to 2.3 for *Rattus*, 2.2 for *Arvicathis* and *Crocidura* and 1.8 for *Mastomys*. This study confirms the role of the spleen as a microbial repository and its suitability for studying microbial pathogens, and utility of 16S rRNA deep sequencing in characterizing the complex microbiota in the spleens of wild rodents and shrews. An inherent problem with the V3-V4 region of 16S rRNA is the inability to classify the bacteria reliably beyond the genera. Future studies should utilize the newer long read methods of 16S rRNA analysis that are able to delimit the species composition.

TABLE OF CONTENTS

ABBREVIATIONS AND ACROYMNS

µl	Microliter
16s rRNA	16 Svedbergribosomal Ribonucleic Acid
Bp	base pair
CDC	Center for disease control
CO1	Cytochrome C oxidase subunit 1 gene
Cyt B	Cytochrome B mitochondrial gene
DNA	Deoxyribonucleic acid
mtDNA	Mitochondrial Deoxyribonucleic acid
MALDI-TOF	Matrix-assisted laser desorption ionization time-of-flight mass spectrometry
ELISA	Enzyme-linked immunoassay
NGS	Next generation sequencing
PCR	Polymerase chain reaction
RNA	Ribonucleic Acid
SSU	Small subunit ribosomal ribonucleic acid
WHO	World Health organization
STG	Scrub typhus Group
PBS	Phosphate-buffered Saline

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Infectious diseases remain the leading cause of death and morbidity among humans worldwide. More than half of these diseases are vector borne zoonoses that originate from domestic and wild animals (Galan *et al*., 2016). Small mammals which represent 40% of mammalian species are distributed in all continents except Antarctica(Delić *et al*., 2013). They are among the vertebrate species that have benefited from human movements enabling them to spread and establish themselves worldwide. Their wide distribution and accelerated anthropogenic environmental changes will undoubtedly lead to increase in opportunities for disease transmission to humans (Diagne et al., 2017)(Diagne et al., 2017) (Diagne *et al*., 2017). Small mammals have been identified as major host of zoo noses (Wilson & Reeder, 2005). The small mammals include rats, mice, squirrels, porcupines, beavers, guinea pigs and hamsters among others. These small mammals are kept as pets(squirrels and hamsters), research animals (guinea pigs), or provide food to certain communities such as certain parts of Thailand, Ghana and Laos(Gruber, 2016).These small mammals play an important role in the ecosystems and in some countries they have been classified as endangered as humans now hunt them for bush meat or for medicinal products(Ripple *et al*., 2016). They however harbor disease causing micro-organisms that include a vast range of bacteria, protozoa, viruses and helminthes. These micro-organisms can be transmitted to human and cause devastating consequences(Lindahl & Grace, 2015).

Both rats and mice have been reported to transmit over 35 diseases worldwide(Center for Disease Control and Prevention, 2017). Their proximity to humans habitation favors pathogen

transmission to human that cause diseases such as Hantavirus pulmonary syndrome, hemorrhagic fever with renal syndrome, plague, leptospirosis, salmonellosis, brucellosis, rat-bite fever, tularemia, Lassa fever, lymphocytic chorio-meningitis, Omsk hemorrhagic fever, and South American arena-viruses among others (CDC, 2016).Although small mammals may transmit diseases such as leptospirosis, salmonellosis and brucellosis by contaminating the environment, the main route of disease transmission between small mammals and humans is ectoparasites such as ticks, lice, fleas, and mites. Small mammal borne ectoparasites are known to transmit diseases such as rickettsiosis, Lyme disease, scrub typhus, babesiosis, cutaneous and visceral leishmaniasis, Colorado tick fever, murine typhus, relapsing fever and human granulocytic anaplasmosis among others(CDC, 2019).

Anaplasma, Ehrlichia, bartonella, and other spotted fever group (STG), are some of the bacteria that are most commonly transmitted to mammalian hosts via vectors. They target the cells and white blood cells of there hosts. These bacteria have evolved and adapted strategies to help them evade or suppress host protective immune responses and thus causing febrile illness in both animals and humans (Ge *et al*., 2018).

The increasing incidences of both known and novel pathogens identified in small mammals have led to renewed interest in rodents and rodent-borne pathogens (Schmidt *et al*., 2014). Rise in human population and the resultant encroachment into wildlife habitats may lead to increase prevalence of rodent-borne zoonotic infections. Proper understanding of small mammals and their associated pathogens is therefore important for control strategies of these zoonoses. This has occasioned an increase in microbial surveillance and pathogen discovery studies to inform strategies geared towards breaking these disease transmission cycles(Diagne *et al*., 2017).In Kenya, the prevalence of this tick-borne bacteria has been seen to be on the rise. For instance,

Coxiella burnetti, *Orientiachuto*, *Rickettsia spp*, *E. chaffeensis*, *Bartonella spp*., and *Babesia spp*., have been detected in vectors, animals and humans in different regions in Kenya.(Masakhwe *et al*., 2018b; Mutai *et al*., 2019; Oswe *et al*., 2018; Thiga, Mutai, Eyako, *et al*., 2015). However, the existence of this bacteria in small mammals is still unknown in Marigat, Baringo County.

According to Muller *et al*., (2013), there is a specific association between pathogens and their host reservoirs and accurate taxonomic assignment of rodents/shrews is essential for better understanding of the probable occurrences of the pathogens they carry (Lu *et al*., 2012; Müller *et al*., 2013).

Morphological classification that relies on structural differences has been found wanting and miss-classification occurs due to similar indistinguishable characteristics. Molecular biology has revolutionized taxonomy and provides a higher resolution and has thus been proven to be better in unraveling hidden mammalian species that are over looked when morphological methods alone are used (Lu *et al*., 2012). Mitochondria genes such as the cytochrome c oxidase subunit 1 (cox1 or CO1) and cytochrome b (cytb) gene are commonly used for mammalian bar-coding (Pentinsaari *et al*., 2016; Tobe *et al*., 2009). The cytb gene is about 1,149 base pairs and it offers more taxonomic information compared to cox1 (Tobe *et al*., 2009) and therefore gives more accurate reconstructions and better resolution for separating species (Nicolas *et al*., 2012). This study used Cytb and CO1 gene for the molecular barcoding of the wild caught small mammals and 16S rRNA deep sequencing to characterize the spleen.

1.2 Statement of the Problem

Wild rodents and shrews host many pathogens that can cause major public health, economic and social damage. They function as reservoirs for approximately 46% of all global zoonoses hence a threat to public health (Lohmus *et al*, 2013). Small mammals are known to harbor ectoparasites such as ticks, mites and fleas which aid in transmission of zoonotic diseases such as Ricketsiosis, Coxiella, and Bartonella. They potentiate risk to both veterinary and human found within their ecosystems. According to Thiga*et al* (2015), it is reported that patients who visited Marigat District Hospital with febrile illness of unknown cause, had seroprevalence of 10% for sported fever group, and <1% for typhus group. These infections are reported to be transmitted by small mammals found within the pastoralist ecosystems where they interact between goats, sheep, cattle, dogs and ultimately humans. Patients who presented with Febrile illness also had a high immunoglobulin (IgG) titer ranging between 1600-6400/dl of blood analyzed (Thiga*et al*., 2015). This acted as an indication that these patients had repeated exposure to sported fever group or scrub typhus that is hypothesized to be associated with small mammals in Baringo County. However, the association of Febrile illness with the prevalence of small mammals in the region has not been ascertained in Baringo county in order to source tract the infection points. It is for this reason that this study was muted in order to establish the microbial biomes that are associated with Febrile illness in Marigat District Baringo County.

1.3 Study Justification

The purpose of this study was to examine the small mammals captured in different sites in Marigat Baringo County in order to do a profiling of the bacteria diversity they harbor. Acute febrile illness occurs at high incidence not only in Marigat Baringo County but in Kenya as a

whole, and the etiology of many of these illnesses is unknown. Prevalence of bacterial organisms such as bartonellosis (Halliday *et al*., 2013) have been identified and studied in previous research. However, there's lack of knowledge on the bacterial profile of small mammals in Marigat Kenya.

Majority of these infectious diseases carried by animals are shared with humans. This can be attributed to the increasing interactions between animals and humans either at the farms, shared homesteads, during animal trade or due to encroachment to wildlife habitats. Scrub typhus, Q fever, brucellosis, Bartonellosis, tularemia and leptospirosis are among the frequently reported zoonotic infections that may have their origin in small mammals that live in the wild.

According to Halliday2013, Bordetellosis has been seen to have a high prevalence in Kibera, Nairobi Kenya as well as Asembo in the Nyanza region. This study will use deep sequencing to explore the identity and diversity of bacterial biota harbored by small mammals so as to close the gap on the paucity of knowledge on this subject.

1.4 Significance of the study

The current study sort to comprehensively identify bacterial pathogens harbored by wild caught rodents and shrews so as to understand the disease risk to communities living in this region. This knowledge could be used as a basis for implementing disease control and prevention strategies.

Findings of this study shall present useful data to support regional residents and farmers with timely and focused implementation of personal protective measures, disease prevention and control programs.

1.5 Objectives of the Study

1.5.1 General Objectives

To determine the identity and diversity of zoonotic bacteria biota in wild rodents and shrews in Marigat, Baringo County.

1.5.2 Specific Objectives

i. To taxonomically classify the wild small mammals using the Cytb and CO1 genes.

ii. To establish the pathogenic bacteria biota in the wild caught small rodents and shrews using the 16SrRNA gene.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

In this chapter, literature is reviewed and analyzed to corroborate the overall and specific objectives of this study namely: to determine the genera/species of the wild caught rodents and shrews and to establish the identity and diversity of zoonotic bacteria biota found in these animals. Only relevant literature were considered and presented. Where possible, data is presented in figures and/or tables and where appropriate statistical inferences included.

2.2 Definition of Small Mammals

Small mammals are animals that weigh ≤500 g or 1kg when adult. They are terrestrial and arboreal in nature, representing the largest order in class mammalian (Norbu, 2016). Most species are herbivores, some are carnivores and insectivorous. Their population impact humans by destroying vast quantities of food per year. Small mammals damage stored food either by direct damage, consumption and contamination by feces and urine, thus they affect both the quantity and quality of food. According to Singleton*et al*., (2003), the annual loss of food due to rodents corresponds to 11 kg of food per person, which translates to over 77 million metric tons annually, in a world of over 7 billion inhabitants. In Asia for instance, 5–17% of the rice crop is eaten by rodents and by estimates, this amount of food is enough to feed >200 million people(Golden *et al*., 2015). In Kenya, Mutambuki & Ngatia (2006)reported that the main causes of damage and weight loss in maize by small mammals accounted a food crop loss of 1.5 %.In addition to crop loss, small mammals are carriers of many infectious diseases, including more

than 20 viruses such as Hanta virus, Crimean Congo Haemorrhagic Fever Virus, more than forty bacteria with leptospirosis, borreliosis, erhlichiosis, relapsing fever, Q-fever and Lyme disease which they transmit directly or indirectly to humans. This is in addition to other parasites like *babesia* and *theileria* (Golden *et al*., 2015). There are several studies that corroborate these assertions (Byrkjeland *et al*, 2015, Lee *et al*., 2010and Han *et al*., 2015). These devastating effect of small mammals are due to their ability to reproduce rapidly and reach large individual numbers in a wide range of habitats, justifying their epidemiological significance as reservoirs and hosts of pathogens (Hornok *et al*., 2015).

2.3 Taxonomic Classification of Small Mammals

Small mammals are important host for various zoonoses that threaten both the animal and human population worldwide. As observed by Lu *et al*., (2012), there is a specific association between pathogens and host animals. Therefore gaining accurate taxonomic information of species that carry harmful pathogens as reservoirs or hosts is essential for better understanding of occurrences, pathogen proliferation and transmission between animal vectors as well as for epidemiological inferences of rodent/shrew borne diseases (Müller *et al*., 2013). Mammals have been studied previously and their taxonomy and diversity species have been clearly documented in literature (Lu *et al*., 2012). The classification has been done either morphologically by using structural features such as mouth length, shape of feet, eye color, size and shape to classify the organism or using molecular methods such as next generation sequencing and Sanger sequencing(Müller *et al*., 2013).

2.3.1 Order Rodentia

Rodents are warm blooded small mammals that belong to the order Rodentia. The word rodent comes from the Latin word '*rodere'* which means "to gnaw" (Kay & Hoekstra, 2008). In the order are animals such as rats, mice, hamsters, guinea pigs and gerbils. In history rodents, have been used as food and for fur and in the recent world, they are also used as study animal models in research. In the wild, they live in close proximity to human habitation as they benefit from humans on shelter and food.

2.3.2 Order Eulipotyphyla (Shrews)

The order Eulipotyphyla contains small mammals such as shrews. They have the same characteristics as rodents but smaller in size and with longer mouths. They also live in close proximity to humans where they derive food and shelter. They are very social animals and have large reproductive potential. Most small mammals may reach large individual numbers in a wide range of habitats, justifying their epidemiological significance as reservoirs and hosts of pathogens (Hornok *et al*., 2015).

2.4 Morphological Classification of Rodents and shrews

Morphological identification of rodents and shrews can be done using identification keys based on phenotypic characteristics such as weight and body length, tail and hind-paw(Backhans, 2011).Other structural features used in classification include mouth length, shape of feet, eye color, size and shape.

For morphological classification, characteristics used in taxonomy should be easily observed, conservative and contrasting. However, miss-classification occurs due to characteristics similarities and thus, structural taxonomy is not reliable. Molecular biology has revolutionized

taxonomy and has shown ability to identify hidden mammal species that are over looked when using morphological methods alone. It is therefore of major importance to classify small mammals using standard molecular techniques such as DNA bar-coding that uses mitochondrial genes so as to complement morphological methods and thus reducing uncertainties in the identification of organisms (Lu *et al*., 2012).

2.5 Molecular classification of small mammals (DNA Bar-Coding)

DNA barcoding is a molecular tool for species discovery based on the analysis of standardized gene sequences.(Dudu *et al.*, 2016). Rapid and accurate identification of small mammalian species has been achieved by use of molecular DNA bar-coding techniques. These methods utilize the sequence diversity present in short universal DNA sequences to discriminate among known species and potentially discover new ones (Hebert *et al*., 2003a).The genetic loci of choice for many taxonomic and phylogenetic studies are primarily found on the mitochondrial genome. The standardized molecular identification is therefore known as "DNA bar-coding" and it has been used widely in the recent years. These methods are able to provide genetic information to confirm identifications that may have been done earlier in the fields by scientist who may have limited taxonomic background (Lu *et al*., 2012). DNA bar-cording has several advantages among them being small number of biological samples is needed and its applicability for all life stages and differentiation among phenotypically alike species. It is a reliable method to identify species in the absence of morphological techniques, to discover species mislabeling and even intentional species substitutions. The DNA barcoding method is fast, cost efficient, and allows the identification of species from different biological samples, thus being a useful tool for conservation purposes. On the other side however, DNA barcoding may be totally inefficient

when is dealing with hybrids and thus it should be avoided for species that can naturally hybridize(Dudu *et al*., 2016).

Commonly utilized DNA barcodes include cytochrome c oxidase 1 (*CO1)*, cytochrome "b" (*Cytb)*, 16S and 12SrRNA.The cytochrome b (*Cytb*) and Cytochrome C oxidase 1 (*CO1*) have been used to determine species boundaries in mammals. These genes have been used in multiple cases for taxonomic classification. However a study comparing the values of *Cytb* and *CO1* for phylogenetic reconstruction and identification of mammal species showed that the *Cytb* gene gave more accurate reconstructions and better resolution for separating species (Nicolas *et al*., 2012).

Therefore, the *CO1* gene if used on its own is limited and insufficient for reliable molecular phylogenetic reconstruction. In this study, we incorporated both the *CO1* and *Cytb*genetic markers gene for small mammal species phylogenetic identification so as to increase the discrimination success.

