DIVERSITY, DISTRIBUTION AND ABUNDANCE OF PROKARYOTIC COMMUNITIES WITHIN SELECTED FORESTS ECOSYSTEM IN KENYA

LORINE AKINYI ONYANGO

MASTER OF SCIENCE

(Molecular Biology & Bioinformatics)

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

2024

Diversity, Distribution and Abundance of Prokaryotic Communities within Selected Forests Ecosystem in Kenya

Lorine Akinyi Onyango

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Molecular Biology and Bioinformatics of the Jomo Kenyatta University of Agriculture and Technology

2024

DECLARATION

This thesis is my original work and has not been presented for a degree in any other university

Signature……. Date……

Lorine Akinyi Onyango

This thesis has been submitted for examination with our approval as the University Supervisors

Signature……. Date……

Dr. Florence Ng'ong'a, PhD
JKUAT, Kenya

Signature……. Date……

Dr. Josiah Ochieng Kuja, PhD
JKUAT, Kenya

Signature……. Date……

Dr. Anne Kelly Kambura, PhD
Taita Taveta University, Kenya

DEDICATION

This work is dedicated to my parents, Dr. Francis Odawa and my late mother Jenipher Apondi. My sincere gratitude goes to my manager, Zachary Mwangi for giving me an easy time during my studies and my siblings for the exclusive support. I appreciate the support you have accorded me during the course of my studies. Without your encouragement and support this journey would have been tough.

ACKNOWLEDGEMENT

I am indebted to Almighty God for bringing me this far from the commencement of this long and challenging journey.

I wish to sincerely thank the following institutions for their assistance in this research. The Jomo Kenyatta University of Agriculture and Technology, Taita Taveta University, University of Pretoria USAID, and Oppenheimer Foundation for a research grant, Kenya Wildlife Service for access permit to the sampling sites, National Commission for Science, International Centre of Insect Physiology and Ecology for the financial support.

I wish to acknowledge my supervisors, Dr. Anne Kelly Kambura for the introduction to the bioinformatics world, inspiration, support and guidance, Dr. Florence Ng'ong'a and Dr. Josiah Ochieng Kuja for their indispensable reach, inspiration, support and guidance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS AND ACRONYMS

ARG	Antimicrobial Resistance Genes
ASVs	Amplicon Sequence Variants
BLAST	Basic Local Alignment Search Tool
C	Carbon
DNA	Deoxyribonucleic Acid
GBA	Guilt-By-Association’
ICPOES	Inductively Coupled Plasma Optical Emission Spectrometry
KEPHIS	Kenya Plant Health Inspectorate Authority
KWS	Kenya Wildlife Service
N	Nitrogen
N2	Nitrogen gas
NACOSTI	National Commission for Science, technology and Innovation
NH3	Ammonia
OUT	Operational Taxonomic Unit
P	Phosphorus
PCA	Principal Component Analysis
PCoA	Principal Coordinate Analysis
PERMANOVA	Permutational Multivariate Analyses of Variance

QIIME	Quantitative Insights into Microbial Ecology
RNA	Ribonucleic Acid
RNA-Seq	Ribonucleic Acid Sequencing
SOM	Self-Organizing Map
TOC	Total Organic Carbon
TN	Total Nitrogen
UTR	Untranslated regions
UPGMA	Unweighted Pair Group Method with Arithmetic mean
LPDAAC	Land Processes Distributed Active Archive Center's
LDL	Linear discriminant analysis

ABSTRACT

Soil microbiomes in forest ecosystems play a crucial role in serving as either sources or sinks of nutrients by participating in activities such as decomposing organic matter, cycling nutrients, incorporating humic compounds into the soil, and facilitating the connection between plant and ecosystem functions. Prokaryotic communities colonize numerous habitats within forests ecosystem; comprising litter, deadwood, rhizosphere and bulk soil where populations are shaped by nutrient availability and biotic interactions. This study determined the composition, diversity and distribution of prokaryotes within selected forests ecosystem in Kenya. Thirty-one (31) soil samples were collected from selected forests ecosystem in Kenya. To identify the possible abiotic drivers for prokaryotic distribution, physiochemical characteristics for the soil samples were analyzed. This was followed by total DNA extraction, purity assessment, amplification and sequencing of the hypervariable region (V4 - V5) of the 16S rRNA gene using Illumina platform. Demultiplexing of high throughput sequence and statistical analysis was done using QIIME2 and R programming language. Linear discriminant analysis (LDA) effect size (LEfSe) was used to detect prokaryotic taxa that were differentially abundant within and between soil samples. Biodiversity metrics (alpha diversity) and community structure dissimilarity (beta diversity) were calculated using the vegan (version 2.5.7) and phyloseq (version 1.16.2) packages in RStudio. The environmental drivers of prokaryotic community structure were estimated using Redundancy analysis. The meta data file and their associated sequence datasets from selected forests around the globe were downloaded from publicly available databases and processed using the QIIME2 pipeline as described above. From this study, the key prokaryotic community drivers included sodium, silt, magnesium, calcium, potassium, pH and carbon whereas aluminium, phosphorus, iron, clay and sand negatively influenced diversity in both Principal Component Analysis 1 and Principal Component Analysis 2. A total of 1,944,316 high quality sequence reads were generated and clustered into 41,901 ASVs (Amplicon Sequence Variants) at 3% genetic distance. Taxonomic classification of the obtained ASVs were assigned to a prokaryotic Kingdom, 2 Phyla, 120 Classes, 280 Orders, 450 Families, 873 Genera and 2313 Species within selected forests ecosystem. Archaeal groups recovered from the obtained ASVs within the selected forests ecosystem were distributed among seven phyla with *Crenarchaeota* as the most abundant Archaeal phylum represented across all samples, with 91.6% mean relative abundance. Analysis of sample alpha-diversity showed that soils from Western and Taita Taveta regions had significantly different (P = 0.0124603) levels of Archaeal richness, Western and Aberdare regions soil displayed Archaeal Shannon diversity index (P=0.00399513) but there were no significant differences between bacterial communities displayed within various forests ecosystems. However, beta-diversity analysis of soil samples from Western, Aberdare and Taita Taveta regions revealed a significant difference (P = 0.0010998) on bacterial and archaeal community structure (Bacteria R^2= 0.19; Archaea R^2 = 0.22). Samples from the different ecoregions showed significant differences (p-value= 0.001998, R^2= 0.45) in soil physiochemical properties, specifically in soil pH, soil texture, macro- and micro-nutrient composition and Enhanced Vegetation Index-2. Taita Taveta forest soil were highly distinct from those obtained from the Nairobi, Aberdare and Western regions. Nairobi and Western region soils exhibited

the least variability. The examination of beta-diversity scores across these datasets, utilizing the Bray-Curtis index, uncovered distinctions in the community structures of forest soil microbiomes, partly influenced by their country of origin ($R2 = 0.63$; p-value = 0.0098). The notable variations in composition between national datasets were further corroborated by Linear Discrimination Analysis, which highlighted 177 taxa distributed disparately among the datasets. The study demonstrates that Kenyan forest soils are unique and harbor potentially distinct soil microbiomes. However, more studies on forest microbiome should be done with focus on revealing vulnerability to possible future losses in forest soil microbial diversity and productivity due to climate change.

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Soil microorganisms exist as single cells or as multi-cellular organisms and they can only be seen under the microscope (Jacobsen *et al.*, 2012). They are classified as eukaryotes and prokaryotes, which help in regulation of plant productivity in nutrient poor ecosystems where plant symbionts contribute in limitation of nutrients (Van Der Heijden *et al.*, 2008).

Kenyan forests are located in high potential areas in which more than 70% of the population is settled (Bleher *et al.*, 2006). Some of these forests are located within Taita Taveta, Nairobi, Western, Aberdare and Mt. Kenya ecoregions. Mt. Kenya forest reserve is located to the east of the Great Rift Valley, along Latitude 0' 10' south and longitude 37' 20' east. It bestrides the equator in the central highland zones of Kenya and situated in two Forest Conservancies and five forest management zones. These include; Meru South, Nyeri, Meru Central, Embu in Eastern Conservancy and Kirinyaga in Central Highlands Conservancy (Kilonzi & Ota, 2019). From this forest the areas of interest were Mt. Kenya Castle Kimunge, Kyang'ondu, Kinondoni, Sirimon and Mt. Kenya State lodge.

The Aberdare Forest Reserve is located to the east of the Great Rift Valley, between 360 30' E, 00 05'S and 360 55'E, 00 450S (Kilonzi & Ota, 2019) .It's situated within Central Kenya in four counties, which include; Nyandarua, Nyeri, Kiambu and Murang'a. Its climate ranges is determined by the altitude (Wambugu, 2018). Taita Taveta region in southern Kenya ,is an area that is rich in biodiversity, comprising Taita Hills in Chawia, Ngangao and Vuria (Abera *et al.*, 2022).The Hills are part of the Eastern Arc Mountains, known for their rich endemism, and represent one of the world's biodiversity hotspots (Platts *et al.*, 2011). The hilltops are covered with indigenous evergreen montane forest, forest plantations like Susua plantation, agroforestry, and croplands (Pellikka *et al.*, 2009). The lowlands are characterized by *Acacia-Commiphora* thickets and shrublands, croplands, sisal farming, grassland,

and wildlife conservation areas harboring megafauna like elephants, giraffes, rhinos, buffalos (Pellikka *et al.*, 2018).

Western region comprised of Londiani junction, Mt. Elgon-Kapsokuong, Kitale school, Moi Barack's Eldoret, Eldoret Kapsabet junction, Equator and Kakamega forest which is the only rainfall forest in Kenya and its notably abundant in biodiversity but it faces the imminent risk of agricultural encroachment and various other human activities that pose a threat to its ecological integrity (Otieno & Analo, 2012). This region holds considerable significance in Kenya as it ranks among the country's vital bird habitats, while also serving as an essential source of hub for adjacent rural communities. These communities greatly rely on the area for a wide array of natural offerings, including medicinal plants, food sources, timber and various fibers (Otieno & Analo, 2012). For the Nairobi region, the areas of focus included Ngubi, Muguga in Kiambu and Ngong hills as shown in Figure 3.1. The hill is among the few significant wind resource areas near Nairobi (Nordman & Mutinda, 2016).

Forests ecosystem offer a wide array of habitats for bacteria such as soil, plant tissues and surfaces, rocks and streams among others, though bacteria seem to be more abundant in forest soil, litter and floor (Hardoim *et al.*, 2015). However, other than the pH which is the most important driver of bacterial community part of soil, factors such as type of soil, climatic conditions, organic matter content, biotic interactions, effect of vegetation and nutrient availability also affects the composition of bacterial communities in the forests ecosystem (Prévost-Bouré *et al.*, 2011). A study by (Zechmeister-Boltenstern *et al.*, 2011) revealed that tree species and composition significantly influence the diversity, distribution, function and structure of microbial communities in forests soil. This suggests the existence of crucial relationship between above-ground and below-ground processes. Such information could be utilized in the conservation of forests whose root network stabilizes huge amount of soil, bracing the entire ecosystem's foundation against water and wind erosion. Soil erosion has been reported to trigger life threatening problems such as landslides and dust storms that killed people in Pokot, Kenya (Krause & Nielsen, 2019).

The forests soil bacteria inhabit multiple habitats with certain properties such as litter, deadwood habitat, rhizosphere and bulk soil where they are shaped by biotic interactions and nutrients availabilities (Lladó *et al.*, 2017). Therefore, novel species might be present in different forests ecosystem and when bacterial taxa show a preference for specific niches with particular nutrient contents and organic matter quality, it indicates their ecological role or function (Fierer *et al.*, 2007).