2.6 Cytochrome c oxidase I (*CO1*) mitochondrial DNA gene

CO1 mitochondrial DNA geneis one of three mitochondrial DNA (mtDNA) encoded subunits of respiratory complex IV.COX is the last enzyme in the electron transport chain, reducing oxygen and pumping protons across the inner mitochondrial membrane. It codes for a protein that has an essential role in cellular respiration. The approximately 650bp gene was proposed as a universal marker for species identification that is being used as a "DNA barcode" in the animal kingdom (Pentinsaari *et al*., 2016).

Mitochondria are organelles, remnants of ancestral bacterial endosymbionts, and are usually found in almost all eukaryotic cells. Along with plastids in plants, mitochondria are the only

cytoplasmic organelles in the eukaryotic cell that carry genetic elements (Ladoukakis & Zouros, 2017). The mitochondrial genome in animals is preferred for taxonomic classification since unlike the nuclear genome; it lacks introns, has limited exposure to recombination and has a haploid mode of inheritance(Ladoukakis & Zouros, 2017).Herbert *et al*,. (2003) described a technique that uses universal primers to amplify a 650 bp region of the *COI* gene. With this technique, they reported 100% success in bar-coding Lepidoptera(Hebert *et al*., 2003).

DNA bar-coding technique uses a short DNA sequence on the 5′ half of the cytochrome c oxidase I (*CO1*) gene(Nicolas *et al*., 2012). However, the systematic information content of the *CO1* barcode is, and therefore this fragment alone is insufficient for reliable molecular phylogenetic reconstruction and the assignment of new species. It therefore has to be used in conjunction with other genes for best taxonomic classification.

2.7 Cytochrome B (*Cytb)*

Cytochrome b *(Cytb)* is a protein found in the mitochondria of eukaryotic cells. It functions as part of the electron transport chain and is one of the 11 components of a group of proteins called complex III. Within the mitochondria, complex III proteins perform one step of oxidative phosphorylation, in which oxygen and simple sugars are used to create adenosine triphosphate (ATP), the cell's main energy source. During oxidative phosphorylation, the protein complexes, including complex III, drive the production of ATP through a step-by-step transfer of electrons. *Cytb*is involved in the transfer of electrons through complex III (Beattie, 1989).The gene is about 1149 bp and it offers more taxonomic information compared to *COI*(Tobe *et al*., 2009).

Figure 1 below shows the classification of the small mammals based on both morphological and molecular techniques.

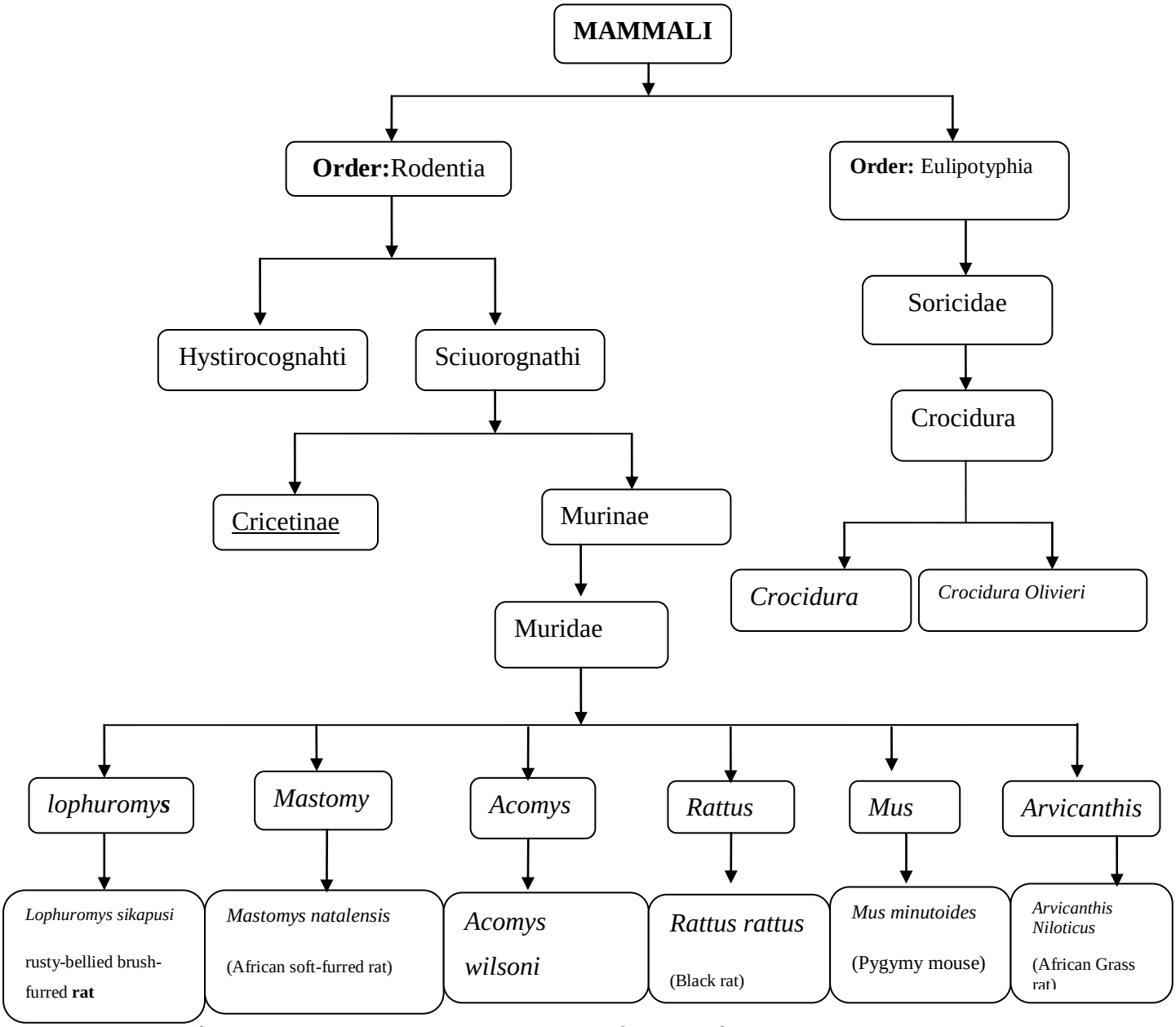

Figure 1 A flow chart showing the taxonomic classification of Rodents and Shrews.

(Taxonomic classification information obtained from NCBI on 22nd February 2018.)

2.8 Pathogens harbored by Small mammals

Zoonotic diseases are of major concern worldwide and form an important proportion of emerging human pathogens in the world(Rabozzi *et al*., 2012). The word 'Zoonosis' was introduced by Rudolf Virchow in 1880 to include collectively the diseases shared in nature by man and animals. Soon after that the World Health Organization (WHO, 1959)expanded on the definition as "those diseases and infections which are naturally transmitted between vertebrate animals and man"(Gezmu *et al*., 2017). In Africa and particular Kenya, there is limited information on zoonotic diseases (Belay *et al*., 2017). This is majorly due to the cost associated with new techniques of determining these diseases.

Zoonotic diseases can be transmitted to humans in several ways: animal bites and scratches, contaminated food, and vectors like mosquitoes, tick, fleas, and lice's (Morwal, 2017).The consequence of zoonotic infections may range from very mild to potentially life threatening diseases such aslymphocytic chorio-meningitis and plague(Bruce H. Williams , 2013).Commonly reported zoonotic and viral diseases include Q- fever, scrub typhus, trench fever, brucellosis, bovine tuberculosis, Lyme disease, leptospirosis, salmonellosis, and spotted fever (Venkatesh *et al*., 2016).

Small mammals carry different species of both Gram negative and positive bacteria that they could transmit to a susceptible host through bites or handling of infected carcasses, inhalation of the bacteria, or through eating or drinking food or water that is contaminated by infected feces or urine. Bacterial zoonoses of public health importance that are transmitted by these small mammals include, among others tularemia, salmonellosis, leptospirosis, rat bite fever, rickettsiosis, bartonellosis, Q-fever and plague(Hornok *et al*., 2015). Human infections with

thèse Bacteria zoonoses are easily treatable with antibiotics if diagnosed early. However with the currently global threat of antimicrobial resistance, even easily treatable infections can become life threatening.

2.9 Transmission of zoonotic infections

Worldwide, rats, mice and shrews spread over 60infectious diseases either directly or indirectly(Guterres & de Lemos, 2018).Small mammals may disseminate infectious agents to humans by serving as reservoirs (or carriers) of the infectious agents in nature (Golden *et al*., 2015). Transmission of infections such as hantaviruses, leptospirosis may happen from small mammals to humans through contact with feces, urine, or saliva or through bites(Lõhmus *et al.*, 2013).

Along with other mammals, small mammals can function as short-term carriers that amplify infectious agents which are in turn spread to humans by intermediate predominantly arthropod vectors. Arthropod vectors that can transmit zoonotic infections include fleas, ticks and sand flies*(*Golden *et al*., 2015).

2.10 Detection of bacterial pathogens

Numerous methods including antibody and antigen detection assays that include; Indirect fluorescent antibody(IFA), enzyme-linked immunosorbent assay (ELISA) and Western blotting, culture isolation and PCR based assays for the detection of bacterial pathogens have been described(Ndao, 2009). PCR assays, especially the Real-time PCR, have been shown to be specific and sensitive overcomes the challenges of culture isolation, serological and conventional PCR assays (ref Courtney).Matrix-Assisted Laser Desorption Ionization-Time of flight Mass Spectrometry **(**MALDI-TOF MS) has also been used in the identification of bacterial species.

The technique uses protein mass to charge values obtained from mass spectral peaks to generate path genomic sequences that can be used for identification of a particular species (Croxatto *et al*., , 2012).Despite the fact that MALDI-TOF is a high throughput, accurate and sensitive method in characterizing a wide array of viruses and bacteria, the identification of these pathogens is limited by database for sequencing of genetic markers, its time consuming and above all, is very expensive compared to other methods.

2.11 Next Generation Sequencing (NGS)

Most of molecular and microbiology techniques used in detection of microbes are limited by targeting only known pathogens. However, the use of next generation sequencing (NGS) allows unbiased detection of pathogens in a sample in a rapid and highly parallel sequencing approach (Motro & Moran-Gilad, 2017). Targeted NGS metagenomics has been used to profile microbial communities in samples using gene markers and microbe databases. Detection of bacterial microbes involves amplification of 16S rRNA gene which has conserved and variable regions, sample bar-coding by PCR, parallel deep sequencing and bioinformatics sequence analysis (Rosselli *et al*., 2016; Couper & Swei, 2018).

16S rRNA is a universal gene found in all bacteria and has a sequence of approximately 1550 bp, large enough for bio-informatics purposes(Chen *et al*.,1989;Janda and Abbott 2007). The 16S rRNA amplicon sequencing technique is based on the amplification of one or more of the hyper variable regions of the 16S rRNA genes. The amplicon library is then sequenced and compared with reference sequences in curated databases for taxonomic identification(Galan *et al*., 2016).

A schematic of the 16S rRNA gene is shown in **Figure 2**. The gene consists of both conserved and variable regions. The conserved regions make universal amplification possible, while sequencing the variable regions allows discrimination between different microorganisms such as bacteria, archaea and microbial eukaryotes. Universal PCR primers have been designed to target the conserved regions of 16S rRNA gene, making it possible to amplify the gene in a wide range of different microorganisms from a single sample.

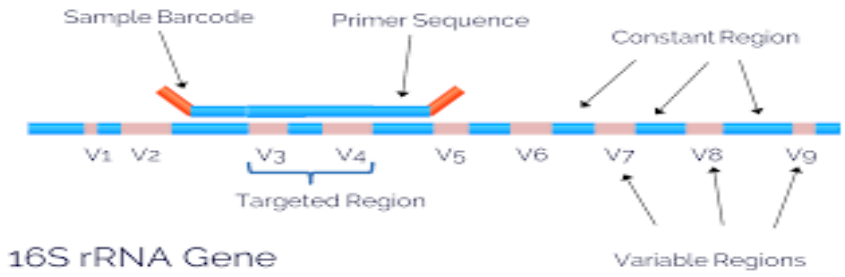

Figure 2A representation diagram showing the 16S rRNA (1500-bp gene).

Blue areas represent conserved regions that serve as gene targets for PCR amplification and DNA sequencing of bacteria. Targeted regions are the V3 V4 regions (Putignani & Conti, 2015)

CHAPTER THREE

MATERIALS AND METHODS

3.1 Study design and Study sites

This was a prospective surveillance study that used already archived samples from a previous study by (Masakhwe et al., 2018a) that evaluated bacterial diversity in wild captured rodents and shrews. The specimens used came from wild caught small mammals that were caught in Marigat County in June 2017. Climate in Marigat County is arid to semi-arid located approximately 260 km north-west of Nairobi and covers an area of 1,514.9 km^2. It is located between latitude 0°49'60N of equator and latitude 36°28'60E. The County has a population of 80,274 as per the 2009 national census with annual population growth rate of 1.93% (Kiarie-makara, 2016).

The study area was selected for surveillance following an earlier report of high sero-prevalence of antibodies against scrub typhus group antigens in febrile patients reporting to Marigat District Hospital (Thiga*et al*., 2015).Five rodent traps were installed which included cropped fields, grass fields, gardens, orchards, and around buildings(site 1 to 5) were selected as shown in Figure (**See Figure 3**).

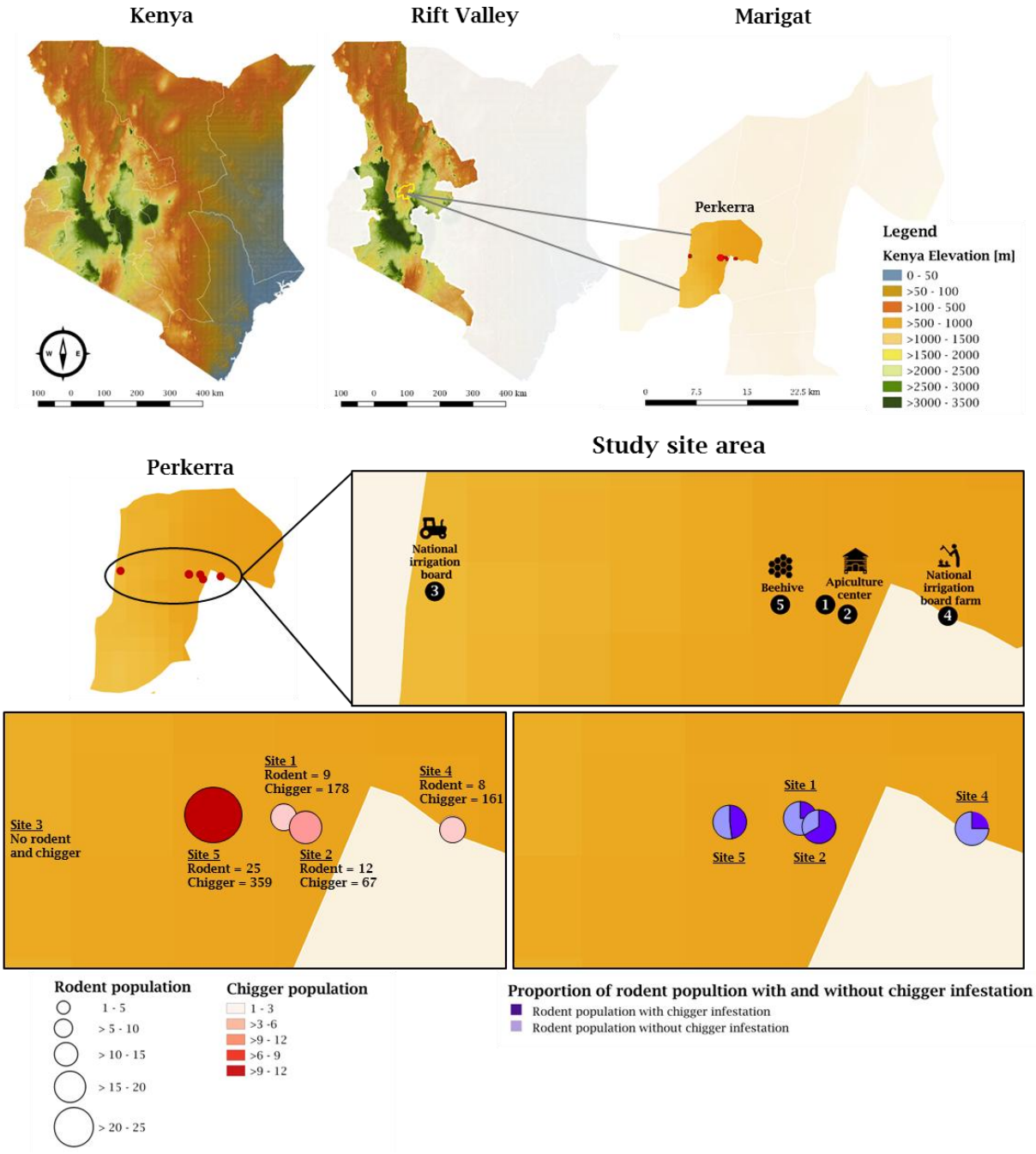

Figure 3 Rodent collection sites

Map of Kenya showing the locations of the 5 trapping sites (marked numbers 1-5) in Marigat, Baringo County.