However, several factors such as soil type, cultivation practices and plant host genotype are key drivers in shaping the unique soil microbial communities in forest ecosystems (Lakshmanan *et al.*, 2014). Therefore, the aim of this study was to investigate the microbial ecology including the composition, diversity and distribution of prokaryotes in specific forest soil ecosystems in Kenya. Additionally, we sought to identify potential abiotic factors influencing the distribution of prokaryotes. To place the obtained diversity in a global context, a comparison was made with amplicon data from forest biomes worldwide. The study sought to establish and document how Kenyan microbial diversity within forests ecosystem differs across different ecoregions, unravel its uniqueness compared to northern hemisphere thus the need to conserve. These forests have not been well studied, the sampling sites are unique and mostly under government protection.

1.2 Statement of the Problem

Alterations in forest land use are believed to profoundly affect soil microbial communities over an extended period (He *et al.*, 2017). Because of their sensitivity to changes in forest land-use, which may result in variations and extinctions of novel species due to environmental influence, leading to loss of biodiversity, thereby rendering the ecosystem more susceptible to disturbances and less resilient to environmental shifts (Romaniuk *et al.*, 2010).

Forest type conversion can induce a great shifts in soil microbial community through abiotic and biotic factors as well as species makeup, above and below ground litter, and quality and quantity of soil substrate, which are interconnected with the soil microbial community (Krashevska *et al.*, 2015) creating complex ecological interactions that influence the diversity and distribution of prokaryotes.

Moreover, tree species and composition strongly affect the structure and function of the microbial communities in soil forest as suggested by (Weand *et al.*, 2010). This is an indication that, there is a strong relationship between above -and below-ground processes that can affect the diversity and distribution on the forest soil microbial communities (Li *et al.*, 2014).

In addition, change in environmental variables like carbon, phosphorus, nitrogen, climate, soil acidity and vegetation, has also been shown to be a strong influence of soil microbial diversity and composition (Eldridge *et al.*, 2017). Nonetheless, the field of forest microbial ecology is firmly established, predominantly relying on culture-dependent techniques with only limited utilization of molecular methods. Consequently, there remains a significant gap in our understanding of forest microbial ecology, despite its potential biotechnological importance.

1.3 Justification of the Study

Forest ecosystems are increasingly being challenged to withstand disturbances and maintain their functions as the pace of change accelerates due to climate change and other human activities. Microorganisms play a crucial role in supporting forest ecosystem functions, and their recovery from disturbances is a key factor in determining the overall resilience of ecosystems. However, despite the numerous pressures associated with global environmental change, the impact of multiple disturbances on microbial community stability and the subsequent effects on ecosystem functions remain unclear and controversial (Philippot *et al.*, 2021).

Therefore, it is necessary to explore potent soil microbes for efficient nutrient recycling and to identify eco-friendly alternatives to reduce the use of chemical fertilizers and their negative impacts. In this context, maintaining soil fertility and crop productivity through natural microbial diversity could be the best approach to enhance nutrient bioavailability and improve soil health and productivity (Prasad *et al.*, 2021).

In addition, forest soil microorganisms react to various impacts of climate change, such as global warming, elevated CO2 levels, drought, anthropogenic nitrogen deposition, increased precipitation, and flooding. Addressing these challenges poses significant difficulties for researchers studying the soil microbiome. This review highlights the current understanding of how climate change affects living soil ecosystems in different forest climate-sensitive areas and the implications for vegetation-soil-climate feedbacks to helps in developing adaptive management practices to sustain forest health under changing conditions (Meena *et al.*, 2023). Therefore, this study demonstrates the importance of managing and conserving forest ecosystems, enhancing soil health, mitigating climate change, and harnessing the benefits of microbial processes for various applications.

1.4 Research Questions

1. How diverse are the prokaryotic communities and their population structure within selected forests ecosystem?
2. What is the alpha and beta diversity of prokaryotic communities within selected forests ecosystem?
3. How does the soil physicochemical parameters affect the prokaryotic diversity within the selected forests ecosystem?
4. How does the Kenyan soil microbiome differ in reference to other selected countries around the globe?

1.5 Hypothesis

1.5.1 Null Hypothesis

There is no significant variation in soil prokaryotic population structure within selected forests ecosystem in Kenya

1.6 Objectives

1.6.1 General Objective

To determine the diversity, distribution and abundance of prokaryotic communities within the selected forests ecosystem in Kenya

1.6.2 Specific Objectives

1. To evaluate the diversity and population structure of the prokaryotic communities in the selected forests ecosystem in Kenya.
2. To determine the alpha and beta diversity of prokaryotic communities in the selected forests ecosystem in Kenya.
3. To assess the effects of soil physiochemical parameters on prokaryotic diversity within the selected forests ecosystem in Kenya.
4. To assess the uniqueness of soil microbial communities in Kenya compared to the selected countries around the globe.

CHAPTER TWO

LITERATURE REVIEW

2.1 Microbial Diversity and Soil Functions

Microbes are unit or multicultural microscopic organisms and they are cosmopolitant in their distribution. They are widely distributed in bodies of living plants, animals, humans, mountains, water, hot springs, sea and soil (Jakobsson *et al.*, 2014). The pores and soil aggregates within and around them create microhabitats that support different microbial communities (Bach *et al.*, 2018). However, the differences in quantity and chemisty of organic substrates within the macrohabitats are likely to cause microbial comminity difference (Thies & Rillig, 2012). For instance, more organic carbon and greater concentrations of less chemically complex and new organic matter input are contained in large macroaggregates (Kong *et al.*, 2005). The environmental conditions within and between aggregates such as oxygen concentration can also vary, resulting in diverse niches that habor different guids of microoganisms (Stief *et al.*, 2016).

Microbial activity, including the secretion of extracellular substances like polysaccharides and proteins, helps in soil aggregation and stabilization. This contributes to soil structure formation, preventing erosion and improving water infiltration and retention (Ali *et al.*, 2024). They play a key role in nutrient cycling processes such as decomposition, mineralization, and nutrient uptake by plants.

Diverse microbial communities contribute to the efficient cycling of carbon, nitrogen, phosphorus, and other essential elements in the soil. They possess a wide range of metabolic capabilities, enabling them to degrade pollutants and contaminants in the soil through processes such as biodegradation and biotransformation (Sahu *et al.*, 2017). However, high microbial diversity can enhance the resilience of soil ecosystems to environmental disturbances such as drought, pollution, and land-use changes. Diverse microbial communities are better equipped to adapt and recover from disturbances, maintaining soil functions and ecosystem stability (Griffiths & Philippot, 2013).

2.2 Indicators of Soil Conservation

Indicator organisms serve as microbial indicator species of soil quality since these are very rare organism in the soil. For example, microbes involved in the process of conducting nitrification are good candidates of indicator microbial species since there are limited number of microbes that are involved in this process (Vermue *et al.*, 2013).

The indicator species provide early information in the soil degradation process so as to assist in land management. This can be done by understanding of soil ecology through research on all aspects that deal with forest soil as well as broadening the taxonomic knowledge and relationship between the diversity of soil microorganism and their ecological role, since there is over a thousand taxa of microorganisms in each gram of soil (Everard *et al.*, 2020).

2.3 Effects of Land Use on Soil Microbial Diversity

Use of the land can lead to change in soil nutrient cycle and soil microbial activity through heat,water and other enviromental conditions which can change the soil microbial community structure (Zhang *et al.*, 2019).Through application of animal manure and biosolids, wastewater treatment effluent and inappropriate disposal of unused medicines, antibiotics are introduced into the soil which can alter the structure and activities of micobial communities and the abundance of antimicrobial resistance genes (Semedo *et al.*, 2018). It has also been reported to alter microbial community composition with the cascading effects on net N_2O production, for example when the bacterial: fungi ratio is high as a result of antibiotics exposure incase the bacteria are selectively inhibited by the antibiotics (Semedo *et al.*, 2018).

2.4 Factors Affecting Soil Microbial Diversity

2.4.1 pH

Soil pH is a measure of the acidity and alkalinity in soils and it ranges from 0-14, with 7 being nutrial, below 7 acidic and above 7 alkaline (McCauley *et al.*, 2009).The pH influences soil-dwelling organisms which in turn, affects soil

conditions and plant health. The slightly acidic conditions are enjoyed by most plants, earthworms and microorganisms that convert nitrogen into forms that plants can use (Brady *et al.*, 2008).

2.4.2 Pesticide Effects on Microbial in Soil

Use of insecticides for the purpose of protecting the plants against harmful insects and to increase the crop yield can also influence the physiochemical properties and biological properties of soil, which will influence the number and metabolic activity of the soil microbial communities (Filimon *et al.*, 2015). For example, some pesticides have depressive effects on microbial like the butachlor which reduce the population of *Azospirillum* and aerobic nitrogen fixers in non-flooded soil. However, some of the pesticides stimulate the growth of the microbes and others do not have any effect at all, for example carbofuran stimulate increase in the population of *Azospirillium* and other nitrogen fixing in flooded and non-flooded soil (Dunivin & Shade, 2018).

2.4.3 Human Activity

Human agricultural practices like the addition of nitrogen and phosphorus to the soil beneath grasslands as a source of fertilizer tend to shift the natural coomunities of the soil microbial, tins may have unintended environmental consequences (Biederman & Harpole, 2013). The release of high amount of antibiotics into water and soil, creats a potential threat to all mirooganisms in the environment.

The relative abundance of different microbioal groups such as gram positive and negative bacterias as well as the overal microbial biomass can be altered when high amount of antibiotic are released into water and soil, creating a potential threat to all microoganisms in the environment, since it changes their enzymatic activity and ability to metabolize different carbon sources (Cycoń *et al.*, 2019).

2.5 The Significance of Soil Microorganisms

2.5.1 Role of Soil Microbes on Mineral Metabolism

The most active part of the soil is microbes that play an important role in transformation and storage of various nutrients as well as decomposition of mineral, organic matter and release of nutrients (Baldrian Petr *et al.*, 2012). They can be absorbed by providing a variety of nutrients and role of plant root systems (Sylvia *et al.*, 2005). Soil microbes have some effects in the agroecosystem, that is; they promote inorganic element to flow through transformation and promotion of system metabolic process and they can adjust as well as store soil nutrients since they contain elements like P, N, and C which are regarded as effective P, N, C source repositories (Vereecken *et al.*, 2016). Also due to the activities of the soil microbes, carbon and other mineral nutrient decomposition and cycling in soil ecosystems have denominated (Chu, 2018).

2.5.2 Role of Soil Microbes in Soil Structure Formation

The soil environment is a complex system that has a variety of small environments with different chemical and physical gradients as well as discontinuous environmental conditions. However, soil microbes play an important role in soil structure formation, in which they adapt to the micro-environment and interact with the other parts of the soil to produce various interactions (Bickel *et al.*, 2019). For instance, *Actinomycetes* produce mycelia that allow soil particles to bind, proving the content of it dominating in fertile soil than barren soil. In addition, different role of microbes differ in different soil types (Ding *et al.*, 2013).

The formation of soil structure is brought about by the combination of soil particles which is as a result of separation of microbes, microbial extracellular polysaccharides and the soil. Therefore, the formation and production of soil human is brought about by the activity of organic matter during polymerization process, which can reduce invasion of soil water for the maintenance of good soil aeration (Totsche *et al.*, 2018)

In addition, some of the soil microbes secrete gums, polysaccharides and glycoproteins which help in glueing the soil minerals together creating soil structure which is essential for plant growth.The fungal hyphae and plant roots ensure that soil aggregates are further boundtogether (Aleklett *et al.*, 2018).