3.2 Ethical statement

Archived spleen tissues from a previous study were used (Masakhwe *et al*., 2018). The animal use protocol for this study was reviewed and approved by Walter Reed Army Institute of Research under protocol number AP-12-001, KEMRI IACUC #2208, and National Museums of Kenya NMK/SCom2013/08.

3.3 Sample collection

This study used a non-probabilistic purposive sampling technique to determine the study sample size. The technique was based on the subjective ruling rather than random selection. The captured rodents belonged to two orders, order Rodentia and *Eulipotyphyla*. Animals from Order *Rodentia*, accounted for 70.4% of the total catch while *Eulipotyphyla* accounted for 29.6%.

The wild rodents and shrews were trapped in Marigat (**See Fig. 3**) between 8th and 15th June 2017.We used archived rodent tissue (spleen) from fifty-four small mammals. The spleen is a peripheral lymphoid organ in vertebrates that aids in filtering blood. It plays an important role in the modulation of immune responses and hematopoiesis. The spleen can be infected by bacteria during the blood filtration and can therefore be used as an indicator for microorganisms harbored by wild animals

Marigat is located in the lowlands of Baringo County in the Rift Valley. A total of 64 to 67 Sherman collapsible rodent traps per site were set up each night in a variety of habitats, including cropped fields, grass fields, gardens, orchards, and around buildings. Traps were placed in a grid format every 10 m or a transect, according to the land topography. Rodent traps were baited with appropriate foods, such as green nuts mixed with peanut butter, fried potato chips, sausages, or

other local foods. Geographic coordinates of the sampling sites were recorded using a handheld GPS tracker (Garmin, Olathe, KS) and used to locate trapping sites on the map (**See figure 3**). Rodent traps were set before sunset, and trapped rodents were collected early the following morning before 1000h or checked two to three times throughout the daytime. The wild caught small mammals were taxonomically identified using their morphological characteristics and found to belong to 4 species including *Acomys Wilsoni* (10), *Mastomys spp* (20),*Crocidura spp*(16), *and Rattus rattus*(8). (Appendix 1)

3.4 Sample Processing

3.4.1 Extraction of genomic DNA from the spleen tissue

Total genomic DNA was extracted from the spleen using the QIAamp DNA Mini Kit (Qiagen, CA, USA) as recommended by the manufacturer. Briefly, 10 mg of spleen biopsy was added into a 1.5 mL micro-centrifuge tube containing no more than 80 µL phosphate buffered saline (PBS). The sample was then homogenized using the Fast Prep- 24 TM homogenizer sample preparation system (MP Biomedical, LLC, Sanata Ana California, USA) after which 100 µL of tissue lysis buffer (ATL buffer) supplied with the kit was added to the homogenate. 20 µL of proteinase K was then added and mixed by vortexing followed by incubating the sample at 56^{o} C for 2 hours until the tissue was completely de-proteinized. The micro-centrifuge tube was briefly centrifuged to remove moisture drops on the lid before adding 4 µL of RNase A (100 mg/mL) in order to remove the RNA. The sample was then mixed by pulse-vortexing for 15 s, followed by incubation for 2 min at room temperature. The samples were briefly centrifuged and 200 µL of buffer AL added and then mixed by pulse-vortexing for 15 s followed by incubation at 70^{0}Cfor 10 min. 200 µL of 90-100% ethanol was added and mixed by pulse-vortexing for 15 s then briefly centrifuged to remove drops from inside the lid. The mixture including the DNA

precipitate was then transferred to the QIAamp Mini spin Colum and centrifuged at 8000 rpm for 1 min. The Mini spin column was placed in a clean 2 mL collection tube and the tube containing the filtrate discarded. 500 µL wash buffer (AW1) was added without wetting the rim and centrifuged at 8000 rpm for 1 min and the collection tube changed. 500 µL of wash buffer 2 (AW2) was carefully added without wetting the rim and centrifuged at 14,000 rpm for 3 min. The spin columns were placed on to clean collection tubes and 100µL elution buffer (AE) was added and then incubated at room temperature for 5 min followed by centrifugation at 6000 x g for 1 min. 50 µL of the DNA was aliquoted into a new tube and archived in -80^{0}C. The other 50 µL was then utilized for molecular bar-coding of the small mammals, as well as for pathogen detection in the mammalian spleens.

3.4.2 PCR amplification and sequencing of Cytochrome C oxidase subunit 1 *(COI gene)*

Genomic DNA extracted from small mammals' spleens was used as template for polymerase chain reaction amplification (PCR) and sequencing of 649-bp fragment of the *CO1* gene using the universal primer pairs as shown in Table 1(Nicolas *et al*., 2012). PCR was performed in a 25 µl reaction volume containing 0.25 µl of 2X MyTaq Red Mix, 5 µl Mytaq buffer (Bioline, UK), 10µM of each primer 16.5 µl of Rnase free water and 2 µl of DNA template. PCR amplification was done on an Eppendorf Mastercycler pro 384 (Eppendorf, USA) with an initial denaturation step of 4 min at 94^{0}Cfollowed by 35 cycles of 94^{0}Cfor 30 seconds, annealing at 48^{0}C for 30 seconds and extension at 72^{0}Cfor 1 minutes and a final extension step at 72^{0}Cfor 10 minutes. The amplification was then visualized on a 2% agarose gel (Thermo Fisher Scientific,Canada, USA) stained with GelRed (Biotium, Australia). PCR products from positive samples were purified using AgencourtAMPure XP Beads (Beckman Coulter, CA, USA) as per manufactures instructions. The purified PCR products were sequenced in both directions using the Big Dye

Terminator Cycle Sequencing Kit v 3.1 (Applied Biosystems, CA, USA) and the sequences analyzed by capillary electrophoresis in a 3130 Genetic Analyzer (Applied Biosystems).

3.4.3 PCR amplification and sequencing of Cytochrome B Gene (Cytb)

The *Cytb* gene was amplified using PCR forward primers L14723 and reverse primer H15915 (Table 1). Where amplification could not be achieved with the primary PCR primers, the internal primers L14749 and H14896 were used (Nicolas *et al*., 2012).PCR was performed in a 25 µLreaction volume containing 2 µL of DNA template,0.2 µM of each primer,0.25 µL, MyTaq polymerase (5 U/µL), and 5 µLx of 5X MyTaq buffer(Bioline, UK).Amplification was performed on an Eppendorf Mastercycler pro 384 (Eppendorf, USA) with an initial denaturation step of 94^0C for 2 min, 35 cycles of denaturation at 94^0C for 30sec, annealing at 52^0C for 30 sec, and 72^0C for 1 min followed by a final extension step of 72^0C for 10 min. The amplification products were visualized on a 2% agarose gel (Thermo Fisher Scientific, Canada, USA), stained with GelRed (Biotium, Australia). The purified PCR products were sequenced in both directions using the Big Dye Terminator Cycle Sequencing Kit v 3.1 (Applied Biosystems, CA, USA) and the sequences analyzed by capillary electrophoresis in a 3130 Genetic Analyzer (Applied Biosystems).

Table 1Primers used for Amplification and Sequencing of *Cyt b, CO1* and 16S rRNA

Gene target	Primer	Primer sequence (5' - 3')	Fragment size(bp)	Origin
Cytb	*# L14723	(CCAATGACATGAAAAATCATCGTT)	1140bp	(Nicolas *et al*., 2012)
	*# H15915	(TCTCCATTTCTGGTTTACAAGAC)		
	*# L14749	(ACGAAACAGGCTCTAATAA)		
	# H14896	* (TAGTTGTCGGGGTCTCCTA)		
Co1	*# BatL5310	(CCTACTCRGCCATTTTACCTATG)	648bp	(Nicolas *et al*., 2012)
	*# R6036R	(ACTTCTGGGTGTCCAAAGAATCA		
16s Rrna		*# TCGTCGGCAGCGTCAGATGTGTATAAG AGACAGCCTACGGGNGGCWGCAG	460bp	(Klindworth*et al*., 2013)
		*# GTCTCGTGGGCTCGGAGATGTGTATAA GAGACAGGACTACHVGGGTATCTAATCC		

R = Reverse primer; *= primary PCR; #= secondary PCR and sequencing L=Low H=High

3.5 Unbiased bacterial identification by 16S rRNA deep sequencing

The presence of bacteria in the rodent spleen genomic DNA were detected by PCR amplification using Illumina barcode tagged primers targeting V3 – V4 hyper variable region of the 16S rRNA gene as described earlier (Klindworth *et al*., 2013). Amplification was carried out in a 25 µL reaction volume consisting of 2×NEBNext PCR master mix (New England Biolabs, MA, USA), 0.2 µM of each primer and 2.5 µL of DNA template. PCR was then performed on the Eppendorf Master-cycler pro 384 (Eppendorf, Humberg, Germany) with an initial denaturation at 95^0Cfor 3 min followed by 25 cycles of 95^0Cfor 30 sec, 55^0Cfor 30 sec and 72^0Cfor 30 sec and a final extension of 72^0Cfor 5 min.

Dual indexing to allow multiplexing of samples was done using 5 µL of purified amplicons, 5 µL of Nextera XT i7 Index Primer, 1.5 µL of Nextera XT Index Primer 2 (Illumina), 25 µL of NEBNext High-Fidelity 2× PCR Master Mix (New Englands BioLabs) and 10 µL of PCR grade water (Thermo Fisher Scientific), with thermocycling at 95^0Cfor 3 min, followed by 12 cycles of 95^0Cfor 30 s, 55^0Cfor 30 s, and 72^0Cfor 30 s, and a final extension at 72^0Cfor 5 min. Constructed 16S sRNA amplicon libraries were purified with AgencourtAMPure XP beads (Beckman Coulter Genomics). The libraries were quantified on Qubit Flourometer 2.0 using Qubit dsDNA HS Assay kit (ThermoFsher Scientific). Libraries were normalized and pooled to 4 nM based on Qubit values. Pooled samples were denatured and diluted to a final concentration of 12 pM and spiked with 5% PhiX (Illumina) sequencing control for accurate base calling. All samples were multiplexed in one sequencing run and sequenced using MiSeq Reagent Kit V3 in the Illumina MiSeq System. DNA extracted from R-Nase free water was also sequenced in the same metagenomic pipeline as a negative control sample so as to allow empirical assessment of the contamination background

3.6 Bacterial 16S rRNA metagenomics data analysis

Miseq reporter software version 2.6.3 (Illumina) was used to de-multiplexed sequences and trim sequencing adapters. Mothur version 1.35 was used for read pairing, quality filtering, chimera removal and operational taxonomic units (OTU) clustering. Taxonomic assignment of OTUs was done by mapping the filtered sequences using the naïve bayes classifier against the SILVA (SSU123) 16S rRNA database classification. To normalize the dataset, sequences with less than 1241 of the total number of sequences were discarded before downstream analysis. Alpha diversity indices including Shannon diversity index (which accounts for both taxa abundance and

evenness), number of observed OTUs (species richness) were calculated in R versionr-3.5.1 (2009-2017 R Studio, Inc.)(R Team, 2017).

Non-metric multidimensional scaling (NMDS) with Bray-Curtis similarity matrix values were computed using the *phyloseq* package to test whether microbial communities differed across the small mammals' species and the study sites.

3.7 Small mammal's taxonomic classification data Analysis

Forward and Reverse nucleotide sequences from the small mammals were quality checked and assembled into contigs using CLC MAIN WORK BENCH version 7.5. Small mammals' species identification was done by querying the contigs against the refseq database using the blast algorithm(www.ncbi.nlm.nih.gov).Phylogenetic analysis of the small mammals was then done using MEGA version 7 (Kumar, Stecher, and Tamura 2015).

To determine their phylogenetic placements, corresponding to the four mammalian species determined by gross morphology i.e. *Acomys Wilsoni* (10), *Mastomys spp* (20),*Crocidura spp*(20),*and Rattus rattus*(10),the reference sequences downloaded from Genebank (www.ncbi.nlm.nih.gov) MEGA version 7(Kumar *et al*., 2016)was used to determine the best substitution model, and the best model was used to infer phylogenetic trees using both Neighbor joining and Maximum Likelihood treeing methods.

CHAPTER FOUR

RESULTS

4.1 The species of wild small mammals captured from Marigat, Baringo County

A total of 54 spleen tissues from wild caught small mammals were used for molecular taxonomy and microbiome profiling. The small mammals had previously been classified based on morphological characteristics into two major orders, Rodentia and Eulipotyphyla. The order Rodentia (70.4%) had the following species: *Acomys wilsoni* (n=21), *Lophuromyssikapusi* (n=1), *Mastomys natalensis* (n=8), *Rattus rattus* (n=2), *Arvicanthis niloticus* (n=6). The order Eulipotyphyla (29.6%): *Crociduraolivieri* (n=2), and *Crocidura spp* (n=14) that could not be speciated.

Table 2Summary of the diversity of small mammals trapped for the surveillance study.

Order	Family	Rodent species
Rodentia	*Muridae*	*Acomys wilsoni* (n=21) *
		Lophuromyssikapusi(n= 1) *
		Mastomys natalensis(n=8) *
		Rattus rattus(n= 2) *
Rodentia	*Murinae*	*Arvicanthis Niloticus*(n= 6) *
Eulipotyphla	*Soricidae*	*Crocidura spp* (n= 14) *
		CrociduraOlivieri(n=2) *
2 orders	3 families	Total (N=54) *

(n=x) *represent the total number of species of the small mammals caught.

4.2 Small mammal's identification using *CO1* gene

DNA from 54 spleen samples of the small mammals were analyzed for molecular taxonomy using the *COI* gene. Due to low sequence quality, 17/54(31.5%) sequences were not included in the analysis and therefore only 37 samples were analyzed. By *COI* gene, the small mammals were classified into two orders, *Rodentia* and *Eulipotyphla*. Within the order *Rodentia*, 4 genera were determined: *Acomys clade*51%(19/37), *Mastomys clade*16%(6/37), *Rattus.spp*4%(2/54), Arvicanthis *niloticus*13%(7/54). Within the order *Eulipotyphla*only C*rocidura somalica* 4%(2/54)(**Fig 4.**) was identified.