2.5.3 Role of Soil Microbe in Soil Fertility

The different types of soil microbes are specialised in dicomposition of different types of organic matter lead into conversion of everything into humus, which is a dark brown-jelly-like substance that help the soil retain mosture and encourages the formation of soil strucure as well. It also suppresses plant diseases and its molecules are covered in negatively charged (anion) sites that bind to positively charged ions (cations) of the plant nutrients hence forming an important componet of soils cation exchange capacity (Kittredge, 2015).

The soil microbial communities such as bacteria are responsible in increasing soil fertility and provide nutrients to the soil to help on plant growth. They also soften the food in the seed, hence very important during the early stages of plant development in agriculture. Also, there are certain pesticides that give benefits to the crops that are developed using bacteria, for example; *Bacillus thuringiensis* (Babalola, 2010). Further more, farmers depends heavily on certain microbes like bacterias that have the ability to convert atmospheric nitrogen to ammonia. For example, *Rhizobium* bacteria in the roots of legumes.The biological nitrogen fixation contribute to about 60% of the nitrogen fixed on earth as compred to 25% that is manufactured from the fertilisers used during farming,which is expensive (Shridhar, 2012).

2.5.4 Role of Soil Microbes in Pest and Disease Control

The soil microbes help in pest and diseaes control, for example *Bacillus thuringiensis* bacteria ,is useful in control of caterpillar pests of crops and some of its strain are used in control of beetles and flies. The fungi, for example genus *Trichoderma* have has been developed as the biocontrol agents against fungal diseases of the plants, mainly roots disease (Sundh *et al.*, 2012).

2.6 Geospatial Data of the Selected Forests Ecosystem

Kenyan forests are located in high potential areas in which more than 70% of the population is settled (Bleher *et al.*, 2006). Some of these forests are located within Taita Taveta, Nairobi, Western, Aberdare and Mt. Kenya ecoregions. Mt. Kenya forest reserve is located to the east of the Great Rift Valley, along Latitude 0' 10' south and longitude 37' 20' east. It bestrides the equator in the central highland zones of Kenya and situated in two Forest Conservancies and five forest management zones. These include; Meru South, Nyeri, Meru Central, Embu in Eastern Conservancy and Kirinyaga in Central Highlands Conservancy (Kilonzi & Ota, 2019). From this forest the areas of interest were Mt. Kenya Castle Kimunge, Kyang'ondu, Kinondoni, Sirimon and Mt. Kenya State lodge.

The Aberdare Forest Reserve is located to the east of the Great Rift Valley, between 360 30' E, 00 05'S and 360 55'E, 00 450S (Kilonzi & Ota, 2019) .It's situated within Central Kenya in four counties, which include; Nyandarua, Nyeri, Kiambu and Murang'a. Its climate ranges is determined by the altitude (Wambugu, 2018). Taita Taveta region in southern Kenya ,is an area that is rich in biodiversity, comprising Taita Hills in Chawia, Ngangao and Vuria (Abera *et al.*, 2022).The Hills are part of the Eastern Arc Mountains, known for their rich endemism, and represent one of the world's biodiversity hotspots (Platts *et al.*, 2011). The hilltops are covered with indigenous evergreen montane forest, forest plantations like Susua plantation, agroforestry, and croplands (Pellikka *et al.*, 2009). The lowlands are characterized by *Acacia-Commiphora* thickets and shrublands, croplands, sisal farming, grassland, and wildlife conservation areas harboring megafauna like elephants, giraffes, rhinos, buffalos (Pellikka *et al.*, 2018).

Western region comprised of Londiani junction, Mt. Elgon-Kapsokuong, Kitale school, Moi Barack's Eldoret, Eldoret Kapsabet junction, Equator and Kakamega forest which is the only rainfall forest in Kenya and its notably abundant in biodiversity but it faces the imminent risk of agricultural encroachment and various other human activities that pose a threat to its ecological integrity (Otieno & Analo, 2012). This region holds considerable significance in Kenya as it ranks among the

country's vital bird habitats, while also serving as an essential source of hub for adjacent rural communities. These communities greatly rely on the area for a wide array of natural offerings, including medicinal plants, food sources, timber and various fibers (Otieno & Analo, 2012). For the Nairobi region, the areas of focus included Ngubi, Muguga in Kiambu and Ngong hills as shown in Figure 3.1. The hill is among the few significant wind resource areas near Nairobi (Nordman & Mutinda, 2016).

2.7 Vegetation Cover of the Selected Forests Ecosystem

High number of plant species, legumes, and natural vegetation types tend to support soil microbial communities with higher function. There is a significant correlation between the number of cultured bacteria and catabolic diversity of the bacterial community (Han *et al.*, 2007).

The Aberdare Ranges boast various vegetation zones such as the bamboo zone, the dense closed-canopy forest area, and the sub-alpine and alpine vegetation regions. The forest belt extends over a substantial portion of the range, with the majority of it designated as forest reserves.However, the distribution of vegetation zones and species is delineated by varying climatic zones and altitudes, primarily distinguished by differences in vegetation structure, coverage, and composition (Njeri *et al.*, 2018).

Mt. Kenya belongs to the series of volcanoes that occur along the fault lines of the Rift Valley system. The vegetation is briefly summarized by describing the zonation from the foothills upwards: Montane rain forests, Bamboo zone, Upper montane forest, Ericaceous zone, Páramo, and Nival zone (Niemelä & Pellikka, 2004). Because of abundant rainfall and minimal evaporation in the Taita Taveta forest ecosystem, the hills nourish rivers and streams that flow into the lowlands. Natural mist and secondary forests are common in the Taita Hills, alongside Arid and Semi-Arid Lands (ASAL), which feature grasslands, woodlands, and shrublands hosting Savannah species (Anyona & Rop, 2022). The Nairobi region boasted dense ground cover comprising acacia trees, grass, shrubs and mixed forest whereas Weatern region featured a diverse landscape of mixed forest with indegenous and exotic trees, rainfal frorest, grassland, planted trees and a variety of mixed forest vegetation.

Inconclusion, studying microbial diversity involves a range of methods to explore the vast array of microoganisms inhabiting various environments, for example culture based method where scientists isolate and culture microbes on various growth media to study their diversity. It allows for detailed analysis of individual strains, though it has limitations since many microbes cannot be cultured in laboratory conditions (Wang *et al.*, 2020).

The molecular techniques like the 16S rRNA sequencing is a method that targets the conserved 16S ribosomal RNA gene present in bacteria and archaea. By sequencing this gene, researchers can identify and classify microbes based on their genetic diversity (Cowan *et al.*, 2022).In addition ,next generation sequencing technique such as Illumina sequencing, allow for high-throughput sequencing of microbial DNA, enabling comprehensive analyses of microbial communities in various environments.

Bioinformatics techniques like phylogenetic analysis which is used to compare genetic sequences, researchers can reconstruct evolutionary relationships between different microbial taxa, provinding insights into their diversity and evolutionary history. The Operational Taxonomic Unit clustering groups similar sequences into OTUs facilitate the analysis of microbial diversity within samples and alpha diversity measures the diversity within a single sample whereas beta diversity compares the diversity between diffent samples, providing insights into community composition and structure (Tabish *et al.*, 2013).

CHAPTER THREE

MATERIALS AND METHODS

3.1 Study Site and Sample Collection

This study was conducted as part of a research consortium project that aimed to conduct a comprehensive survey of soil chemistry and microbiology across various regional and climatic zones in sub-Saharan Africa at a primary-scale (Cowan *et al.*, 2022). In Kenya, a microbiome survey of the soils within specific forests ecosystem (Nairobi region, Taita Taveta region, Western region, Mt. Kenya Forest region and Aberdare Forest region) was conducted, targeting forest regions for the study (*http://kws.go.ke/content/overview-0*). Each sampling site was carefully documented capturing data such as GPS location, elevation, vegetation present during sample collection, slope, general soil characteristics and site description. ArcGIS 10.8.1 (Environmental Systems Research Institute software application, 2020) was used to visualize and display the sample sites (Figure 3.1). The distribution and characteristics of the selected forests used in this study are summarized in Table 3.1.

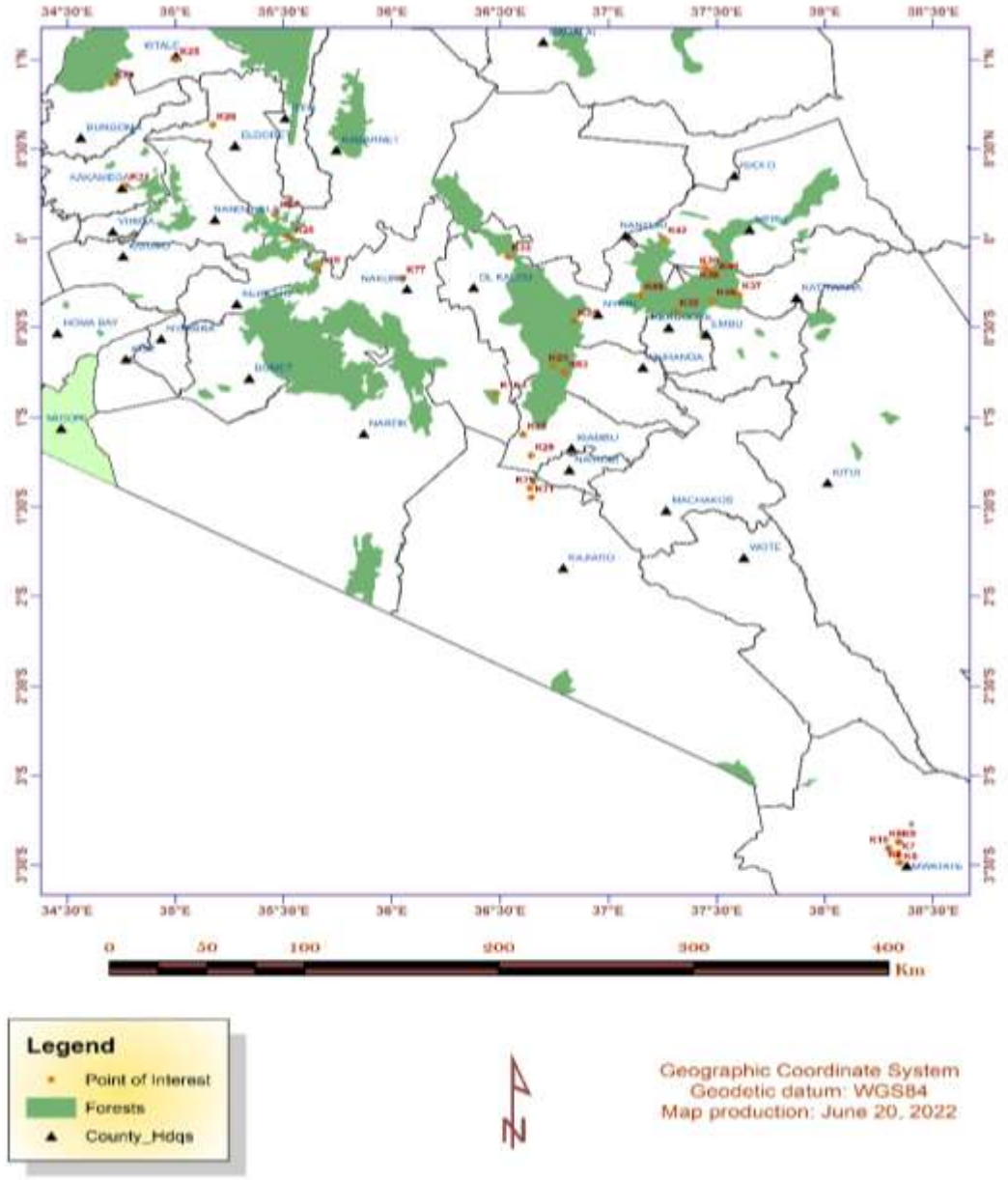


Figure 3.1: Kenyan Forest Sites Where Soil Samples Were Collected.