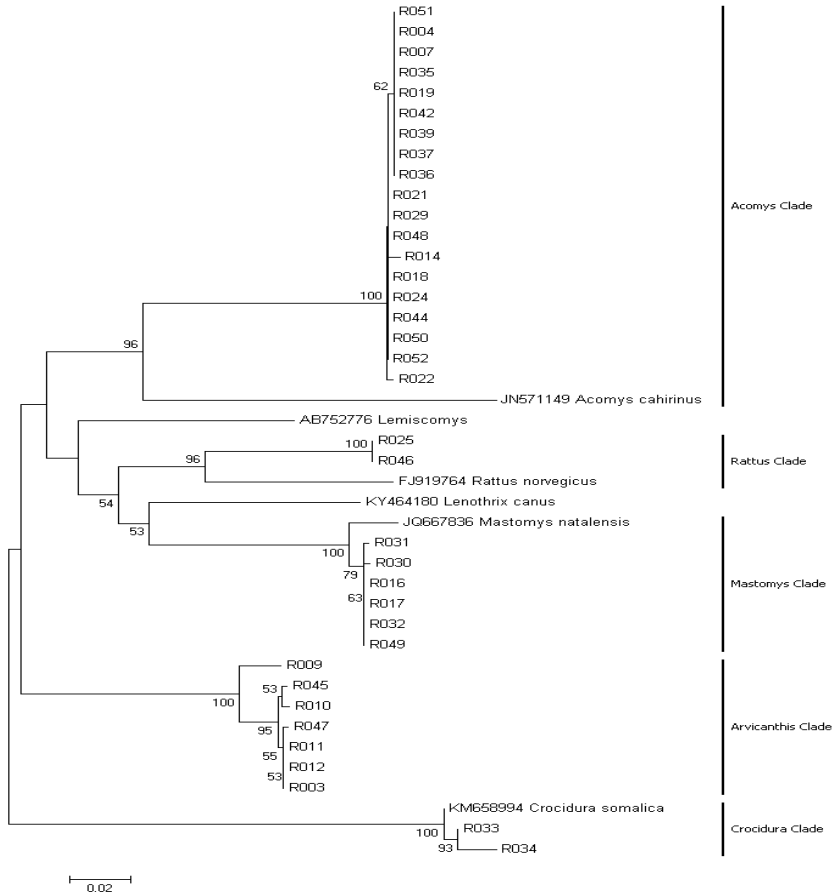

Figure 4 Evolutionary relationships of small mammal taxa using the COI gene.

COI gene tree based on complete amino acid sequences of the wild caught small mammals and those obtained from Genbank database (Accession numbers).Only bootstrap values>50% are shown. Samples from this study are named as "R000".The scale bar indicates 0.02 changes per nucleotide position (evolutionary distance).

4.3 Small rodent identification using *Cyt b* gene

The small mammals had previously been classified based on morphological characteristics into two major orders, Rodentia and Eulipotyphyla. The order Rodentia contained 5 species: *Acomys wilsoni*, *Lophuromyssikapusi*, *Mastomys natalensis*, *Rattus rattus* and *Arvicanthis niloticus* while order Eulipo typhyla had two species: *Crociduraolivieri*, and un-speciatable *Crocidura spp*.

By the *Cytb*gene, 85% (46/54) samples generated usable sequences. Homology searches with BLASTn against the GenBank database classified the small mammals into two main orders: Rodentia(77.6%) comprising *Acomys*46%(21/46), *Mastomys 17%(*8/46), *Arvicanthis*13%(6/46) and *Rattus*7%(3/46), and Eulipotyphyla (22.4%), comprising only of *Crocidura*24%(11/46).

As shown in Figure 4, five well supported clades (*Crocidura*, *Acomys*, *Rattus*, *Arvicanthis* and *Mastomys*) were identified. All *Crocidura*24%(11/46) clustered with *Crocidura somalica* in a well-supported clade (99%). 43.5% (20/46) samples clustered in the *Acomys* clade and branched with *Acomys wilsoni* (100%). 6.5% (3/46) samples clustered with the *Rattus* clade and branched with *Rattus rattus*. 13% (6/46) samples clustered with the *Arvicanthis* clade and branched with *A. niloticus*. 13% (6/46) samples grouped with *Mastomys natalensis* in the *Mastomys* Clade.

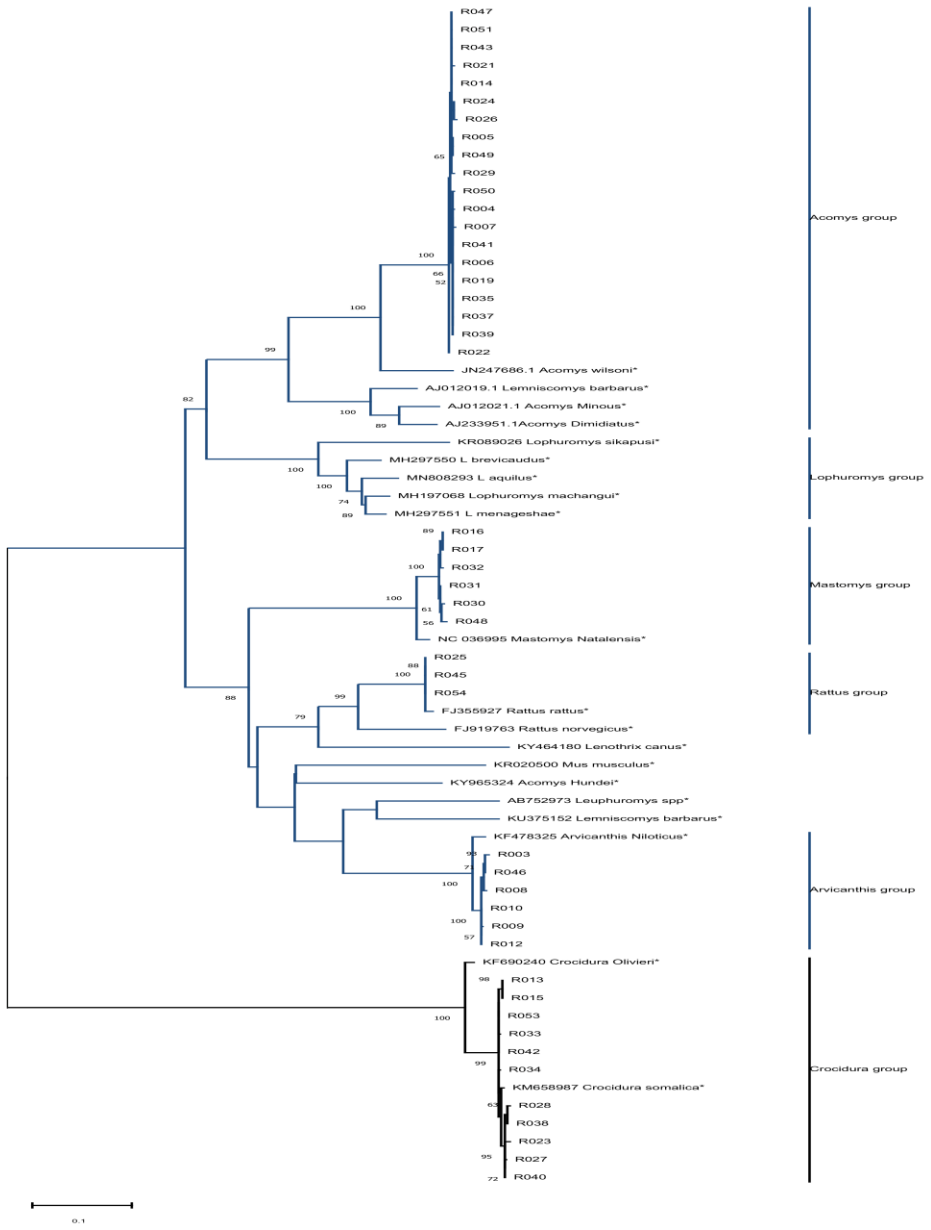

Figure 5 Evolutionary relationships of small mammals using the cytb gene.

46 study samples (all have prefix "R" and 20 reference isolates from genbank marked with an *) were used to infer the Maximum Likelihood tree. Blue lines represent members of order Rodentia, black lines represent Eulipotyphla. Numbers on the branches represent bootstrap support values >50%. The branch lengths represent the number of base substitutions per site.

4.4 Concatenated phylogeny by CO1 Cytb genes

In order to improve the phylogenetic resolution of individual genes, the two genes were concatenated. Out of the 54 samples initially available for analysis, only 30 were included in the concatenation. The choice of the 30 was influenced by availability of both genes (Appendix 1).The concatenated sequences also included validated species available in GenBank. A concatenated tree constructed with Bayesian method is shown in Fig 6.The observed species were *Arvicanthis niloticus*, *Crocidura somalica*, Acomys wilsoni, *MatomysNatalensis* and *Rattus rattus.*

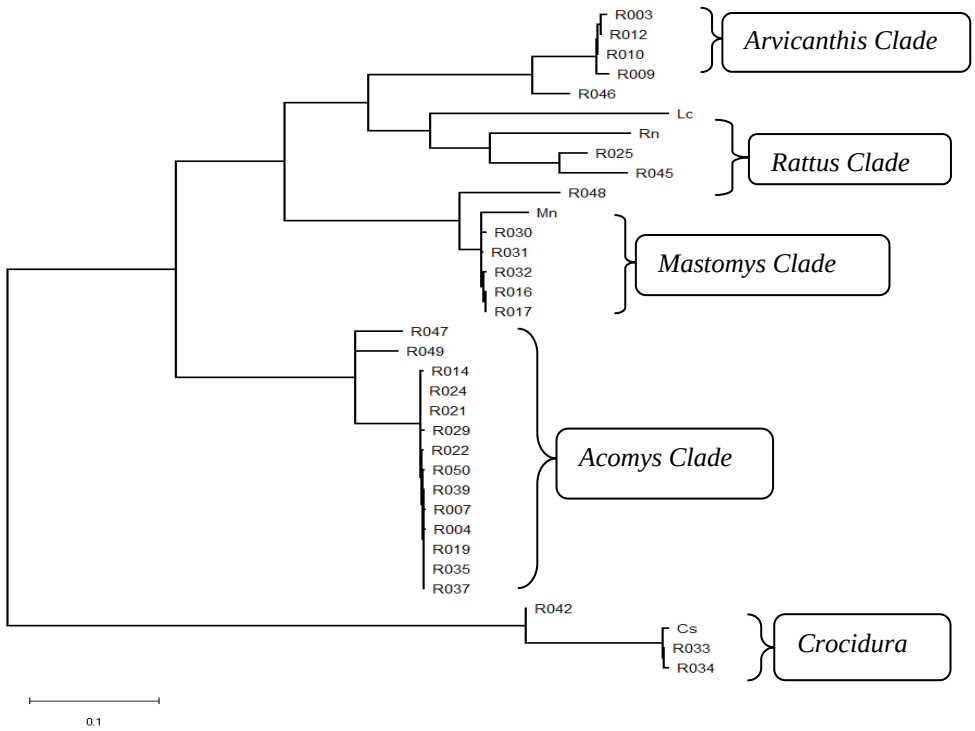

Figure 6 Phylogeny of wild caught small mammals from Marigat, Kenya.

Maximum Likelihood trees were obtained *COI* and *Cytb* nucleotide sequences. Only those with bootstrap values >50% are shown. Samples from this study are named as "R000".The scale bar indicates 0.02 changes per nucleotide position (evolutionary distance).

From both the *CO1* and *Cyt b*gene trees, and from the concatenated tree, the samples were seen to cluster into 5 major grouping involving, *Acomys Wilsoni, Arvicanthis niloticus, Rattus rattus, Mastomys natalensis and Crocidurasomalica.*

However, a few sequences including R003, R042, and R043 R047 and R054 had some discrepancies where by at least two of the methods were not agreeing to one species with both the morphological classification and one molecular marker.

4.5 Bacterial biota identified by 16S rRNA sequencing

16S rRNA amplicon libraries were prepared and sequenced on a MiSeq platform. A total of 3,111,804 raw sequences were generated from paired end reads. After quality filtering, removal of chimera reads and non-bacterial sequences,175,629 sequences were available for analysis. An average of 1241 sequences was used to infer indices of bacterial richness and diversity based on OTUS estimated through a rarefaction curve. This was used to give an adequate depth of coverage (Figure 7).

From the rarefaction curves (*Figure 7*), only samples with $\geq$1,241 sequences were considered adequate for inferring bacterial richness and diversity. Based on this cut off of rarefaction, the sequences clustered into 213 unique operational taxonomic units (OTUs) at 97% sequence similarity (Appendix 2).

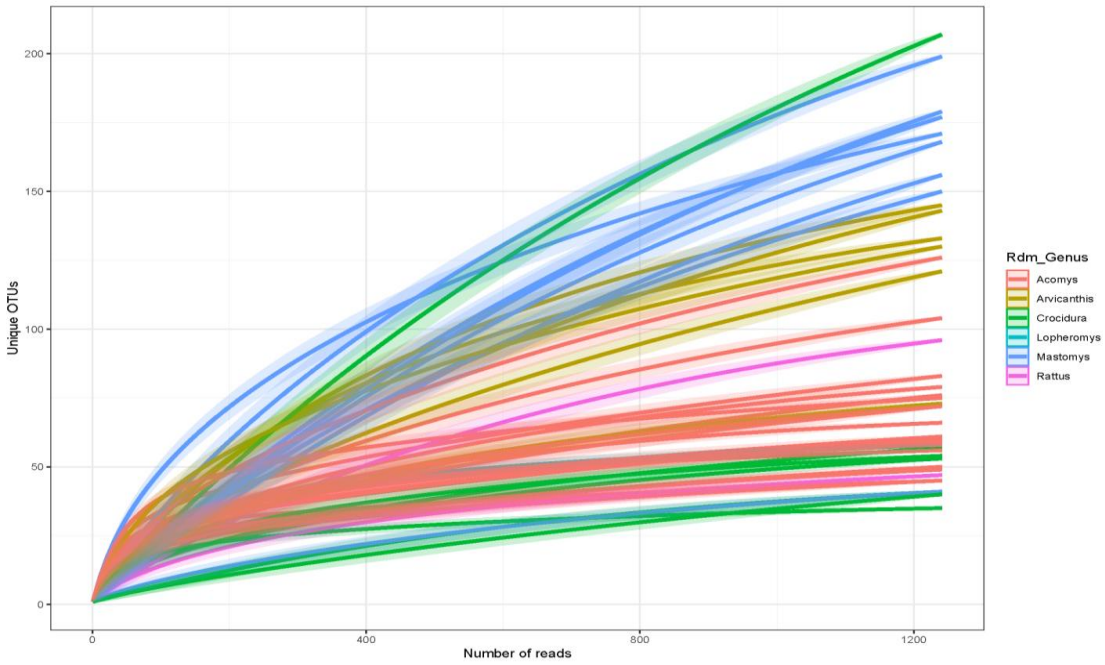

Figure 7 Rarefactioncurvesusedfor sequence count normalization.

A sequencing depth was selected at the point where the curves form a plateau so as to capture the full diversity of OTUs. In the rarefaction curves shown, only samples with an average of >1,241 sequences were considered adequate for inferring indices of bacterial richness and diversity.

4.6 Diversity of bacteria in the small mammals

Of the 54 spleen samples examined, 10 failed to meet the set quality sequence threshold and were therefore not included in downstream analysis. The remaining 44 samples yielded 8,837,475 raw sequence contigs. After quality filtering, collapsing duplicate sequences, removing chimeras and non-bacterial sequences,2,778,822 sequences were considered suitable for further analysis. On querying the SILVA rRNA database, the sequences were grouped into 196 unique operational taxonomic units (OTUs) at 97% sequence similarity.