The samples are indicated sequentially from K5 to K77 within the map. The point of interest represents each sampling site.

Table 3.1: Parameters of Sampling Sites Analyzed in this Study

Sample	Forest Ecosystem	Latitude (S)	Longitude (E)	Elevation (m)	Vegetation (Time of Sampling)	General site Description	Weather (Time of Sampling)
K5	Chawia – Inside	S03°29'09.42'	E038°20'44.0'	1565	Highland mountain forest	Dense forest	Light shower, foggy and cloudy
K6	Chawia-periphery	S03°28'55.42'	E038°20'47.0'	1605	Busy plus pine trees	Litter from pine, organic	Drizzling, cloudy, foggy
K7	Susu plantation	S03° 26'31.7'	E038°20'15.3'	1671	Eucalyptus plantation & wattle trees	Red soil	Thick clouds-foggy
K8	Ngangao – lower	S 03°22'17.0'	E038°20'34.0'	1760	Montane-Tropical	Thick ground cover	Over cast
K9	Ngangao – Upper	S 03° 22'19.2'	E038°20'25.5'	1812	Mixed pine, Cyprus & indigenous tree	Heavy litter from needle leaves	Cloudy
K10	Vuria	S 03°24.316'	E038°20'17.66	1802	Montane	Thick ground cover	Sunny plus cloudy
K15	Ngubi	S 01° 05.769'	E036°36.441'	2226	Thick ground cover, mixed forest	Thick ground cover	Cloudy, dry and cold
K16	Longonot	S 00° 52.088'	E036°29.040'	2171	Acacia, grass and shrubs	Gullies, eroded soil	Windy, dry and sunny
K18	Londiani junction	S 00° 10.078'	E035°39.254'	2435	Mixed forest	Hilly with forest	Cool and cloudy
K21	Kakamega	0°17'43.5'N	34°46'24.2'E	1582	Rain forest	Natural forest	Cool and wet
K23	Arbadare - Kihinganda	S 00° 42.893'	E 036°44.805'	2700	Bamboo forest	Thick litter -Bamboo leaf cover	Cloudy, Cool and Wet
K24	Mt. Elgon-Kapsokwony	N0° 51.49'	E 034°42.10'	2069	Mixed forest with indigenous exotic trees	Hilly, rocky slopes and valleys	Rainy
K25	Kitale School	S0°59' 43.6'N	35°00'13.8'E	1914	Planted forest	School and firming activities	Sunny
K26	Moi Barracks Eldoret	0°37'56.2'N	35°10'24.5'E	1988	Grass land	Rock with natural grasses	Cloudy, Windy and dry
K27	Eldoret-Kapsabet junction	N00° 08'28.4'	E035°28'00.4'	2511	Natural forest	Forest with native trees	Cloudy and wet
K28	Equator Londiani	N0°00'00.0'	E035°32'05.6'	2487	Mixed exotic	Highly depleted forest	Cloudy and Rainy
K29	Muguga – Kiambu	S 01° 12.835'	E036°38.549'	2036	Forest	Thick ground cover	Cloudy
K33	Aberdare Pesi - Ndaragwa	S 00° 06.055'	E 036°32.189'	2499	Very thick natural forest with thick shrub	Thick indigenous vegetation	Cool and cloudy
K34	Aberdare Tosha Forest	S 00° 27.740'	E 036°50.563'	2302	Indigenous forest	Thick vegetation	Cool evening`
K35	Mt. Kenya-Castle Kimunge	S 00° 24.490'	E 037°18.682'	1894	Forest mountain ecosystem	Thick ground and forest cover	Rainy, cloudy and foggy
K36	Mt. Kenya - Irangi	S 00° 20.832'	E 037°28.929'	2002	Indigenous mixed	Thick ground cover	Raining
K37	Mt. Kenya - Kyang'ondu	S00°18'48.0"	E 037°36'00"	1622	Indigenous mixed forest	Invasive shrub	Worm, cloudy and wet

K38	Mt. Kenya- Kinondoni	S 00° 09.683’	E 037°26.769'	2950	Moorland Rose wood trees	Grassy-moorland	Rain showers, foggy and cloudy
K39	Mt. Kenya -Kinondoni	S00°10.33’	E 037°27.157'	2845	Bamboo forest	Bamboo	Raining, cloudy and Foggy
K40	Mt. Kenya - Kinondoni	S 00° 12.250’	E 037°29.658'	2363	Transition based	Bushy - base of forest succession	Raining, cloudy and foggy
K42	Mt. Kenya - Sirimon	S00°00.16’	E 037°15.34'	2718	Thick bush-mixed shrub	Thick ground Cover	Cloudy and dry
K63	Aberdare-Githika block	S 00°45.020’	E 036°47.879'	2368	Indigenous forest	Thick forest	Cool and Wet cloudy evening
K66	Mt. Kenya- state lodge	S 00° 19.089’	E 037°08.900'	2216	Mountainous forest with podo tress	Thick forest	Sunny, Cool-clean sky
K70	Ngong hills nature reserve	S 01° 23.763’	E 036°38.254'	2394	Shrubs and tall trees	Grassy	Sunny, windy, clear sky
K71	Ngong hills Corner Baridi	S 01° 26.842’	E 036°38.682'	2194	Forest, grass & shrubs	Grassy and shrubs	Sunny, windy, clear sky
K77	Menengai crater	S 00° 13.397’	E 036°03.010'	2056	Crater-shrubs patching on rocks	Grass and bushes rocky volcanic	Cloudy cool evening

The sampling process involved retrieving four separate topsoil samples, each weighing 200 grams, from a depth of 0-5 centimeters, at approximately 50 meters apart within each site (Figure 3.2). To create each working sample, a composite of four pseudo-replicate samples, each weighing 50 grams, was collected from the corners of a one square meter virtual quadrant. Subsequently, each sample was carefully placed in a separate Whirl Pak bag, labeled accordingly, and stored at 4ºC for both nucleic acid extraction and soil physiochemical analysis.

These samples were later categorized into specific regions based on their geographical location on the Kenyan map as follows: Aberdare (Sample K23, K33, K34, K63 and K77); Mt. Kenya (K35, K36, K37, K38, K39, K40, K42 and K66); Nairobi (K15, K16, K29, K70 and K71); Taita Taveta (K5, K6, K7, K8, K9 and K10) and Western region (K18, K21, K24, K25, K26, K27 and K28).

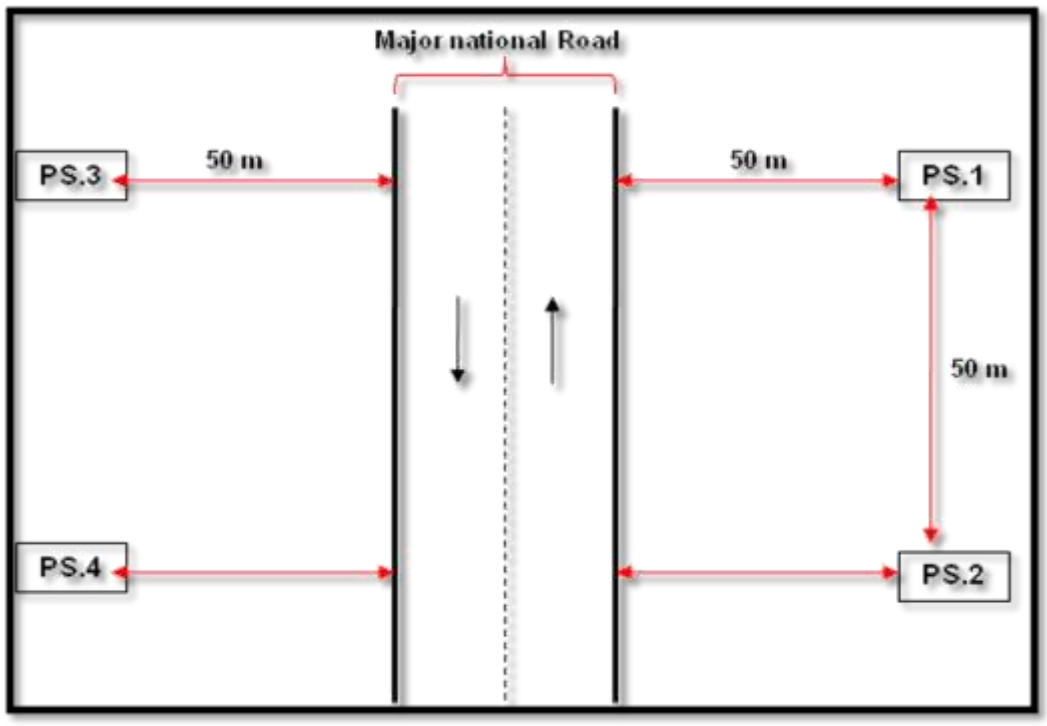


Figure 3.2: Sketch Showing Sample Collection Strategy

3.2 Soil Physicochemical Characteristics

Soil physicochemical characteristics (Table 4.1) were analyzed following the manufacturers procedure as outlined by AgriLASA (2004) protocol. Soil pH was measured using the slurry method at a soil -to-water ratio of 1:2.5, and the pH of the resulting supernatant was recorded using a calibrated bench top pH meter (Crison Basic, +20, Crison, Barcelona, Spain). The concentrations of soluble and exchangeable sodium (Na), potassium (K), carbon (C), magnesium (Mg), and phosphorus (P) were determined using Mehlich 3 test (Mehlich, 1984). The quantification of extractable ion concentration was carried out using high-performance Inductively Coupled Plasma Optical Emission Spectrometry (ICP-OES) instrument called SPECTRO Genesis manufactured by SPECTRO Analytical Instruments GmbH & Co. KG from Germany, (Aşçı *et al.*, 2015). The soil particle size distribution (sand/silt/clay percent) was determined using the Hydrometer method (Bouyoucos, 1962). Total nitrogen (TN) and soil organic carbon (TOC) were measured using the catalyzed high temperature combustion method, specifically the Dumas method (Bremner, 1996). The Enhanced Vegetation Index-2 (EVI2) was

obtained from the NASA Land Processes Distributed Active Archive Center's (LP DAAC) VIIRS Vegetation Indices dataset (Didan, 2018) with a resolution of 500 meters.

3.3 Prokaryotic DNA Extraction and 16srrna Gene Sequencing

Total DNA was extracted from soil samples using the DNeasy PowerSoil Kit (QIAGEN, USA) following the manufacturer's instructions with slight modifications. Before the final elution step, the elution buffer C6 was pre-heated to 55°C for 10 minutes, and the DNA was eluted using 70 µl of the elution buffer. Subsequently, DNA concentration and purity were assessed using the Nanodrop 2000 (Thermo Fisher, USA). Sequencing of the V4/V5 16S rRNA gene was carried out using the 515F (5'-GTGYCAGCMGCCGCGGTAA-3') and 909R (5'-CCCCGYCAATTCMTTTRAGT-3') primers, following a method described previously (Caporaso *et al.*, 2012; Cowan *et al.*, 2022). Prior to library preparation, the target regions were amplified using the Hot Star Taq Plus Master Mix Kit (Qiagen, USA), and purification was conducted using calibrated Ampure XP beads (Beckman Coulter Life Sciences, USA). Sequencing of the pooled samples were carried out at MR DNA laboratory (www.mrdnalab.com, Shallowater, TX, USA) using MiSeq 2 x 300 bp V3 instrument following the guidelines provided by the manufacturer.

3.4 Sequence Analysis and Taxonomic Classification

The raw amplicon sequence reads obtained were filtered, trimmed, and clustered into unique amplicon sequence variants (ASVs) using the QIIME2 pipeline (Dacal *et al.*, 2022). Briefly, raw sequences were demultiplexed, quality checked and a feature table constructed using Divisive Amplicon Denoising Algorithm 2 (DADA2) pipeline (Callahan *et al.*, 2016) within QIIME2 (Bolyen *et al.*, 2019). The raw sequences were denoised and chimeras removed. Sequences that were less than 200 base pairs in length after phred20- base quality trimming, as well as those containing ambiguous base calls and homopolymer runs exceeding 6 base pairs, were excluded from the analysis. Both forward and reverse reads were truncated at 324 base pairs. Denoising statistics were calculated and representative sequences were selected to

create the ASVs feature table. A total of 1,944,316 high quality sequence reads were obtained, which were then clustered into 41,901 ASVs at 3% genetic distance threshold. Representative sequences were aligned using MAFFT program (Rozewicki *et al.*, 2019) and highly variable regions were masked to reduce the noise during phylogenetic analysis (Katoh *et al.*, 2013). Phylogenetic tree was created and rooted at midpoint on QIIME2 to help in downstream analysis. Taxonomic classification of ASVs was done using QIIME feature-classifier (Bolyen *et al.*, 2019) against the untrained SILVA 138.1 (release 2022.2) (McLaren *et al.*, 2021; Quast *et al.*, 2012). Demultiplexed high-quality sequence reads were curated and deposited in the National Centre for Biotechnology Information (NCBI) Sequence Read Archive (SRA), as BioProject ID: PRJNA851255 and study accession numbers are available for download at http://www.ncbi.nlm.nih.gov/bioproject/851255.

3.5 Data Processing of Amplicon Datasets from Other Countries

Sequence datasets from selected forests around the globe were downloaded from publicly available databases (accession numbers at Table 3.1) and processed using the QIIME2 pipeline as described above. Raw reads from the downloaded datasets spanned the 16S rRNA gene hypervariable regions v3-v4, v4, and v4-v5, depending on the study. To keep the sample sizes between countries comparable, a subset of between 28 to 30 samples was chosen for each dataset (This was based on Kenyan sample size 31(+ or -3). To accommodate the variable quality scores of the different datasets, quality threshold was set to 20 and all reads were truncated at 220 bps. After DADA2 processing, the resulting representative sequence file and ASV table were merged with the Kenyan dataset. Read counts for the combined dataset ranged from 10877 to 346157 reads (**Figure 3.3**). The merged representative sequence file was taxonomically annotated using the un-trained SILVA database 138.1 (release 2022.2), (McLaren *et al.*, 2021; Quast *et al.*, 2012).

Table 3.2: Accession Numbers from Publicly Available Databases of Different Countries

Dataset			
Bio project	**Country**	**Number of samples**	**Reference**
PRJNA315018	Canada	30	https://doi.org/10.1093/femsec/fiw239
PRJNA431064	China	30	https://doi.org/10.3389/fmicb.201
PRJNA701565	Czech Republic	30	https://doi.org/10.1111/ele.13896
PRJEB38904	Estonia	30	https://doi.org/10.3389/fmicb.20
PRJNA823643	Finland	30	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA823643
PRJNA291212	Kenya	31	https://doi.org/10.5281/zenodo.7827432 (This study)
PRJNA512531	New Zealand	30	https://doi.org/10.1111/1462-2920.14792
PRJNA595488	Switzerland	28	https://doi.org/10.3389/ffgc.2020
PRJNA308872	USA	30	https://doi.org/10.1111/1462-2920.14303

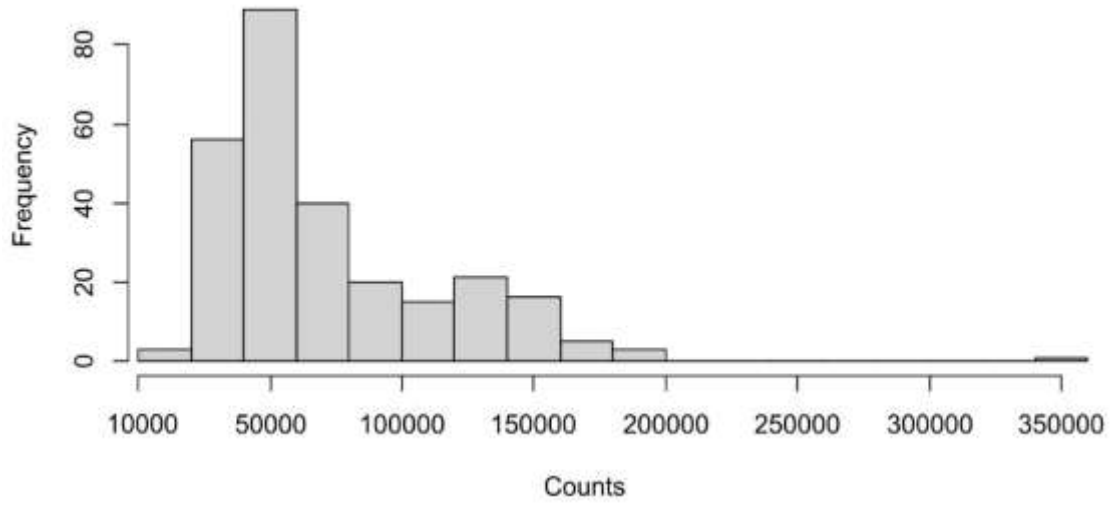


Figure 3.3: Read Counts for the Combined Dataset of Sequences Downloaded from Publicly Available Databases.

3.6 Statistical Analysis

ASVs table and associated metadata were exported from QIIME2 and imported to RStudio for modification for analysis with the phyloseq package (version 1.36.0) in RStudio (McMurdie & Holmes, 2013; Team, 2013). The taxonomy table was merged with the feature table, and the relative abundance and bar plots were plotted using the ggplot2 package (version 3.3.5) (McMurdie & Holmes, 2013; Team, 2013). The normality of the dataset was first tested with the Shapiro-Wilk test (Royston, 1982). The Kruskal-Wallis Rank Sum test (Hollander & Wolfe, 1973) was subsequently used to calculate the significance of mean differences in soil variables between forest soil samples (adj. p. value = 0.07176). Tukey *post hoc* analysis test (Armstrong & Hilton, 2004) were used to compare significant differences between regions where soil environmental variables were normally distributed (adj. p. value = 0.001657).

Significant differences in soil physicochemical characteristics were calculated using the stats package version 3.6.2 (Team, 2021) in RStudio version 4.0.3 (RStudio, 2020). The distribution of soil physiochemical variables across different forest sites was calculated on log-standardized data using the "*decostand*" function (Venables & Ripley, 2002) from vegan package version 2.5.7 (Jari *et al.*, 2020), which performs principal component analysis of the data (PCA) (Venables & Ripley, 2002). The resulting distance matrix between samples was plotted in a PCA plot, with the projected direction and magnitude of the distribution for each variable plotted in a separate loading plot. The hmisc (version 4.5) package (Team, 2021) was subsequently used to evaluate Pearson correlations (Hollander & Wolfe, 1973) between variables (adj. p-value < 0.01), which were plotted as a bubble graph using the corrplot (version 0.9) package (Team, 2021).Biodiversity metrics (alpha diversity) and community structure dissimilarity (beta diversity) were calculated using the vegan (version 2.5.7) (Jari *et al.*, 2020) and phyloseq (version 1.16.2) (McMurdie & Holmes, 2013) packages in RStudio. Observed richness, Inverse Simson (Chao *et al.*, 2014) and the Shannon indexes (Shannon, 1949) were used as metrics for alpha-diversity (Chao *et al.*, 2014). The prokaryotic ASV table was split into Archaea and Bacteria using the "subset_taxa" function in phyloseq before calculating the diversity indexes. Differences in alpha-diversity between designated

regions were assessed as described for the soil physicochemical variables. Beta-diversity index of each soil sample was calculated from the log-transformed ASV tables using the "vegdist" function in vegan, based on Bray-Cutis distance estimation method (Bray & Curtis, 1957). Ordination of the beta-diversity scores was plotted on a principal component analysis plot (PCoA) (Jolliffe, 2002), and the significance of beta-diversity dissimilarity between forest regions was calculated using Permutational Mult ivariate Analyses of Variance (PERMANOVA) (Anderson, 2005) with 999 permutations. Comparison of beta-diversity distribution between the samples of different countries datasets was also performed using the methodology described above.

The environmental drivers of prokaryotic community structure were estimated by Redundancy analysis (RDA) (McArdle & Anderson, 2001). The soil physiochemical dataset was z-score standardized and tested for multicollinearity using the "vif" function from the car (version 3.0.11) package (Dormann *et al.*, 2013). The best models for explanatory variables were calculated using forward step-wise regression model selection method with the ordistep () function in the vegan package, with 1000 permutations, and significant variables with vif values above 10 were removed. The significance of the best fitted models and each predictor variable in the model were calculated using the ANOVA permutation test, (Legendre *et al.*, 2011) with 1000 permutations. The relative taxonomic abundances of prokaryotic taxa were compared between regions using Linear Discriminant Analysis (LDA) effect size (LEfSe) algorithm (Segata *et al.*, 2011) for high-dimensional biomarker discovery and explanation of differentially abundant organisms. This analysis was implemented using the package Microbiome Marker in RStudio (Cao *et al.*, 2022) . Differences were analyzed using Kruskal–Wallis sum-rank test (Kruskal & Wallis, 1952) to detect significant differentially abundant taxa at genus level (adj. P value= 0.011541). The biological consistency was investigated using a set of pairwise tests among genera using the Wilcoxon rank-sum test (Mann & Whitney, 1947) (Wilcoxon, 1945), with an LDA threshold of 2. The same LDA method was used to detect differently abundant taxa across the country datasets.

3.7 Ethical Approval

Research authorization was obtained from the National Commission for Science, technology and Innovation (NACOSTI)) and permission to conduct research in the selected forests ecosystem, Access Permit, and Export Permit was obtained from the Kenya Wildlife Service (KWS) and National Environmental Management Authority (NEMA) while phytosanitary certificate was obtained from Kenya Plant Health Inspectorate Authority (KEPHIS) (Appendix IV, V, VI & VII).

CHAPTER FOUR

RESULTS

4.1 Soil Physicochemical Properties for the Different Sites

In this study, prokaryotic community composition was assessed in 31 soil samples collected from the selected forests ecosystem in Kenya. Soil physicochemical properties and macro-environmental variables were used to interpret the microbial community composition variations. The soil sample with the highest acidity was from site K23 in the Aberdare region (Aberdare Forest, Gatare). Conversely, the majority of the Taita Taveta area exhibited high acidity, except for site K10 (Vuria). In contrast, the Nairobi, Western, and Mt. Kenya regions had alkaline soils, with the sample from site K66 at Mt. Kenya being the most alkaline (Table 4.1).