The taxonomic assignment to phyla and genera for the 196 OTUs is shown in Figure 8 bacteria phyla were detected, with Proteobacteria being the single most abundant phylum and contributed

by 33,752 of 52,146 total contigs (64.7%). Other phyla included Actinobacteria at 18.0% of the sequence abundance, Firmicutes(6.1%), Chlamydiae (3.8%), Chloroflexi (2.6%) and Bacteroidetes (1.9%).Less abundant phyla (<1% of the sequence abundance) included Acidobacteria, Verrumicroba, Planctomycetes, Fusobacteria, Deinococcus-Thermus, Armatimonadetes, Gemmatimonadetes, TM7, OD1, Spirochaetes, SR1 and Tenericutes

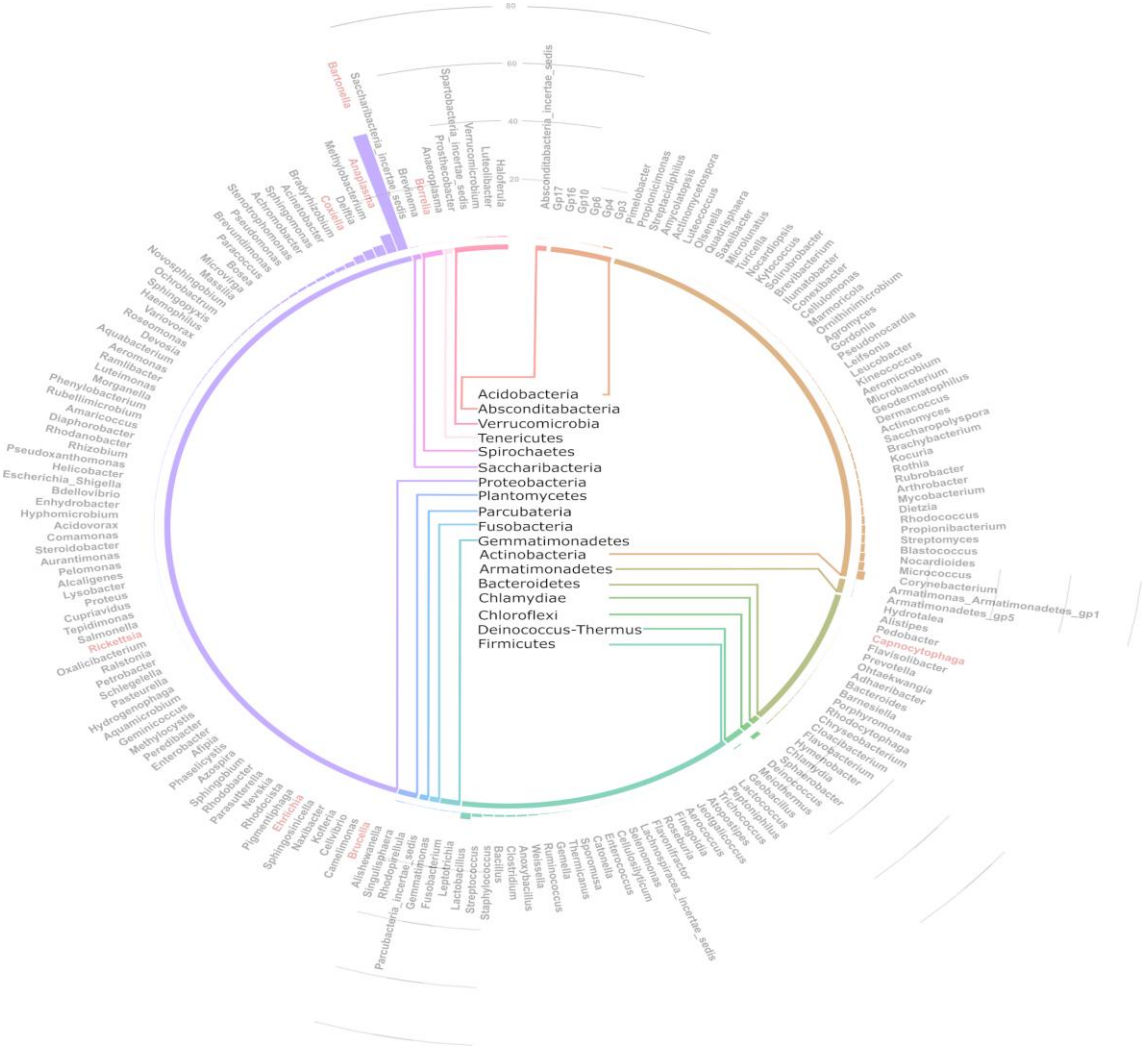

Figure 8 Circular bar plot showing taxonomic assignment for the 196 OTUs in the spleen samples from small animals of the bacteria genera identified, Bartonella was the most abundant and contributed 41.5% of the total contigs.

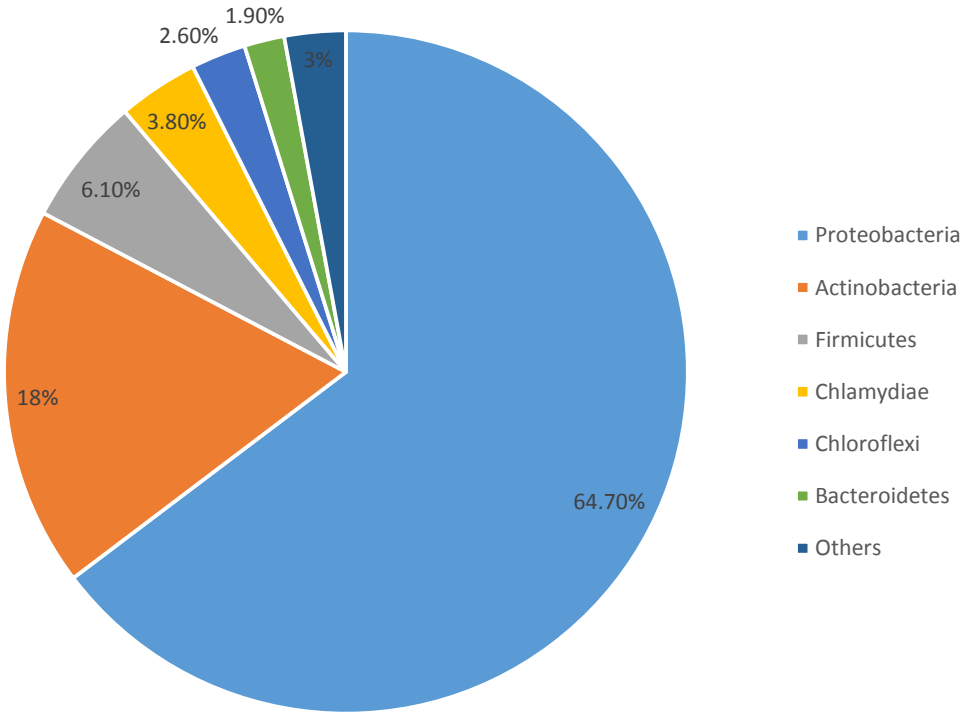

Figure 9color-coded pie chart showing the abundance of bacteria at phylum level present in the small mammals.

4.7 Bacterial pathogens recovered in the different genera of wild caught small mammals

At the genus level, nine major bacterial pathogens were identified in more than 95% of the 44 samples that met the sequence quality threshold. Of these, *Bartonella* was the most abundant at a prevalence rate of 41.5%. *Anaplasma* was the second most abundant at 6.5%,*Delftia* was observed at an abundance of 3.2% (Fig 9).Delftia *spp* has been known to be a nonpathogenic environmental organism and is seldom clinically significant. However, *D acidovorans* infection has been reported to occur in hospitalized or immune-compromised patients (Bilgin *et al*., 2015). *Coxiella* was also detected at an abundance of 2.6%. In particular, Coxiella was highly abundant in one of the *Rattus rattus* at abundance of 99.03%.

Lactobacillus, Corynobacterium, Micrococcus, and *Bradyrhizobium* were detected although in low abundance of *2*% each. Other relatively less abundant bacteria genera but worth noting include *Rickettsia, Acinetobater, Leptospira, Borrelia,* and *Brucella, Chlamydia and Streptococcus* that accounted for36% (<1000 reads).

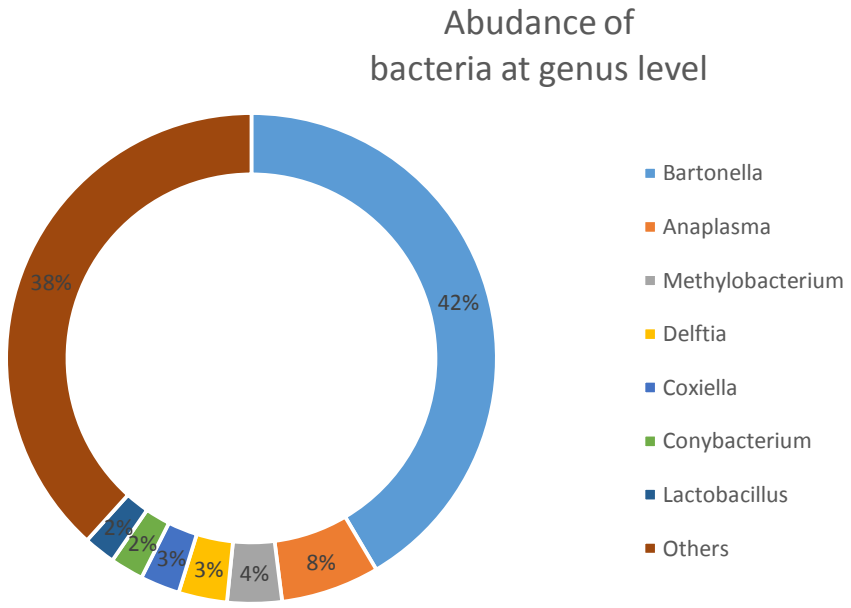

Figure 10A pie chart showing the abundance of bacteria at genus level in the spleen tissue of the small mammals

Thirty four out of forty-four(77.3%) small mammals tested positive for *Bartonella* genus with abundances ranging between 1 to 1223 reads. Of the34, samples, (47.1%) were recovered from *Acomys* .The small mammals of the genus *Crociduran* had only *Bartonella*.

On a heat map, *Bartonella* was seen to have more heat followed by *Anaplasma* and then *Coxiella* (Fig 11).

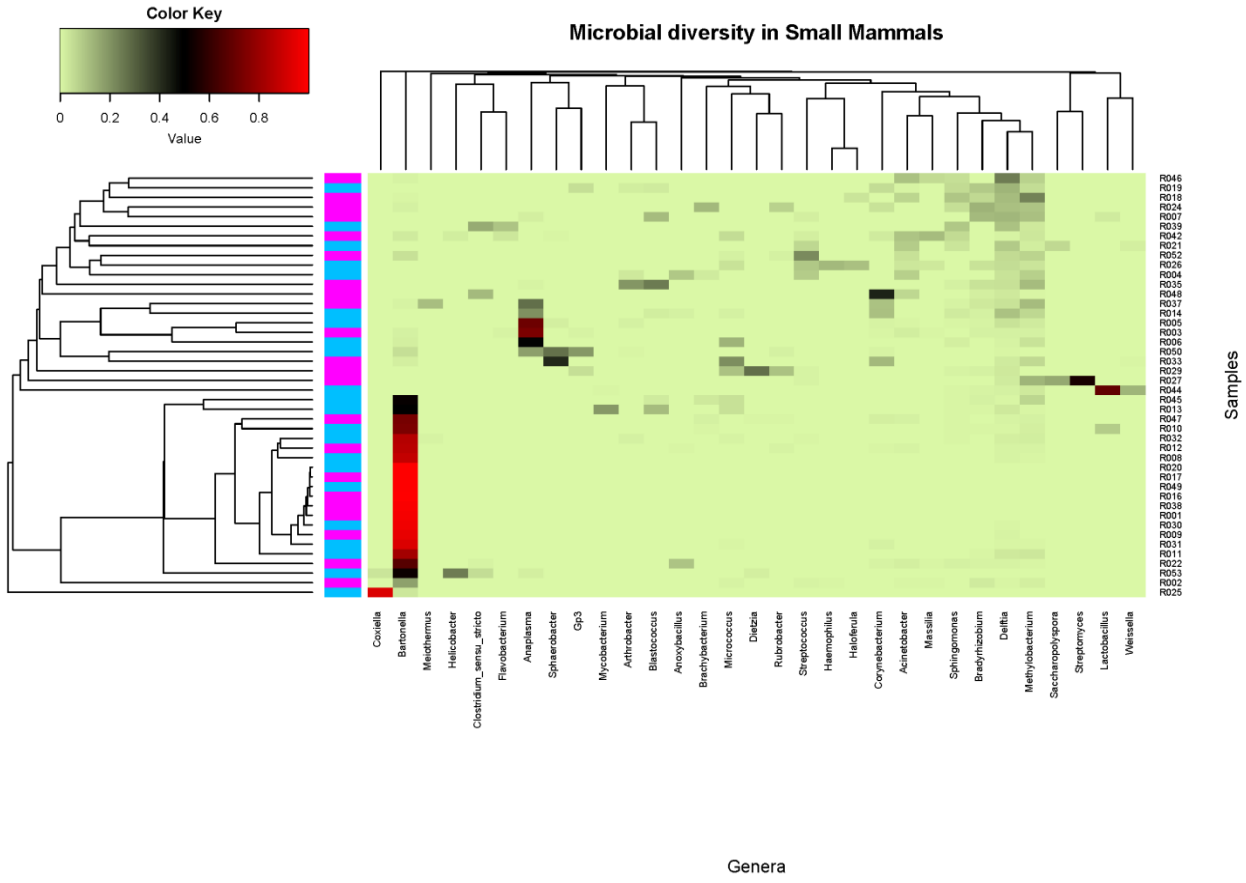

Figure 11 Heat map of the pathogenic spleen microbiome recovered from wild caught small mammals. x axis indicates the pathogenic genera while y axis indicates individual small mammal ID. The color scale (log 10%) is on the top left corner.

4.8 Alpha diversity of pathogenic bacteria in the wild caught small mammals

Alpha diversity indicates the species richness in each of the wild caught small mammals. From the Shannon diversity indices (Figure 12), *Acomys* had the highest diversity of 3.09, *Arvicathis* 2.41, *Crocidura* 2.18, *Lopheromys*2.13, *Mastomys* 1.75. *Rattus* 1.42 had the least.