The nutrients present in the sampling sites included C, K, Ca, Mg, Na, P, Mn, Fe, and Al. Different sites exhibited high concentrations of various nutrients: site K33 had the highest carbon concentration at 8.31%; site K23 had a potassium concentration of 21.06 Mg/Kg; site K25 had a calcium concentration of 85.68 Mg/Kg; site K7 had magnesium and sodium concentrations of 54.70 Mg/Kg and 23.63 Mg/Kg, respectively. Additionally, site K38 had a phosphorus concentration of 25.99 Mg/Kg; site K29 had a manganese concentration of 350.37 Mg/Kg; site K9 had an iron concentration of 34.20 Mg/Kg; and site K23 had an aluminum concentration of 1938.67 Mg/Kg (Table 4.1). The various soil types—sand, silt, and clay—were identified in different sampling sites with specific percentage distributions. For example, sample K8 was composed of 92% sand, sample K29 contained 43% silt, and sample K28 consisted of 24% clay (Table 4.1).

Table 4.1: Soil Physiochemical Characteristics as an Influence to Soil Prokaryotic Communities

Sample No.	pH	C %	K (mg/kg)	Ca (mg/kg)	Mg (mg/kg)	Na (mg/kg)	P (mg/kg)	Mn (mg/kg)	Fe (mg/kg)	Al (mg/kg)	Sand (%)	Silt (%)	Clay (%)
K5	4.09	5.37	2.38	74.64	21.29	1.70	17.58	65.20	187.56	769.48	83	2	15
K6	4.53	3.13	2.49	77.15	18.40	1.97	9.63	15.18	218.82	1097.39	72	14	14
K7	3.75	4.96	14.75	6.91	54.70	23.63	3.82	8.79	201.52	1429.14	84	5	11
K8	4.97	2.85	1.66	82.59	14.81	0.93	9.62	39.50	110.16	942.03	92	2	6
K9	4.17	5.32	2.33	74.26	20.90	2.51	4.94	18.44	341.20	1202.48	89	2	9
K10	6.41	2.67	4.91	71.29	23.28	0.51	10.71	44.89	75.80	640.78	62	19	20
K15	6.33	2.61	6.65	76.27	16.57	0.51	1.36	142.45	105.12	382.72	32	47	20
K16	6.57	0.67	12.83	70.34	14.78	2.04	2.70	9.23	110.34	399.06	77	11	12
K18	6.23	2.2	6.02	75.03	18.05	0.89	1.45	170.62	112.95	527.64	40	41	20
K21	5.96	4.51	0.96	82.71	16.02	0.31	1.81	122.55	76.39	372.02	65	20	15
K23	3.42	6.57	21.06	41.49	23.80	13.65	11.28	7.17	251.99	1938.67	84	1	15
K24	6.41	2.15	2.67	71.36	25.59	0.37	2.50	71.90	85.08	370.97	56	23	21
K25	6.43	2.47	2.83	85.68	11.05	0.44	8.55	65.51	53.99	378.77	77	9	14
K26	5.72	1.66	8.70	60.74	29.32	1.25	1.81	125.35	68.75	782.02	50	29	21
K27	6.14	2.73	7.87	66.02	25.60	0.51	3.12	99.06	62.64	537.71	42	36	23
K28	5.01	3.09	7.95	71.40	19.87	0.78	6.30	152.00	100.63	853.52	34	42	24
K29	6.57	2.77	4.50	81.17	13.98	0.35	0.70	350.37	74.19	419.71	39	43	18
K33	6.84	8.31	4.53	70.65	24.57	0.25	9.94	142.83	56.18	207.55	63	17	20
K34	4.39	6.29	5.81	70.89	22.04	1.26	16.67	69.88	124.25	1680.56	75	5	20
K35	4.61	5.5	6.59	62.08	29.49	1.83	10.02	22.39	77.24	1890.99	63	22	15
K36	3.92	5.62	9.57	66.57	21.86	2.00	9.57	228.86	134.02	1395.40	72	7	21
K37	5.2	3.07	4.46	57.30	37.88	0.35	4.67	60.10	70.50	693.72	68	11	21
K38	5.37	3.48	5.92	81.86	11.64	0.58	25.99	41.91	145.38	959.44	67	15	18
K39	5.29	6.42	5.47	79.91	13.01	1.60	4.42	36.97	94.31	1705.09	58	20	21
K40	3.9	6.43	30.77	42.83	14.05	12.35	8.02	12.32	46.60	2196.58	46	30	24
K42	6.69	6.13	4.13	72.20	23.31	0.35	1.43	149.20	69.00	441.51	18	68	14
K63	4.25	6.54	3.97	78.09	16.38	1.57	17.58	67.79	276.99	1634.29	60	6	34
K66	6.84	3.74	3.97	84.08	11.63	0.31	12.37	248.79	88.58	283.82	44	32	24
K70	6.37	4.14	4.89	75.33	19.29	0.50	42.17	119.14	120.74	820.01	32	45	23
K71	6.83	3.25	6.73	76.79	16.07	0.40	92.35	256.12	97.02	459.00	25	45	30
K77	6.56	4.78	3.82	81.52	13.60	1.05	0.65	26.90	155.65	1496.84	27	54	19

4.2 Data Preprocessing

After demultiplexing, quality filtering, denoising and removal of potential chimeras, a total of 1,944,316 high quality sequence reads (non-chimeric sequences) were obtained from soil samples collected from selected forests ecosystem. Samples from Mt. Kenya had the highest number of reads at 522,111, followed by Taita Taveta with 487,935 reads, the Nairobi region with 337,373 reads, Western with 301,478 reads, and the least reads were from Aberdare, totaling 295,419. The number of input sequences, filtered sequences, denoised sequences and non-chimeric sequences were distributed into samples as shown in table 4.2.

Table 4.2: Distribution of High-Quality Sequences across Samples

Sample ID	Raw Sequences	Filtered Sequences	% Raw sequences passed filter	Denoised Sequences	Non-chimeric Sequences	%non-chimeric Sequences
K5	85544	82110	95.99	77410	74672	87.29
K6	100448	96249	95.82	90867	88435	88.04
K7	104387	99545	95.36	94723	90394	86.6
K8	135276	129263	95.56	121284	117657	86.98
K9	35826	34261	95.63	30605	29521	82.4
K10	101060	96677	95.66	88697	87256	86.34
K15	90273	86144	95.43	79542	77702	86.07
K16	101571	96597	95.1	89780	86921	85.58
K18	95254	91285	95.83	83651	82204	86.3
K21	72391	69314	95.75	63903	63156	87.24
K23	37053	35455	95.69	32700	31646	85.41
K24	14548	13939	95.81	11732	11698	80.41
K25	39688	38239	96.35	34262	33782	85.12
K26	53593	51315	95.75	46834	45988	85.81
K27	41098	39524	96.17	35201	34857	84.81
K28	35407	33986	95.99	30310	29793	84.14
K29	44352	42547	95.93	38381	37816	85.26
K33	76113	72845	95.71	66479	64809	85.15
K34	70828	68119	96.18	62753	61451	86.76
K35	78327	75057	95.83	69679	67789	86.55
K36	82735	79092	95.6	74393	72684	87.85
K37	67238	64513	95.95	59151	58228	86.6
K38	67270	64701	96.18	59080	57676	85.74
K39	79895	76792	96.12	70896	68584	85.84
K40	74286	70944	95.5	65794	63811	85.9
K42	76805	73794	96.08	67885	66581	86.69
K63	75786	72599	95.79	66384	63772	84.15
K66	76729	73562	95.87	67903	66758	87
K70	83146	79712	95.87	73125	71360	85.82
K71	74810	71467	95.53	66249	63574	84.98
K77	87659	84347	96.22	75675	73741	84.12

These were clustered into 41,901 ASVs at 3% genetic distance. Composition and diversity assessment of prokaryotic communities within the selected forests' ecosystem displayed Vuria and Ngangao forests in Taita Taveta to harbor the highest number of ASVs, 1838 and 1832 respectively. These were followed by samples collected from Menengai Crater (1795), Londiani junction (1720), and Longonot forest with 1648 ASVs (Table 4.3). Soils displaying the lowest number of ASVs included Mt. Elgon Kapsokwony with 386, Aberdare-Kihinganda (639), Ngangao upper forest (720), Kitale School (780) and Equator Londiani with 839 ASVs (Table 4.3).

4.3 Taxonomic Composition of Prokaryotic Taxa within Selected Forest Ecosystems.

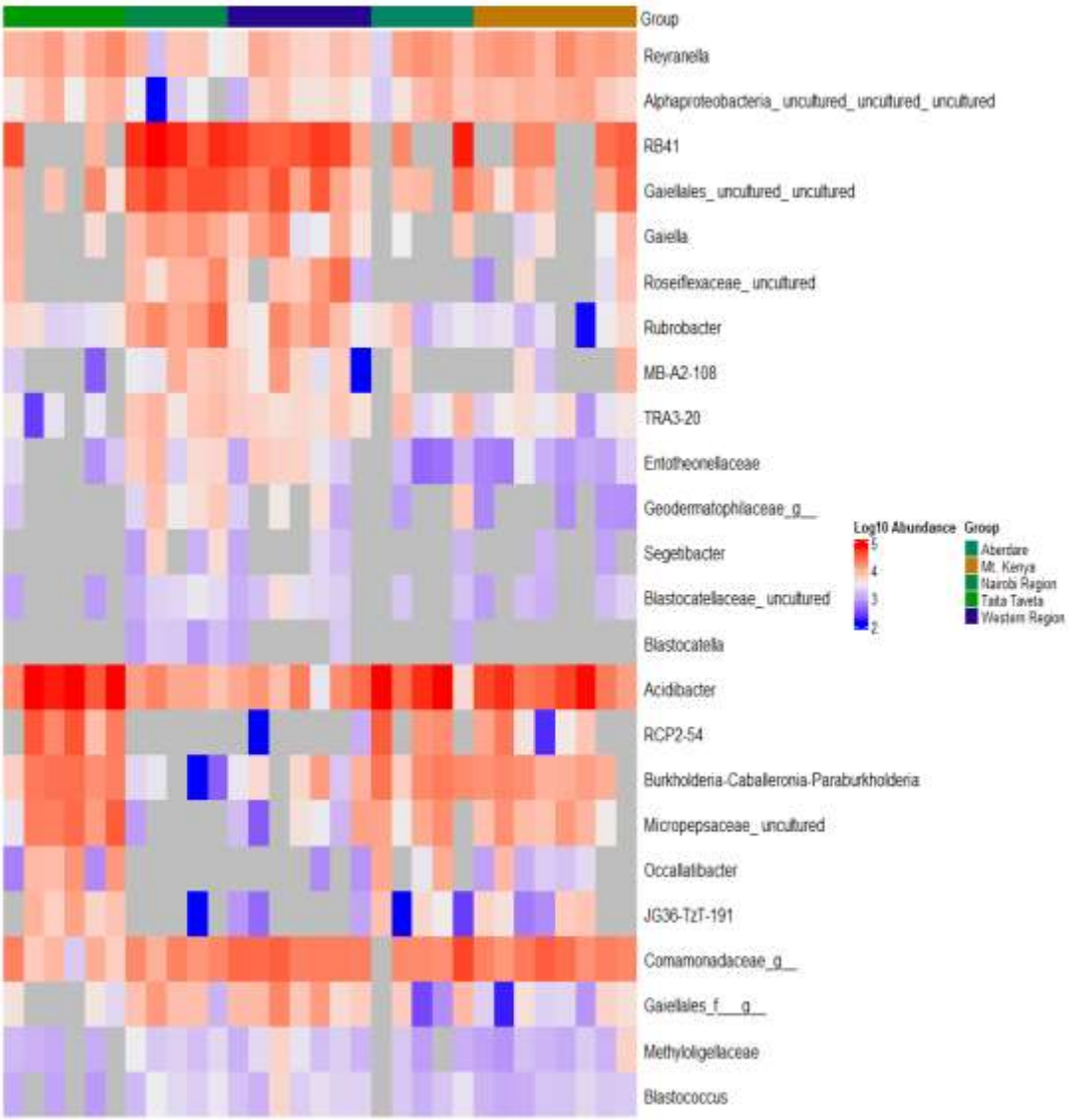
Composition and diversity assessment of prokaryotic communities within the selected forests ecosystem displayed the 41,901 ASVs obtained from the high-quality sequences to comprise of Bacteria (83.41%), Archaea (14.58%) and unassigned phylum which scored 2.08% (Appendix 1). This suggests that members of bacterial kingdom are the most abundant taxa within the selected forests ecosystem. Taxonomic classification of the obtained ASVs were assigned to a Kingdom, 2 Phyla, 120 Classes, 280 Orders, 450 Families, 873 Genera and 2313 Species within selected forests ecosystem (Table 4.3).