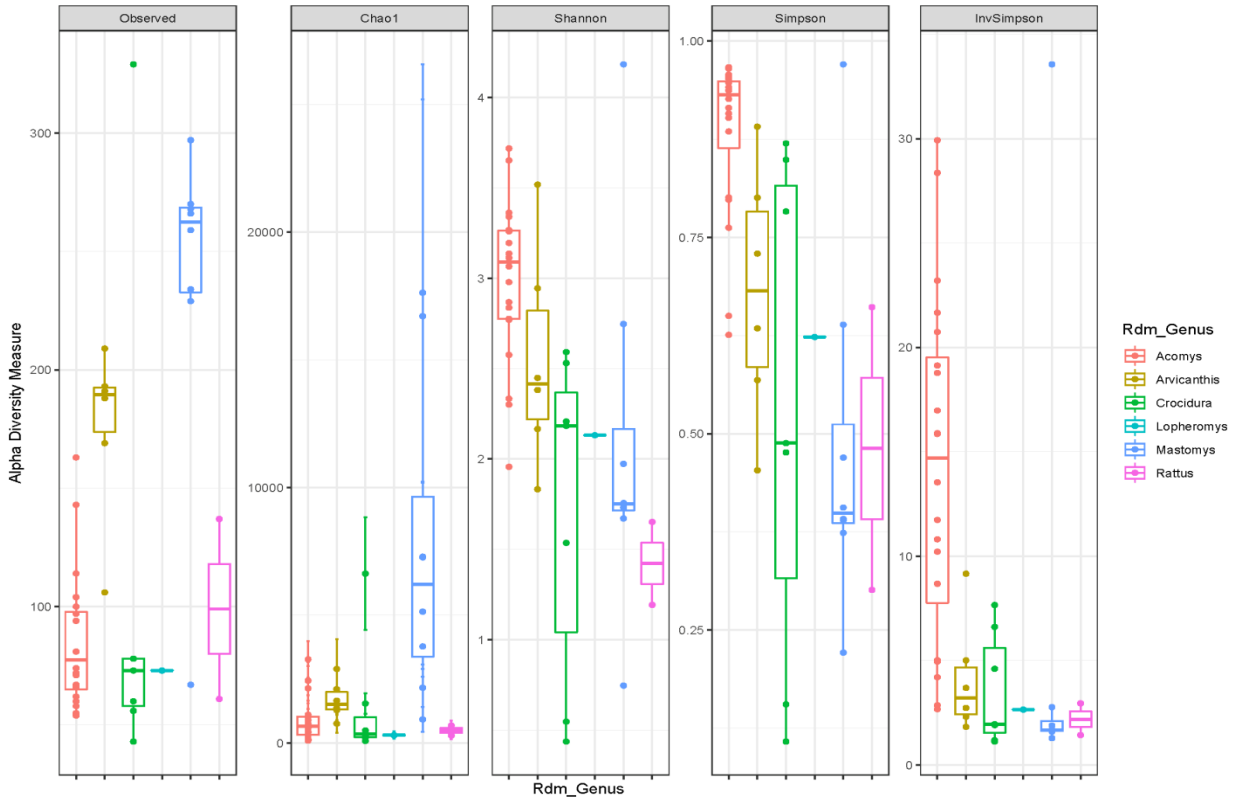

Figure 12 Alpha diversity indices. (Chao1 and Shannon) for rodent spleen bacteria biota. x-axis shows the Rodent genus while the y- axis shows the alpha diversity measures

We also identified 9/54 (17%) small mammals that were hyper-reservoirs. That is, these small mammals were carrying more than two zoonoses, the highest being *Crocidura* which harbored 4 pathogens namely *Anaplasma, Bartonella, Coxiella* and *Erhlrichia.*

CHATER FIVE

DISCUSSION

Globally, zoonotic diseases account for 15.8% of all deaths and 43.7% of deaths in low-resource countries. It is projected that zoonoses are responsible for 2.5 billion cases of human illness and 2.7 million human deaths worldwide each year (al., 2017).These diseases are a major public health problem that pose a threat not only to the well-being of animals and humans but also to global health security. These diseases present themselves mainly with fever.

In Kenya, the knowledge and information on zoonotic diseases is insufficient and limited with no attention to control and prevention programs. The present study aimed to characterize bacterial pathogens present in different small mammals as well as to identify what exact species they belonged to. The spleen is a peripheral lymphoid organ in vertebrates that acts as a blood filter. It plays an important role in the modulation of immune responses and hematopoiesis (Ge *et al*., 2018). For this reason, its role may change its niche and thus making it an ideal organ for these bacteria to thrive in.

The results obtained suggest the potential involvement of small mammals in the cycle of zoonotic diseases as they are one of the biggest reservoirs of the pathogens. We present a detailed profile of the bacterial biota of the small mammals' caught in Marigat, Kenya. To our knowledge this is the first time that bacterial profiles in rodents and shrews is being explored in Marigat Kenya and thus this has provided a wealth of information on the bacterial pathogens present in these rodents and shrews as well as the association of the different species with the pathogens. In addition, it has advanced our understanding on the pathogens present in these small wild mammals.

5.1 Taxonomic/Molecular Classification of Small mammals captured.

In this study, a combination of two genetic markers (CO1 and Cyt b) genes were used for molecular classification of the small mammal species using DNA samples obtained from the spleen. Six genera comprising of five (5) rodent and one (1) shrew genera were identified by phylogenetic analysis of the amino acids and sequences of the mitochondrial Cyt b and COI genes **(Fig 4 & 5**).

While most of the species were identified morphologically using the identification keys by (Taylor, 2010), *Crocidura spp*, *Acomys spp* and *Rattus spp* had similar characteristics which led the organisms in being named as either *Crocidura* while they were *Acomys spp* according to molecular classification or being named *Acomys spp* while they are *Crocidura spp.*

CO1 and Cytb gene were concatenated in order to improve the phylogenetic resolution of individual genes. Out of the 54 samples that were available for study, only 31 were included in the concatenation and the sequences were successfully aligned to the reference sequences available in the gene bank. The choice of the 31 was influenced by availability of both genes (Appendix 3). It was however noted that the CO1 gene did not give have a high resolution for speciation. The sequences generated by the CO1 gene were poor and some had a match of less than 95% when compared to the reference sequences on gene bank. This is in contrast to Hebert *et al*., 2003, where their CO1 sequences possessed a high level of diversity. This could be attributed to the primers used or degrade of the nucleic acid. The Cyt b gene on the other hand however, showed higher level of genetic divergence. This may be due to the fact that the gene is 1149 bp long compared to CO1 which is 650 bp long. This allows the Cyt b gene to offer more taxonomic information when sequenced.

In this study, we found that small mammals of the genus *Acomys spp*, *Mastomys spp*, and *Crocidura spp* are the dominant carriers of zoonotic pathogens with a detection rate of 46%, 56% and 35% respectively. Acomys wilsoni and *Crocidura spp* were seen to be the dormant species which harbored Bartonella spp. Prevalence of Anaplasma spp was seen to be highest in Acomys genus.

5.2 Pathogen Prevalence in small mammals

As a powerful sequencing method, the next generation sequencing was used to reveal the complexity of bacterial communities in humans and animals. A total of 213 bacteria were detected in the small mammals captured. Of the fifty-four samples collected, 44 were seen to be reservoirs harboring more than 80 zoonoses caused by bacteria. From previous findings, it is seen that rodents can be hyper-reservoirs carrying between 2 to 11 zoonoses (Han *et al*., 2015). This study depicts that of the forty small mammals, nine (15%) were seen to harbor between 3-4 bacterial genera.

In the current study, the two major bacterial phyla, among the taxa of tested wild rodents and shrews' samples were Proteobacteria (64.7%) and Actinobacteria (18.0%). Earlier reports on the similar studies have shown that microbiota of blood and the spleen from humans or rodents mainly consisted of *Proteobacteria* and sometimes *Firmicutes* as well (Ge *et al*., 2018). However, in this study the percentage of Actinobacteria (18.0%) was higher than that firmicutes (6.1%).

This is the first report on the detection of rodent borne pathogens detected from the spleen in Marigat Baringo County. 77% of the samples tested positive for Bartonella with the highest abundance being in a *Crocidura spp*. The infection rate of Bartonella was 41.5%, which was the

highest among the rodent-shrew borne bacteria detected. These results propose that the wild captured rodents and shrews serve as animal reservoirs of zoonoses that are of medical importance.It has also been observed on other investigations that the prevalence of Bartonella spp. infection can reach high rates in rodent populations (Kosoy *et al*., 2010).

This current study also suggests the presence of *Coxiella spp* in the small mammals found in Marigat, consistent with previous studies that have been done on vectors that have been collected on goats, and cows. (Koka*et al*., 2018; Njeru *et al*., 2016). Based on the analysis of the 16s rRNA, Coxiella spp was detected in 9 samples with abundance ranging from 1 to 1119. Specifically, we observed Rattus rattus having the highest abundance of 1119 reads.

Q fever, a human infection caused by *Coxiella burnetti* has been previously reported in semi-arid to arid regions(Marigat, Mai Mahiu, Ijara, Garissa, and Isiolo), Kenya. The prevalence of Q fever was reported at 12.1% in both livestock and human populations in seven provinces which are known for large nomadic pastoral communities (Koka *et al*., 2018) and the risk of Q fever transmission is allegedly higher in grazed animals.

The infection rate of *Bartonella* was 41.5%, which was the highest among the rodent-shrew borne bacteria detected. This may be as a result of changes in social, dietary or cultural mores and environmental changes. Molecular methods especially the use of quantitative polymerase chain reaction are currently being preferred as the diagnostic tests kits of the choice taking over culturing hence easier identification of this bacteria avoiding biasness(Nachum-biala & Harrus, 2015).Through the application of new sequencing technologies, it is possible to observe the occurrence patterns of individual populations upon which the structure of microbial communities is based.

Furthermore bacteria in the genera Anaplasma, Ehrlichia and Rickettsia of the order Rickettsiales, which can cause febrile illnesses in human beings as well as animals was observed. This establishes elevated incidences of the rodent-borne zoonoses in Marigat County, Kenya. Marigat is known for having a large population of pastoralists who practice large scale farming of traditional livestock keeping, and they freely interact with these small mammals in the different environments where they herd their animals. This therefore increases the chance of zoonoses being transferred either directly i.e. by consuming contaminated food or water or indirectly by vectors such as ticks, and fleas. Much more animals are often chronically infected, shedding bacteria in feaces, milk, urine and especially birth products of mammals (Ouvery & Odolakis, 2005).

Rodents of the genus Mastomys harbor more pathogens and shrews, (order Eulipotyphla) are seen to be hyper-reservoirs with more than 2 pathogens. This is to enable them to increase the capability of transmission of infections to a wide range of different hosts rather than other agents that have restricted specific transmission dynamics because of limited host ranges.

This study establishes elevated incidences of the vector-borne zoonoses in Marigat County, Kenya. Marigat is known for having a large population of pastoralists who practice large scale farming of traditional livestock keeping, and they freely interact with these small mammals in the different environments where they herd their cattle. This therefore increases the chance of zoonoses being transferred either directly i.e. by ingestion of raw or contaminated food, milk or water or indirectly by vectors such as ticks, and fleas. The aerosol route is the primary mode of human contamination with most of zoonosis diseases. Contamination by this aerosols may occur directly from parturient fluids of infected animals, which may contaminate newborn animals, placenta or wool (Cardeñosa et al., 2006).

Finally, according to Ge *et al*., (2018), Anaplasma, Elrichia, Rickettsia and Coxiella are vector borne bacteria and are therefore transmitted by ticks. Ticks are considered to be the second most important vectors of infectious agents globally. They transmit pathogens from wild animals (small mammals) to humans and the transmission can also be made through domestic animals. They are everywhere in our environment; however, their prevalence is greater in areas with warm and humid climates (Betancur *et al*., 2015).However, the extent to which they regulate survive within these warm and humid environments remains unknown. In Marigat, the community is at higher risk of getting bites from ticks and other vectors that transmit diseases to both animal and humans.

CHAPTER SIX

SUMMARY OF FINDINGS, CONCLUSION, LIMITATIONS AND RECOMMENDATIONS

6.1 Summary of findings

The objectives of this study were to taxonomically Identify small rodents and shrews that were captured in Marigat, Baringo county as well as to determine the Bacterial biota in the spleen tissue of this small mammals. Two genes the CO1 and Cytb Mitochondrial gene were used to identify the small mammals and together these genes were able to resolve a few discrepancies on the naming of organisms using the phenotypic characteristics. The bacterial 16S rRNA of 54 wild rodents and shrews was deep sequenced to unravel the diversity of bacteria communities harbored in their spleens. 196 unique bacteria OTUs were identified. Top 5 phyla by abundance included Proteobacteria (64.7%), Actinobacteria (18.0%), Firmicutes (6.1%), Chlamydiae (3.8%) and Chloroflexi (2.6%). Top 5 pathogenic bacteria genera detected included Bartonella (41.5%), Anaplasma (6.5%), Coxiella (2.6%), Acinetobacter (1.3%) and Ehrlichia (0.5). This diversity of bacteria communities observed in this study could be attributed to the fact that some bacteria exist in association with animals either as hosts or in symbiotic relationship.

6.2 Conclusion

This study has comprehensively characterized the complex bacteria diversity present in the small mammals at genus level. A number of zoonoses with Bartonella having the highest prevalence were detected in the small mammals collected in Marigat County. *Acomys wilsoni* and *Mastomys Natalensis* species had the highest prevalence of these zoonotic pathogens.

It is hoped that the findings of this study have presented useful data that will help local clinicians, the county government and the appropriate personnel to support regional residents and farmers with timely information and focused implementation of personal protective measures such as sanitation and mouse proof construction, disease prevention and control programs.

The information will also educate the clinicians on the bacterial pathogens circulating around the pastoral communities and thus they will be able to know what diseases these small mammal cause for ease of their management.

6.3 Limitation

We recognize that some limitations were present in our study which may have affected our final findings. 16s rRNA metagenomics does not sequence bacteria to the species level. We therefore could not speciate the present pathogens further than the genus level. We recommend that future studies go further to speciate on these bacterial pathogens.

6.4 Recommendations

From the findings of this study, it is recommended that the results from this study should inform the community in Marigat and the general public at large on the zoonotic pathogens harbored by small mammals in their environment. It should also inform them on what rodent or shrew has the highest abundance of the bacterial pathogens.

6.5 Suggestions for future studies

Coxiella and Rickettsia are considered vector borne pathogens. However, with the advances in molecular biology, some of their members are gradually recognized as non-pathogenic intracellular bacteria and may be endosymbionts to their host (Ge *et al*., 2018). We therefore

recommend that the bacterial pathogens such as Bartonella, Coxiella, Acinetobacter and Brucella just to mention a few that were present in the small mammals of this study to be further tested to what do you mean, provide a one stop type of technique that you think would assist in their proper identification determine the species

REFERENCES

al., S. J. S. et. (2017). Ahead of Print -Prioritizing Zoonoses for Global Health Capacity Building—Themes from One Health Zoonotic Disease Workshops in 7 Countries, 2014–2016 - Volume 23, Number 13—December 2017 - Emerging Infectious Disease journal - CDC. *Emerging Infectious Diseases*, *23*(13), 2014–2016. http://wwwnc.cdc.gov/eid/article/23/13/17-0418_article.htm%0Ahttps://wwwnc.cdc.gov/eid/article/23/13/17-0418_article

Backhans, A. (2011). *Wild rodents as carriers of potential pathogens to pigs, chickens and humans*. http://pub.epsilon.slu.se/8105/

Beattie, S. (1989). *Electron Transfer through Center*. *264*(24), 13994–13997.

Belay, E. D., Kile, J. C., Hall, A. J., Barton-Behravesh, C., Parsons, M. B., Salyer, S., & Walke, H. (2017). Zoonotic disease programs for enhancing global health security. *Emerging Infectious Diseases*, *23*(December), S65–S70. https://doi.org/10.3201/eid2313.170544

Bilgin, H., Sarmis, A., Tigen, E., Soyletir, G., & Mulazimoglu, L. (2015). *Case Report Delftia acidovorans : A rare pathogen in immunocompetent and immunocompromised patients*. *26*(5), 277–279.

Bruce H. Williams DVM, D. (2013). Zoonoses of small mammals. *Clinicians Brief*. www.cliniciansbrief.com/sites/default/.../Zoonoses of Small Mammals.pdf%0A%0A

Byrkjeland, R. (2015). *The role of small rodents and shrews as hosts for ticks and reservoirs of tick-borne pathogens in a northern coastal forest ecosystem*. 1–64.

Cardeñosa, N., Sanfeliu, I., Font, B., Muñoz, T., & Nogueras, M. M. (2006). *SHORT REPORT : SEROPREVALENCE OF HUMAN INFECTION BY COXIELLA BURNETII IN BARCELONA (NORTHEAST OF SPAIN)*. *75*(1), 33–35.

CDC. (2016). *Annual Report 2016*. https://doi.org/https://www.techcombank.com.vn/investors/annual-reports

CDC. (2019). *MMWR: Summary of Notifiable Infectious Diseases*. https://www.cdc.gov/mmwr/mmwr_nd/index.html

Center for Disease Control and Prevention. (2017). *Zoonotic Diseases*. *March*. https://www.cdc.gov/onehealth/basics/zoonotic-diseases.html

Couper, L., & Swei, A. (2018). Tick Microbiome Characterization by Next-Generation 16S rRNA Amplicon Sequencing. *Journal of Visualized Experiments*, *138*. https://doi.org/10.3791/58239

Croxatto, A., Prod'hom, G., & Greub, G. (2012). Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiology Reviews*, *36*(2), 380–407. https://doi.org/10.1111/j.1574-6976.2011.00298.x

Delić, T., Luštrik, R., Kljun, F., & Potočnik, H. (2013). Species diversity and composition of small mammal communities in Goteniška gora (S Slovenia). *NATURA SLOVENIAE*, *15*(1), 13–26.

Diagne, C. A., Charbonnel, N., Henttonen, H., Sironen, T., & Brouat, C. (2017). Serological Survey of Zoonotic Viruses in Invasive and Native Commensal Rodents in Senegal, West Africa. *Vector-Borne and Zoonotic Diseases*, *17*(10), 730–733. https://doi.org/10.1089/vbz.2017.2135

Dudu, A., Barbalata, T., Popa, G. O., Georgescu, S. E., Barbălată, T., Popa, G.-O., Georgescu, E., & Costache, M. (2016). Advantages and Limitations of DNA Barcoding in Identifying Commercially-Exploited Fish Species. *Scientific Papers: Animal Science and Biotechnologies*, *1*, 49. https://www.researchgate.net/publication/303699632

Galan, M., Razzauti, M., Bard, E., Bernard, M., Brouat, C., Charbonnel, N., Dehne-garcia, A., Loiseau, A., Tatard, C., Tamisier, L., & Vayssier-taussat, M. (2016). 16S rRNA amplicon sequencing for epidemiological surveys of bacteria in wildlife : the importance of cleaning post-sequencing data before estimating positivity , prevalence and co-infection. *MSystems*, *1*(4), e00032-16. https://doi.org/http://dx.doi.org/10.1101/039826

Ge, Y., Guo, G., Ge, B., Yin, H., & Yin, H. (2018). The spleen microbiota of small wild mammals reveals distinct patterns with tick-borne bacteria. *PLoS Neglected Tropical Diseases*, *12*(7), 1–21. https://doi.org/10.1371/journal.pntd.0006499

Gezmu, M., Bsrat, A., & Mekuria, A. (2017). *Assessment of Community Knowledge , Attitude and Practice on Zoonotic Disease in and Around Dodola Town , West Arsi Zone , Ethiopia*. *1*(1), 57–65.

Golden, J. W., Hammerbeck, C. D., Mucker, E. M., & Brocato, R. L. (2015). Animal Models for the Study of Rodent-Borne Hemorrhagic Fever Viruses: Arenaviruses and Hantaviruses. *BioMed Research International*, *2015*. https://doi.org/10.1155/2015/793257

Gruber, K. (2016). Rodent meat – a sustainable way to feed the world? *EMBO Reports*, *17*(5), 630–633. https://doi.org/10.15252/embr.201642306

Guterres, A., & de Lemos, E. R. S. (2018). Hantaviruses and a neglected environmental determinant. *One Health*, *5*(October 2017), 27–33. https://doi.org/ 10.1016/j. onehlt.2017.12.002

Han, B. A., Schmidt, J. P., Bowden, S. E., & Drake, J. M. (2015). Rodent reservoirs of future zoonotic diseases. *Proceedings of the National Academy of Sciences*, *112*(22), 7039–7044. https://doi.org/10.1073/pnas.1501598112

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003a). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, *270*(1512), 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003b). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rspb.2002.2218

Hornok, S., Földvári, G., Rigó, K., Meli, M. L., Gönczi, E., Répási, A., Farkas, R., Papp, I., Kontschán, J., & Hofmann-Lehmann, R. (2015). Synanthropic rodents and their ectoparasites as carriers of a novel haemoplasma and vector-borne, zoonotic pathogens indoors. *Parasites & Vectors*, *8*(1), 27. https://doi.org/10.1186/s13071-014-0630-3

K Chen, Neimark H, Rumore P, S. C. (1989). Broad range DNA probes for detecting and amplifying eubacterial nucleic acids. *FEMS Microbiol Lett*, *1*(49), 19–24.

Kay, E. H., & Hoekstra, H. E. (2008). Rodents. *Current Biology*, *18*(10), 406–410. https://doi.org/10.1016/j.cub.2008.03.019

Kiarie-makara, M. (2016). *A Descriptive Survey on Knowledge , Attitude , Practices and Beliefs on Kala Azar Among the Residents of ...January*. https://doi.org/10.5455/209302/ijtmph

Kosoy, M., Bai, Y., Lynch, T., Kuzmin, I. V., Niezgoda, M., Franka, R., Agwanda, B., Breiman, R. F., & Rupprecht, C. E. (2010). Bartonella spp. in bats, Kenya. *Emerging Infectious Diseases*, *16*(12), 1875–1881. https://doi.org/10.3201/eid1612.100601

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

Ladoukakis, E. D., & Zouros, E. (2017). Evolution and inheritance of animal mitochondrial DNA : rules and exceptions. *Journal of Biological Research-Thessaloniki*, 1–7. https://doi.org/10.1186/s40709-017-0060-4

Lee, I.-Y., Kim, H.-C., Shringi, S., Foley, J., Yu, D.-H., Chong, S.-T., Klein, T. A., & Chae, J.-S. (2010). Microbial pathogens in ticks, rodents and a shrew in northern Gyeonggi-do near the DMZ, Korea. *Journal of Veterinary Science*, *9*(3), 285. https://doi.org/10.4142/jvs.2008.9.3.285

Lindahl, J. F., & Grace, D. (2015). *The consequences of human actions on risks for infectious diseases: a review*. *1*, 1–11.

Lõhmus, M., Janse, I., van de Goot, F., & van Rotterdam, B. J. (2013). Rodents as Potential Couriers for Bioterrorism Agents. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, *11*(S1), S247–S257. https://doi.org/10.1089/bsp.2012.0085

Lu, L., Chesters, D., Zhang, W., Li, G., Ma, Y., Ma, H., Song, X., Wu, H., Meng, F., Zhu, C., & Liu, Q. (2012). Small Mammal Investigation in Spotted Fever Focus with DNA-Barcoding and Taxonomic Implications on Rodents Species from Hainan of China. *PLoS ONE*, *7*(8), 1–13. https://doi.org/10.1371/journal.pone.0043479

Masakhwe, C., Linsuwanon, P., Kimita, G., Mutai, B., Leepitakrat, S., Yalwala, S., Abuom, D., Auysawasi, N., Gilbreath, T., Wanja, E., & Waitumbi, J. (2018a). Identification and characterization of Orientia chuto in trombiculid chigger mites collected from wild rodents in Kenya. *Journal of Clinical Microbiology*, *47*(October), JCM.01124-18. https://doi.org/10.1128/JCM.01124-18

Masakhwe, C., Linsuwanon, P., Kimita, G., Mutai, B., Leepitakrat, S., Yalwala, S., Abuom, D., Auysawasi, N., Gilbreath, T., Wanja, E., & Waitumbi, J. (2018b). Identification and characterization of Orientia chuto in trombiculid chigger mites collected from wild rodents in Kenya. *Journal of Clinical Microbiology*, *56*(12). https://doi.org/10.1128/JCM.01124-18

Morwal, S. (2017). Bacterial Zoonosis - A Public Health Importance. *Journal of Dairy, Veterinary & Animal Research*, *5*(2). https://doi.org/10.15406/jdvar.2017.05.00135

Motro, Y., & Moran-Gilad, J. (2017). Next-generation sequencing applications in clinical bacteriology. In *Biomolecular Detection and Quantification* (Vol. 14, pp. 1–6). https://doi.org/10.1016/j.bdq.2017.10.002

Müller, L., Gonçalves, G. L., Cordeiro-Estrela, P., Marinho, J. R., Althoff, S. L., Testoni, A. F., González, E. M., & Freitas, T. R. O. (2013). DNA barcoding of Sigmodontine rodents: Identifying wildlife reservoirs of zoonoses. *PLoS ONE*, *8*(11), 1–12. https://doi.org/10.1371/journal.pone.0080282

Mutai, B., Njaanake, K., Gathii, K., Estambale, B. B., & Waitumbi, J. N. (2019). *A Duo 4-Plex Real Time PCR for Detection of Eight Tick-Borne Zoonoses in Kenya*. 1–15. https://doi.org/10.4236/ojcd.2019.91001

Mutambuki, K., & Ngatia, C. M. (2006). Losss assessment of on-farm stored maize in semi arid area of Kitui District , Kenya. *9Th International Working Conference on Stored Product Protection*, 15–23. http://spiru.cgahr.ksu.edu/proj/iwcspp/iwcspp9.html

Nachum-biala, Y., & Harrus, S. (2015). *Relationship between the Presence of Bartonella Species and Bacterial Loads in Cats and Cat Fleas (Ctenocephalides felis) under Natural Conditions*. *81*(16), 5613–5621. https://doi.org/10.1128/AEM.01370-15

Ndao, M. (2009). Diagnosis of Parasitic Diseases: Old and New Approaches. *Interdisciplinary Perspectives on Infectious Diseases*, *2009*, 1–15. https://doi.org/10.1155/2009/278246

Nicolas, V., Schaeffer, B., Missoup, A. D., Kennis, J., Colyn, M., Denys, C., Tatard, C., Cruaud, C., & Laredo, C. (2012). Assessment of three mitochondrial genes (16S, Cytb, CO1) for identifying species in the Praomyini tribe (Rodentia: Muridae). *PLoS ONE*, *7*(5), 1–11. https://doi.org/10.1371/journal.pone.0036586

Norbu, L. (2016). *Small Mammal : Rodents and Shrews of Bumdeling Ramsar Site , Trashi Yangtse- Bhutan*.

Oswe, M., Odhiambo, R., Mutai, B., Nyakoe, N., Awinda, G., & Waitumbi, J. N. (2018). Zoonotic Pathogens in Ticks Collected from Livestock in Kenya. *Open Journal of Preventive Medicine*, *08*(08), 248–259. https://doi.org/10.4236/ojpm.2018.88021

Ouvery, N. A. R., & Odolakis, A. R. (2005). *Review article Is Q Fever an emerging or re-emerging zoonosis ?36*, 327–349. https://doi.org/10.1051/vetres

Pentinsaari, M., Salmela, H., Mutanen, M., & Roslin, T. (2016). Molecular evolution of a widely-adopted taxonomic marker (COI) across the animal tree of life. *Scientific Reports*, *6*(September), 1–12. https://doi.org/10.1038/srep35275

Putignani, L., & Conti, S. (2015). *Bacterial Pangenomics*. *1231*(May). https://doi.org/10.1007/978-1-4939-1720-4

Rabozzi, G., Bonizzi, L., Crespi, E., Somaruga, C., Sokooti, M., Tabibi, R., Vellere, F., Brambilla, G., & Colosio, C. (2012). Emerging Zoonoses: the "One Health Approach." *Safety and Health at Work*, *3*(1), 77–83. https://doi.org/10.5491/SHAW.2012.3.1.77

Ripple, W. J., Abernethy, K., Betts, M. G., Chapron, G., Dirzo, R., Galetti, M., Levi, T., Lindsey, P. A., Macdonald, D. W., Machovina, B., Newsome, T. M., Peres, C. A., Wallach, A. D., Wolf, C., & Young, H. (2016). Bushmeat hunting and extinction risk to the world ' s mammals Subject Category : Subject Areas : *Open Science*, *3*, 160498. https://doi.org/10.1098/rsos.160498

Rosselli, R., Romoli, O., Vitulo, N., Vezzi, A., Campanaro, S., De Pascale, F., Schiavon, R., Tiarca, M., Poletto, F., Concheri, G., Valle, G., & Squartini, A. (2016). Direct 16S rRNA-seq from bacterial communities: A PCR-independent approach to simultaneously assess microbial diversity and functional activity potential of each taxon. *Scientific Reports*, *6*(August), 1–12. https://doi.org/10.1038/srep32165

Singleton, G. (2003). Impacts of rodents on rice production in Asia. *IRRI Discuss. Pap. Ser*, *45*, 30. http://books.irri.org/971220183X_content.pdf

Team, R. C. (n.d.). *R: A Language and Environment for Statistical Computing*.

Thiga, J. W., Mutai, B. K., Eyako, W. K., Ng'Ang'A, Z., Jiang, J., Richards, A. L., & Waitumbi, J. N. (2015). High seroprevalence of antibodies against spotted fever and scrub typhus bacteria in patients with febrile illness, Kenya. *Emerging Infectious Diseases*, *21*(4). https://doi.org/10.3201/eid2104.141387

Thiga, J. W., Mutai, B. K., Richards, A. L., & Waitumbi, J. N. (2015). High Seroprevalence of Antibodies against in Patients with Febrile Illness, Kenya. *Emerging Infectious Diseases*, *21*(4), 688–691.

Tobe, S. S., Kitchener, A., & Linacre, A. (2009). Cytochrome b or cytochrome c oxidase subunit I for mammalian species identification-An answer to the debate. *Forensic Science International: Genetics Supplement Series*, *2*(1), 306–307. https://doi.org/10.1016/j.fsigss.2009.08.053

Venkatesh, S., Chhabra, M., Balakrishnan, Gupta, N., Singhai, M., Tiwari, S., Cordelia, & Sharma, P. (2016). *Zoonotic Diseases of Public Health Importance*.

Wilson, D. E., & Reeder, D. M. (2005). *Mammal Species of the World* (D. E. Wilson (ed.); Third). John Hopkins University Press.

APPENDICES

Appendix 1Table showing the small mammals captured with metadata included.