4.4 Bacterial Taxonomic Composition

To evaluate the diversity and population structure of the prokaryotes between the different forest areas, linear discriminant analysis (LDA) effect size (LEfSe) was used to detect prokaryotic taxa that were differentially abundant within and between soil samples. In a comparison of samples from the five forest regions (Aberdare, Mt. Kenya, Nairobi, Taita Taveta and Western), several taxa were identified as differentially abundant (P adj.= 0.001998): 6 genera(*Acidibacter, Burkholderia-Caballeronia-Paraburkholderia,RCP2-54, Reyranella, RB41 and Occallatibacter*) in Taita Taveta, 4 (*RB41,Rubrobacter,Gaiella and Acidibacter*)in Nairobi, 3(*Reyranella, Acidibacter* and *Burkholderia-Caballeronia-Paraburkholderia*) in Mt. Kenya, 2 (*Burkholderia-Caballeronia-Paraburkholderia* and *Acidibacter*) in

Aberdare region and *RB41* in Western (Figure 4.1 a). The genus *Acidibacter* was over-represented in Taita forest soils, possibly due to the acidic soil pH observed in this region. *RB41* was over-represented in western and Nairobi region. *Burkholderia-Caballeronia-Paraburkholderia taxa,* which typically have a very broad ecological diversity and metabolic versatility were the most abundant in Aberdare Forest soils, while *Reyranella* was the most abundant in Mt. Kenya region soil samples (Figure 4.1a).

a



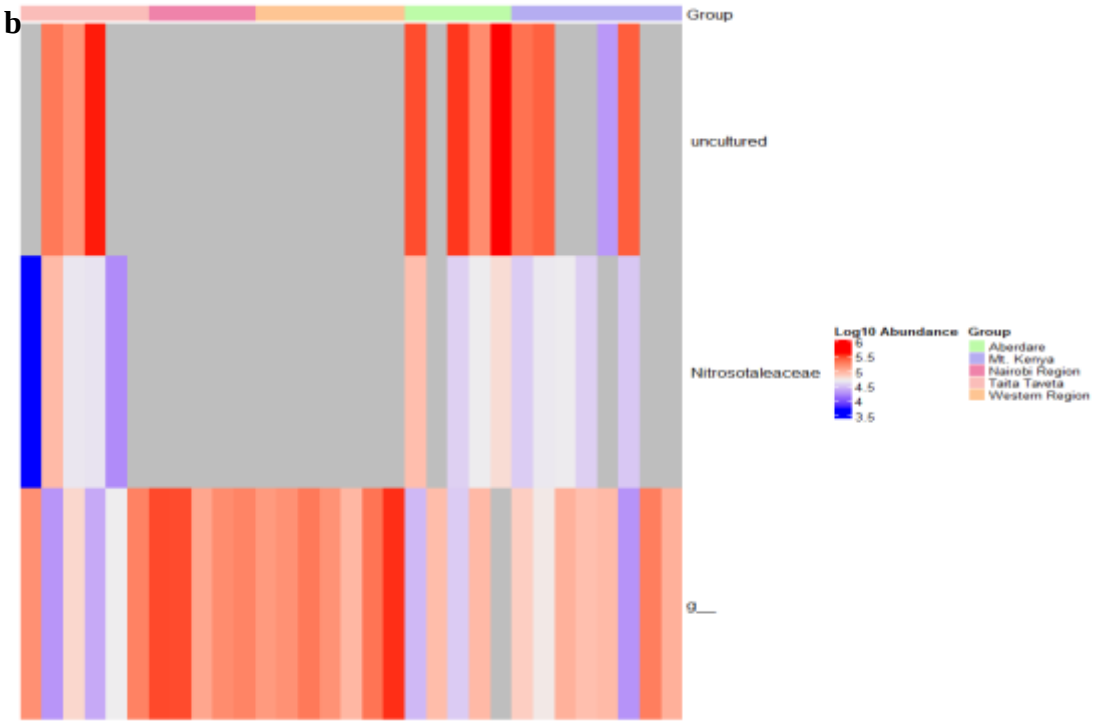


Figure 4.1 a and b: Lefse Plot at the Last Known Taxon of Prokaryotic Communities within Soil Samples Collected from Selected Forest Regions. a represents LEfSe analysis of Bacterial communities while b represents LEfSe analysis of Archaea communities within soil samples collected from the five regions in Kenya. The bars represent the effect size (LDA) for a particular taxon in each forest region. The length of the bar represents a log10 transformed LDA score. The colors represent the group in which that taxon was found to be more abundant compared to the other groups. f_ indicates that the ASV was not able to be classified to a family while g_ indicates that the ASV was not able to be classified to a genus.

Bacterial groups were the most abundant within the forest's ecosystem comprising of 34 Phyla (Appendix 1). These were further grouped into 104 Classes, 255 Orders, 418 Families, 835 Genera and 2238 species. Analysis of Bacterial diversity in forest soil samples indicated the presence of 34 phyla, of which 12 were dominant (i.e., represented by >1% of ASV reads in at least 87% of the samples) (Figure 4.2a). The most abundant of these was *Proteobacteria* (30.3% mean relative abundance), followed by *Acidobacteriota* (23.4% mean relative abundance) and *Actinobacteria* (13.1% mean relative abundance) (Figure 4.2a). Other major bacterial phyla included

Myxococcota, *Gemmatimonadota*, *RCP2-54*, *Firmicutes*, *Methylomirabilota*, *WPS-2*, *Nitrospirota* and the unknown bacterial phylum among others as shown in Figure 4.3.

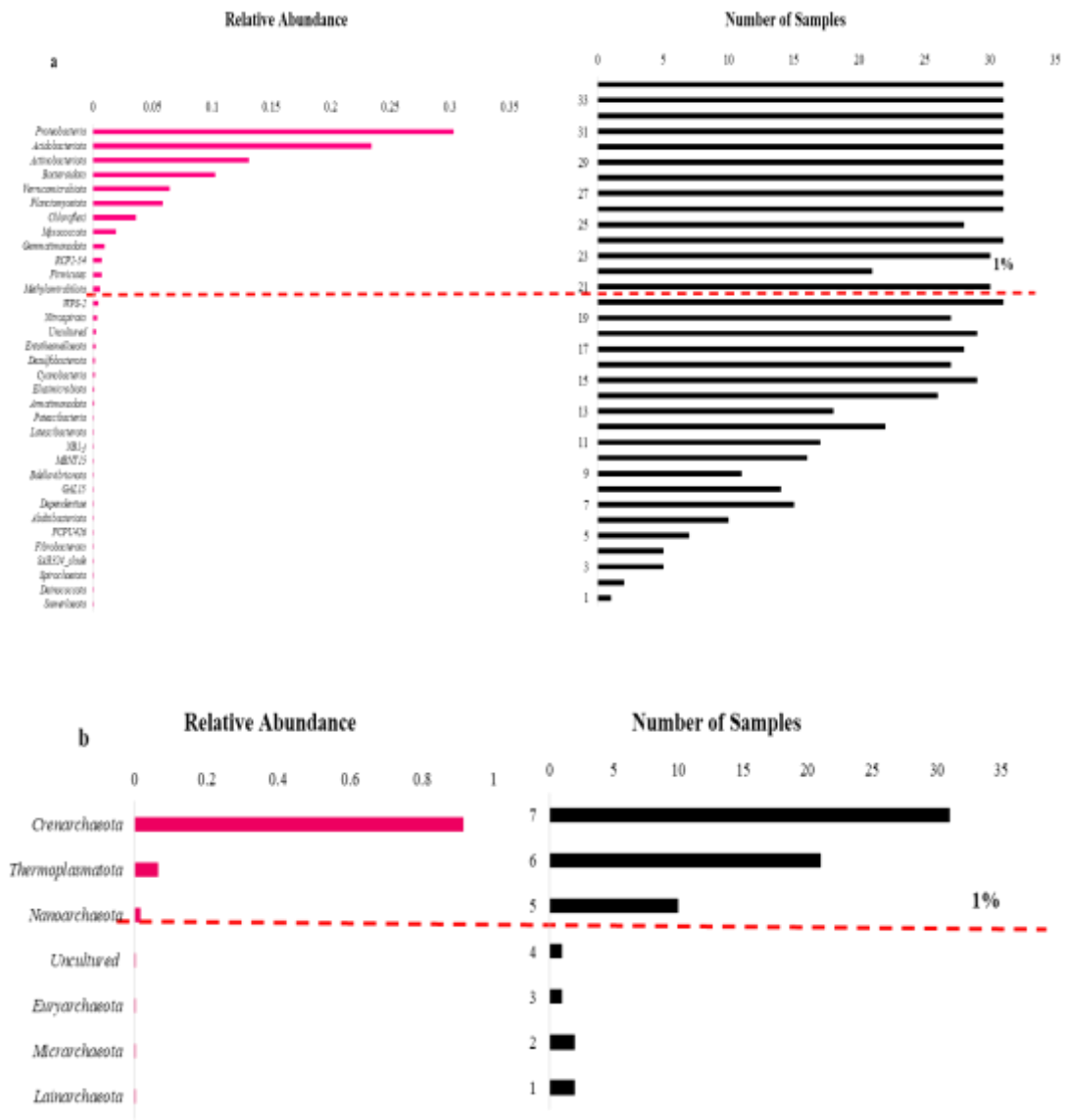


Figure 4.2 (a and b): Mean Relative Abundances of Prokaryotic Phyla across Forest Soil Samples, together with the Number of Samples in which they Were Identified (a- Bacteria) (b - Archaea). The phyla with mean relative abundance scoring >1% and above red dashed line are dominant.