Sample	Species (morphological identification)	Sex	Weight	Body(cm)	Site of collection
R001	Mastomys Natalensis	M	94g	28 cm	Marigat Kalro
R002	Mastomys Natalensis	M	63g	22 cm	Marigat Kalro
R003	Lopheromys	F	85g	29 cm	Marigat Kalro
R004	Acomys Wilsoni	F	31g	21 cm	Marigat Kalro
R005	Acomys Wilsoni	F	21g	13 cm	Marigat Kalro
R006	Acomys Wilsoni	F	24g	8.4 cm	Marigat Kalro
R007	Acomys Wilsoni	F	20 g	8 cm	Marigat Kalro
R008	Arvicanthis	M	78g	27.8 cm	Marigat Kalro
R009	Arvicanthis	F	48g	19.5 cm	Marigat Kalro
R010	Arvicanthis	F	63 g	19 cm	Marigat Kalro
R011	Crocidura	F	10g	12.3 cm	Marigat Kalro
R012	Arvicanthis	M	100g	23cm	Marigat Kalro
R013	Crocidura	F	7g	15.5 cm	Marigat Kalro
R014	Acomys Wilsoni	M	20g	13 cm	Marigat Kalro
R015	Crocidura	F	14g	13.5 cm	Marigat Kalro
R016	Mastomys Natalensis	M	54g	22.5 cm	Marigat Kalro
R017	Mastomys Natalensis	M	89g	27.5 cm	Marigat Kalro
R018	Acomys Wilsoni	F	30g	9 cm	Marigat Kalro
R019	Acomys Wilsoni	F	20g	8 cm	Marigat Kalro
R020	Crocidura	F	10g	13.4 cm	Marigat Kalro
R021	Acomys Wilsoni	M	21g	8cm	Marigat Kalro
R022	Acomys Wilsoni	M	17g	8.5cm	Marigat Kalro
R023	Crocidura	F	10g	13.9cm	Marigat Kalro
R024	Acomys Wilsoni	F	18g	11cm	Marigat Kalro
R025	Rattus Rattus	M	142g	34.5 cm	Marigat Kalro
R026	Acomys Wilsoni	F	19g	12.4 cm	Marigat Kalro
R027	Crocidura	F	8g	12.4 cm	Marigat Kalro
R028	Crocidura	M	17g	14.6 cm	Marigat Kalro
R029	Acomys Wilsoni	M	23g	12cm	Marigat Kalro
R030	Mastomys Natalensis	F	70g	23.2 cm	Marigat Kalro

R031	Mastomys Natalensis	F	52g	21.2 cm	Marigat Kalro
R032	Mastomys Natalensis	M	79g	24.5cm	Marigat Kalro
R033	Crocidura	F	10.4g	13.4 cm	Marigat Kalro
R034	Crocidura	F	8g	12cm	Marigat Kalro
R035	Acomys Wilsoni	F	18g	11.5 cm	Marigat Kalro
R036	Crocidura	M	10g	13.6 cm	Marigat Kalro
R037	Acomys Wilsoni	M	19g	12.4 cm	Marigat Kalro
R038	Crocidura	M	11g	14 cm	Marigat Kalro
R039	Acomys Wilsoni	M	18g	13 cm	Marigat Kalro
R040	Crocidura	F	7g	11.3 cm	Marigat Kalro
R041	Crocidura	F	11g	14.2 cm	Marigat Kalro
R042	Acomys Wilsoni	F	18g	11.2 cm	Marigat Kalro
R043	Crocidura	F	8g	12.7 cm	Marigat Kalro
R044	Acomys Wilsoni	F	22g	9cm	Marigat Kalro
R045	Arvicanthis	M	62g	23 cm	Marigat Kalro
R046	Rattus Rattus	M	98g	32 cm	Marigat Kalro
R047	Arvicanthis	M	105g	23.5cm	Marigat Kalro
R048	Acomys Wilsoni	M	18g	8.0 cm	Marigat Kalro
R049	Mastomys Natalensis	F	24g	19.0 cm	Marigat Kalro
R050	Acomys Wilsoni	M	23g	12.6 cm	Marigat Kalro
R051	Acomys Wilsoni	F	24 g	13.5 cm	Marigat Kalro
R052	Acomys Wilsoni	M	22g	12.1 cm	Marigat Kalro
R053	Crocidura	M	14g	13.2 cm	Marigat Kalro
R054	Crocidura	F	6g	11.5 cm	Marigat Kalro

Appendix 2 Table showing the total number of bacterial pathogens detected in the small mammal and their abundances

Bacterial Genus species	Abundance	Bacterial Genus Species	Abundance
Bartonella	18204	Hyphomicrobium	30
Anaplasma	2814	Comamonas	29
Methylobacterium	1925	Conexibacter	29
Coxiella	1778	Asaccharobacter	27
Delftia	1649	Aurantimonas	27
Lactobacillus	1000	Prevotella	27
Corynebacterium	974	SR1_genus_incertae_sedis	27
Micrococcus	973	Steroidobacter	27
Bradyrhizobium	856	Ilumatobacter	26
Nocardioides	660	Alcaligenes	25
Acinetobacter	606	Brevibacterium	25
Sphaerobacter	593	Pelomonas	25
Sphingomonas	591	Borrelia	24
Achromobacter	524	Nocardiopsis	23
Streptomyces	480	Oscillibacter	22
Rhodococcus	479	Armatimonas_Armatimonadetes_gp1	21
Stenotrophomonas	478	Sporomusa	21
Blastococcus	470	Tepidimonas	21
Streptococcus	470	Adhaeribacter	20
Propionibacterium	438	Kytococcus	20
Brevundimonas	375	Solirubrobacter	20
Pseudomonas	353	Lysobacter	19
Staphylococcus	311	Clostridium_IV	18
Bacillus	273	Flavonifractor	18
Anoxybacillus	249	Ohtaekwangia	18
Dietzia	247	Cupriavidus	16
Clostridium_sensu_stricto	231	Proteus	16
Mycobacterium	227	Turicella	15
Kocuria	225	Catonella	14
Gp3	224	Spartobacteria_genera_incertae_sedis	14
Ehrlichia	211	Brevinema	13
Bosea	210	Capnocytophaga	13
Paracoccus	207	Flavisolibacter	13
Arthrobacter	201	Microlunatus	13
Rubrobacter	201	Chlamydia	12

Helicobacter	199	Schlegelella	12
Rothia	198	Alistipes	11
Dermacoccus	189	Oxalicibacterium	11
Brachybacterium	187	Pseudochrobactrum	11
Weissella	184	Ralstonia	11
Saccharopolyspora	181	Rickettsia	11
Hymenobacter	178	Salmonella	11
Lactococcus	177	Saxeibacter	11
Microvirga	175	Selenomonas	11
Massilia	152	Jannaschia	10
Novosphingobium	150	Petrobacter	10
Ochrobactrum	147	Amycolatopsis	9
Actinomyces	145	Butyricicoccus	9
Haloferula	142	Enterorhabdus	9
Meiothermus	134	Lachnospiracea_incertae_sedis	9
Ruminococcus	126	Aquamicrobium	8
Sphingopyxis	124	Hydrogenophaga	8
Klebsiella	123	Pasteurella	8
Haemophilus	121	Quadrisphaera	8
Aquabacterium	120	Afipia	7
Cloacibacterium	116	Geminicoccus	7
Geodermatophilus	116	Methylocystis	7
Variovorax	114	Enterobacter	6
Roseomonas	111	Finegoldia	6
Gordonia	109	Peredibacter	6
Microbacterium	106	Roseburia	6
Devosia	105	Aerococcus	5
Luteolibacter	103	Alkanindiges	5
Flavobacterium	102	Coprococcus	5
Aeromonas	99	Deinococcus	5
Barnesiella	97	Gp10	5
Singulisphaera	97	Luteococcus	5
Phenylobacterium	96	Olsenella	5
Rubellimicrobium	91	Pedobacter	5
Fusobacterium	85	Actinomycetospora	4
Aeromicrobium	82	Atopostipes	4
Chryseobacterium	82	Azospira	4
Gp16	79	Clostridium_XI	4
Kineococcus	78	Janibacter	4
Leucobacter	76	Jeotgalicoccus	4
Ramlibacter	74	Macrococcus	4
Leptotrichia	72	Mucilaginibacter	4
Luteimonas	71	Parabacteroides	4

Rhodocytophaga	70	Phaselicystis	4
Gemella	68	Nevskia	3
Gemmatimonas	68	Parasutterella	3
Morganella	68	Prosthecobacter	3
Leifsonia	67	Rhodobacter	3
Gp4	66	Sphingobium	3
Pseudonocardia	63	Trichococcus	3
Thermicanus	62	Allobaculum	2
Diaphorobacter	60	Anaeroplasma	2
Porphyromonas	60	Peptoniphilus	2
TM7_genus_incertae_sedis	60	Pigmentiphaga	2
Verrucomicrobium	54	Pimelobacter	2
Enterococcus	53	Rhodocista	2
Armatimonadetes_gp5	52	Streptacidiphilus	2
Amaricoccus	50	Alishewanella	1
OD1_genus_incertae_sedis	50	Brucella	1
Rhodopirellula	45	Camelimonas	1
Clostridium_XlVa	43	Cellvibrio	1
Rhodanobacter	42	Geobacillus	1
Acidovorax	41	Gp17	1
Cellulomonas	40	Hydrotalea	1
Gp6	40	Kofleria	1
Rhizobium	40	Naxibacter	1
Cellulosilyticum	39	Propionicimonas	1
Pseudoxanthomonas	39	Sphingosinicella	1
Enhydrobacter	38		
Agromyces	35		
Escherichia_Shigella	33		
Ornithinimicrobium	33		
Bdellovibrio	32		
Marmoricola	31		
Bacteroides	30		

Appendix 3Ethical Approval

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **March 12, 2012**

TO: **MAJ JOSHUA BAST (PRINCIPAL INVESTIGATOR)**

THROUGH: FOR- **DR. JUMA RASHID, THE DIRECTOR, CCR, NAIROBI**

Forwarded 13/03/2012

Dear Sir,

RE: SSC PROTOCOL No. 2208 (*RE-SUBMISSION*): PREVALENCE OF RODENT PARASITES (ECTO AND ENDO PARASITES) AND THEIR IMPLICATION FOR ZOONOSES MONITORING IN KENYA

Reference is made to your letter dated March 5, 2012. We acknowledge receipt a copy of the approval letter from the KEMRI Animal Care and Use Committee dated 12th August 2011.

This is to inform you that the Committee determines that the issue raised at the initial review is adequately addressed. Consequently, the study is granted approval for implementation effective this **12th day of March 2012** for a period of one year.

Please note that authorization to conduct this study will automatically expire on **March 11, 2013**. If you plan to continue data collection or analysis beyond this date, please submit an application for continuation approval to the ERC Secretariat by **January 28, 2013**. The regulations require continuing review even though the research activity may not have begun until sometime after the ERC approval.

Please note that any unanticipated problems resulting from the implementation of this study should be brought to the attention of the ERC. You are also required to submit any proposed changes to this study to the SSC and ERC for review and approval prior to initiation and advise the ERC when the study is completed or discontinued.

Work on this project may begin.

Sincerely,

CHRISTINE WASUNNA,
Ag. SECRETARY,
KEMRI ETHICS REVIEW COMMITTEE

Appendix 4 Comparison of morphological taxonomy, molecular phylogeny and concatenated taxonomy

Rodent ID	Sex	Morphological	*Cyt b*	*CO1*	Concatenated Cyt *b* and *C01*
R004	F	Acomys	Acomys	Acomys	Acomys
R005	F	Acomys	Acomys	NA	Acomys
R006	F	Acomys	Acomys	NA	Acomys
R007	F	Acomys	Acomys	Acomys	Acomys
R014	M	Acomys	Acomys	Acomys	Acomys
R018	F	Acomys	Acomys	Acomys	Acomys
R019	F	Acomys	Acomys	Acomys	Acomys
R021	M	Acomys	Acomys	Acomys	Acomys
R022	M	Acomys	Acomys	Acomys	Acomys
R024	F	Acomys	Acomys	Acomys	Acomys
R026	F	Acomys	Acomys	Acomys	Acomys
R029	M	Acomys	Acomys	Acomys	Acomys
R035	F	Acomys	Acomys	Acomys	Acomys
R037	M	Acomys	Acomys	Acomys	Acomys
R039	M	Acomys	Acomys	Acomys	Acomys
R042	F	Acomys	Crocidura	Acomys	NA
R044	F	Acomys	NA	Acomys	Acomys
R048	M	Acomys	Mastomys	Acomys	NA
R050	M	Acomys	Acomys	Acomys	Acomys
R051	F	Acomys	Acomys	Acomys	NA
R052	M	Acomys	NA	Acomys	NA
R008	M	Arvicanthis	Arvicanthis	NA	NA
R009	F	Arvicanthis	Arvicanthis	Lemniscomys	Arvicanthis
R010	F	Arvicanthis	Arvicanthis	Lemniscomys	Arvicanthis
R012	M	Arvicanthis	Arvicanthis	Lemniscomys	Arvicanthis
R045	M	Arvicanthis	Rattus	Rodentia	Rattus
R047	M	Arvicanthis	Acomys	Mus	Acomys
R011	F	Crocidura	NA	NA	NA
R013	F	Crocidura	Crocidura	Crocidura	NA
R015	F	Crocidura	Crocidura	NA	NA
R020	F	Crocidura	NA	NA	NA
R023	F	Crocidura	Crocidura	NA	NA
R027	F	Crocidura	Crocidura	NA	NA
R028	M	Crocidura	Crocidura	NA	NA
R033	F	Crocidura	Crocidura	Crocidura	NA

R034	F	Crocidura	Crocidura	Crocidura	NA
R040	F	Crocidura	Crocidura	NA	NA
R041	F	Crocidura	Acomys	NA	NA
R043	F	Crocidura	Acomys	NA	NA
R053	M	Crocidura	Crocidura	NA	NA
R054	F	Crocidura	Rattus	NA	NA
R036	M	Crocidura	NA	NA	NA
R038	M	Crocidura	Crocidura	Crocidura	NA
R003	F	Lophuromys	Arvicanthis	Lenothrix	Arvicanthis
R001	M	Mastomys	Mastomys	NA	NA
R002	M	Mastomys	Mastomys	NA	NA
R016	M	Mastomys	Mastomys	Mastomys	Mastomys
R017	M	Mastomys	Mastomys	Mastomys	Mastomys
R030	F	Mastomys	Mastomys	Mastomys	Mastomys
R031	F	Mastomys	Mastomys	Mastomys	Mastomys
R032	M	Mastomys	Mastomys	Mastomys	Mastomys
R049	F	Mastomys	Acomys	Mastomys	Acomys
R025	M	Rattus	Rattus	Rattus	Rattus
R046	M	Rattus	Arvicanthis	Rattus	Arvicanthis

Appendix 5 Table showing the Phyla present in each small mammals after analysis.

Phyla\Sample ID	R001	R002	R003	R004	R005	R006	R007	R008	R009	R010	R011	R012	R013	R014	R016	R017	R018	R019	R020	R021	R022	R024	R025	R026	R027	R029	R030	R031	R032	R033	R035	R037	R038	R039	R042	R044	R045	R046	R047	R048	R049	R050	R052	R053	Total
Acidobacteria	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53	0	0	0	0	6	0	0	44	0	0	0	5	0	0	0	0	0	0	21	0	0	16	1	116	65	0	339
Actinobacteria	7	29	55	227	51	366	367	5	3	34	103	24	516	274	0	2	85	169	8	331	24	338	7	177	593	591	2	71	37	424	298	323	1	153	141	38	193	10	58	444	2	130	32	1	6744
Armatimonadetes	0	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	73
Bacteroidetes	2	108	40	93	1	0	0	7	6	40	9	9	0	0	0	0	55	0	0	0	0	0	1	18	19	0	2	0	0	0	0	27	0	91	26	0	20	9	4	54	0	15	20	0	676
Chlamydiae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	2	12
Chloroflexi	0	0	0	0	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	399	0	0	4	0	6	0	0	0	0	0	0	162	0	0	593
Deinococcus-Thermus	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	11	0	0	123	0	0	0	0	0	0	0	0	0	3	0	0	139
Firmicutes	5	96	9	246	49	29	47	49	16	91	6	34	1	112	0	0	0	69	1	82	254	0	9	85	14	8	10	0	18	23	21	7	3	94	14	975	23	8	50	219	2	2	142	4	2927
Fusobacteria	0	150	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	157
Gemmatimonadetes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	0	34	0	4	0	0	0	0	0	68
OD1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	44	0	0	0	6	0	0	0	0	0	0	0	0	50
Planctomycetes	0	0	0	42	19	0	0	0	0	0	0	0	0	0	0	0	45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	25	0	0	0	0	0	0	0	0	7	0	142
Proteobacteria	1196	496	959	452	993	754	643	1011	1082	670	1020	1054	618	665	1216	1222	902	781	1226	426	915	591	1183	490	134	221	1199	1166	1112	113	226	542	1187	243	527	169	652	83	788	247	1204	168	316	48	30910
Spirochaetes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	2	37
SR1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27
Tenericutes	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
TM7	4	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	60
Verrucomicrobia	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	43	0	0	132	0	0	0	99	0	0	0	0	0	0	1	0	0	0	33	0	5	0	0	0	0	0	0	0	316
Grand Total	**1215**	**879**	**1075**	**1060**	**1156**	**1149**	**1057**	**1073**	**1112**	**835**	**1138**	**1128**	**1135**	**1051**	**1216**	**1224**	**1130**	**1124**	**1235**	**995**	**1193**	**929**	**1207**	**896**	**760**	**879**	**1213**	**1237**	**1178**	**964**	**546**	**1066**	**1199**	**606**	**817**	**1188**	**948**	**121**	**904**	**980**	**1209**	**596**	**592**	**57**	**43272**