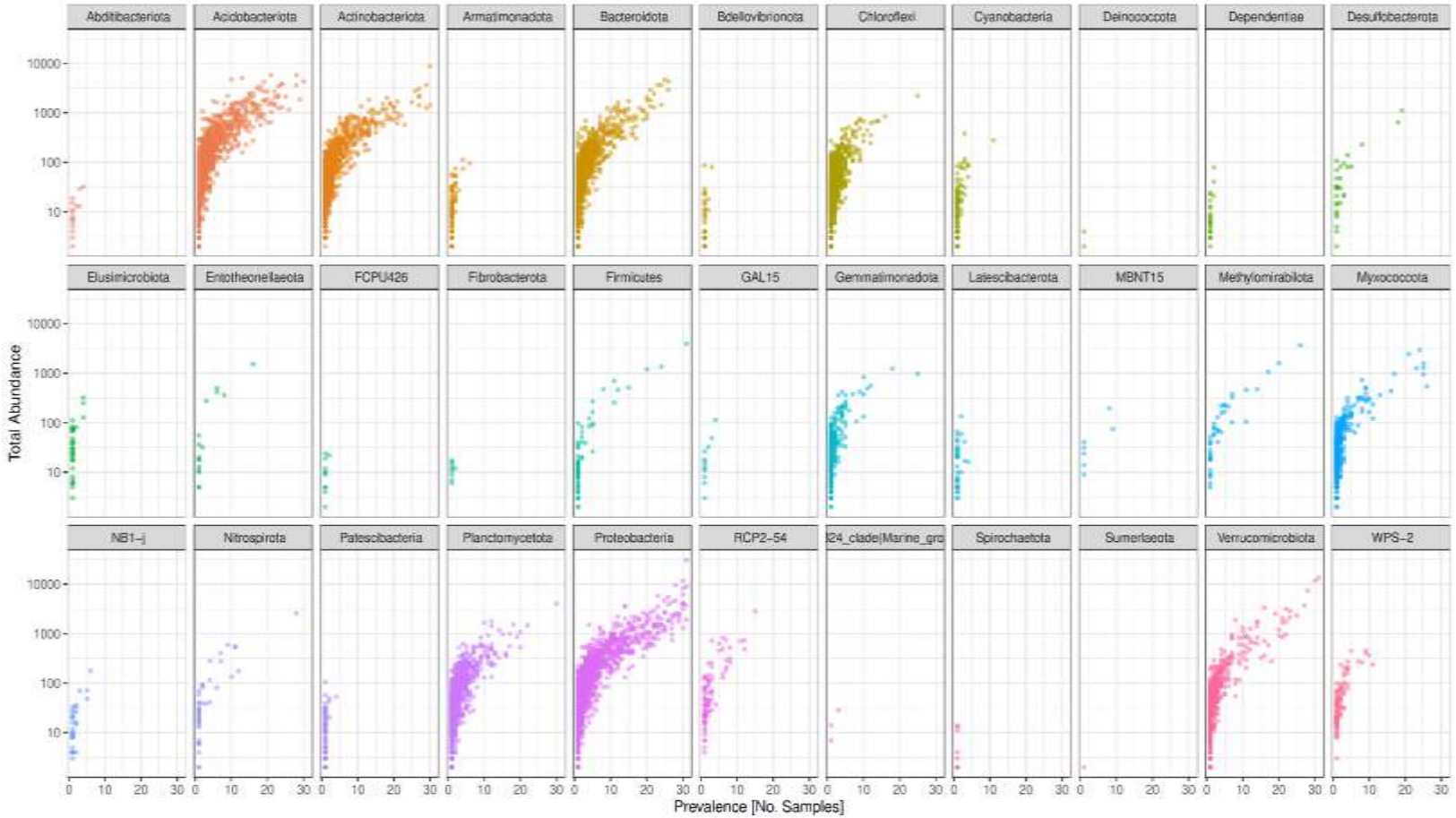


Figure 4.3: Feature Prevalence of Major Bacterial Phyla within Soil Samples Collected from Selected Forests Ecosystem in Kenya.

Table 4.3: Distribution of High-Quality Sequences, ASVs and Prokaryotic Taxa within the Selected Forest Ecosystems

Sample	Ecosystem	High quality sequences	No. of ASVs	Taxonomic classification							Top 3 Most abundant taxa at family level
				Phylum	Class	Order	Family	Genus	Species	Uncultured Genus	
K5	Chawia - Inside	74672	1314	30	54	120	175	280	500	78	*Acidobacteriae: Subgroup_2:Subgroup_2*, *Xanthobacteraceae* and *Nitrososphaeraceae*.
K6	Chawia-periphery	88435	1519	30	60	131	207	345	614	95	*Xanthobacteraceae*, *Acidobacteriae: Subgroup_2:Subgroup_2* and *Gammaproteobacteria_Incertae_Sedis: Unknown_Family.*
K7	Susu plantation	90394	1174	30	55	109	157	250	436	75	*Acidobacteriae: Subgroup_2:Subgroup_2*, *Xanthobacteraceae* and *Gammaproteobacteria_Incertae_Sedis:Unknown_Family.*
K8	Ngangao - lower	117657	1832	35	75	157	231	382	768	150	*Xanthobacteraceae*, *Nitrososphaeraceae* and *Chthoniobacteraceae.*
K9	Ngangao - Upper	29521	720	25	45	99	143	226	364	67	*Xanthobacteraceae*, *Gammaproteobacteria_Incertae_Sedis:Unknown_Family* and *Acidobacae: Subgroup_2: Subgroup _2.*
K10	Vuria	87256	1838	33	74	164	257	432	863	117	*Nitrososphaeraceae*, *Xanthobacteraceae* and *Chitinophagaceae.*
K15	Ngubi	77702	1566	32	67	140	216	359	704	105	*Chthoniobacteraceae*, *Xanthobacteraceae* and *Chitinophagaceae*.
K16	Longonot	86921	1648	26	59	139	211	345	678	109	*Nitrososphaeraceae*, *Chthoniobacteraceae* and *Chitinophagaceae*.
K18	Londiani junction	82204	1720	32	71	156	235	402	808	107	*Nitrososphaeraceae*, *Chthoniobacteraceae* and *Xanthobacteraceae*.
K21	Kakamega	63156	1213	31	66	139	209	326	622	86	*Xanthobacteraceae*, *Nitrososphaeraceae*, *Chitinophagaceae*
K23	Arbadare - Kihinganda	31646	639	23	46	90	113	164	265	49	*Acidobacteriae:Subgroup_2:Subgroup_2*, *Xanthobacteraceae* and *Gammaproteobacteria_Incertae_Sedis:Unknown_Family*
K24	Mt Elgon - Kapsokwony	11698	386	22	46	88	127	185	280	53	*Xanthobacteraceae*, *Vicinamibacterales: uncultured* and *Vicinamibacteraceae*.
K25	Kitale School	33782	780	24	55	114	175	274	487	54	*Flavobacteriaceae*, *Xanthobacteraceae* and *Chitinophagaceae*.
K26	Moi Barracks Eldoret	45988	1000	26	51	111	171	289	502	86	*Chthoniobacteraceae*, *Xanthobacteraceae* and *Chitinophagaceae*.
K27	Eldoret-Kapsabet junction	34857	919	24	53	114	176	296	540	78	*Xanthobacteraceae*, *Chitinophagaceae* and *Chthoniobacteraceae*.
K28	Equator Londiani	29793	839	24	50	107	171	297	498	80	*Xanthobacteraceae*, *Chitinophagaceae* and *Chthoniobacteraceae*.
K29	Muguga - Kiambu	37816	904	26	57	117	179	286	508	80	*Nitrososphaeraceae*, *Xanthobacteraceae* and *Chitinophagaceae*.
K33	Aberdare	64809	1370	28	63	141	211	353	698	102	*Xanthobacteraceae*, *Nitrososphaeraceae* and *Chitinophagaceae*

Sample	Ecosystem	High quality sequences	No. of ASVs	Taxonomic classification							Top 3 Most abundant taxa at family level
				Phylum	Class	Order	Family	Genus	Species	Uncultured Genus	
K34	Pesi - Ndaragwa Aberdare Tosha Forest	61451	1292	34	62	131	198	318	595	86	*Xanthobacteraceae*, *Acidobacteriae: Subgroup_2:Subgroup_2* and *Chitinophagaceae*.
K35	Mt. Kenya - Castle Kimunge	67789	1355	30	66	136	205	342	620	87	*Xanthobacteraceae*, *Acidobacteriae: Subgroup_2 and Acidobacteriales:uncultured*
K36	Mt. Kenya - Irangi	72684	1254	27	50	110	165	282	507	84	X*anthobacteraceae*, *Acidobacteriales: uncultured* and *Acidobacteriae: Subgroup_2:Subgroup_2*
K37	Mt. Kenya - Kyang'ondu	58228	1285	30	61	132	210	353	659	92	*Xanthobacteraceae*, *Chitinophagaceae* and *Vicinamibacterales: uncultured*.
K38	Mt. Kenya - Kinondoni	57676	1305	33	70	141	216	347	682	89	*Xanthobacteraceae*, *Chitinophagaceae* and *Chthoniobacteraceae*.
K39	Mt. Kenya - Kinondoni	68584	1493	35	68	150	223	361	696	92	*Chitinophagaceae*, *Xanthobacteraceae and Acidobacteriae: Subgroup_2:Subgroup_2*.
K40	Mt. Kenya - Kinondoni	63811	1238	34	75	138	193	304	531	81	*Acidobacteriae:Subgroup_2:Subgroup_2*, *Xanthobacteraceae* and *Acidobacteriales:uncultured*
K42	Mt.Kenya - Sirimon	66581	1408	28	56	127	206	368	702	97	*Xanthobacteraceae*, *Chitinophagaceae* and *Chthoniobacteraceae*
K63	Arbadare Githika block	63772	1147	31	61	124	184	295	523	80	*Acidobacteriae:Subgroup_2:Subgroup_2*, *Xanthobacteraceae and Acidobacteriales:uncultured*
K66	Mt.Kenya-state lodge	66758	1204	27	62	133	203	321	606	92	*Xanthobacteraceae*, *Nitrososphaeraceae* and *Vicinamibacterales:uncultured*
K70	Ngong hills nature reserve	71360	1321	26	61	131	200	347	436	95	*Xanthobacteraceae Chitinophagaceae Nitrososphaeraceae*
K71	Ngong hills Corner Baridi	63574	1183	27	55	131	202	331	654	94	*Chitinophagaceae Xanthobacteraceae* and *Nitrososphaeraceae*.
K77	Menengai crater	73741	1795	33	73	164	251	411	599	114	*WD2101_soil_group*, *Xanthobacteraceae* and *Chitinophagaceae*.

4.5 Archaeal Taxonomic Composition

Archaeal groups recovered from the obtained ASVs within the selected forests ecosystem were distributed among seven (7) phyla namely; *Crenarchaeota*, *Euryarchaeota*, *Lainarchaeota*, *Micrarchaeota*, *Nanoarchaeota, Thermoplasmatota,* and Unknown Archaeal phylum (Appendix 189). The LEfSe algorithm identified several differentially abundant archaeal taxa (P adj. = 0.001998) within the three regions (Aberdare, Nairobi and Taita Taveta) each having a taxon (Figure 4.1 b). *Crenarchaeota* was the most abundant Archaeal phylum represented across all samples, with 91.6% mean relative abundance (Figure 4.2b). This phylum was further grouped into two classes; *Nitrososphaeria* (77.1% mean relative abundance) and *Bathyarchaeia* (0.2% mean relative abundance). The second was *Thermoplasmatota* phylum and lastly *Nanoarchaeota* phylum. In a comparison of samples from the five forest regions (Aberdare, Mt. Kenya, Nairobi, Taita Taveta and Western), several taxa were identified as differentially abundant; 5 phylums (*Crenarchaeota*, *Thermoplasmatota* , *Nanoarchaeota, Iainarchaeota* and *Euryarchaeota*) in Taita Taveta, 5 (*Crenarchaeota*, *Thermoplasmatota*, *Nanoarchaeota, Iainarchaeota* and *Micrarchaeota*) in Aberdare, 4 (*Crenarchaeota, Micrarchaeota, Nanoarchaeota* and *Thermoplasmatota)* in Mt. Kenya and 2 (*Nanoarchaeota and Thermoplasmatota*) in Nairobi region (Appendix I). Hierarchical clustering of samples at phylum level revealed clustering of forest sites and prokaryotic taxa into two major groups where one parent comprised more diversified groups as shown in Figure 4.4.

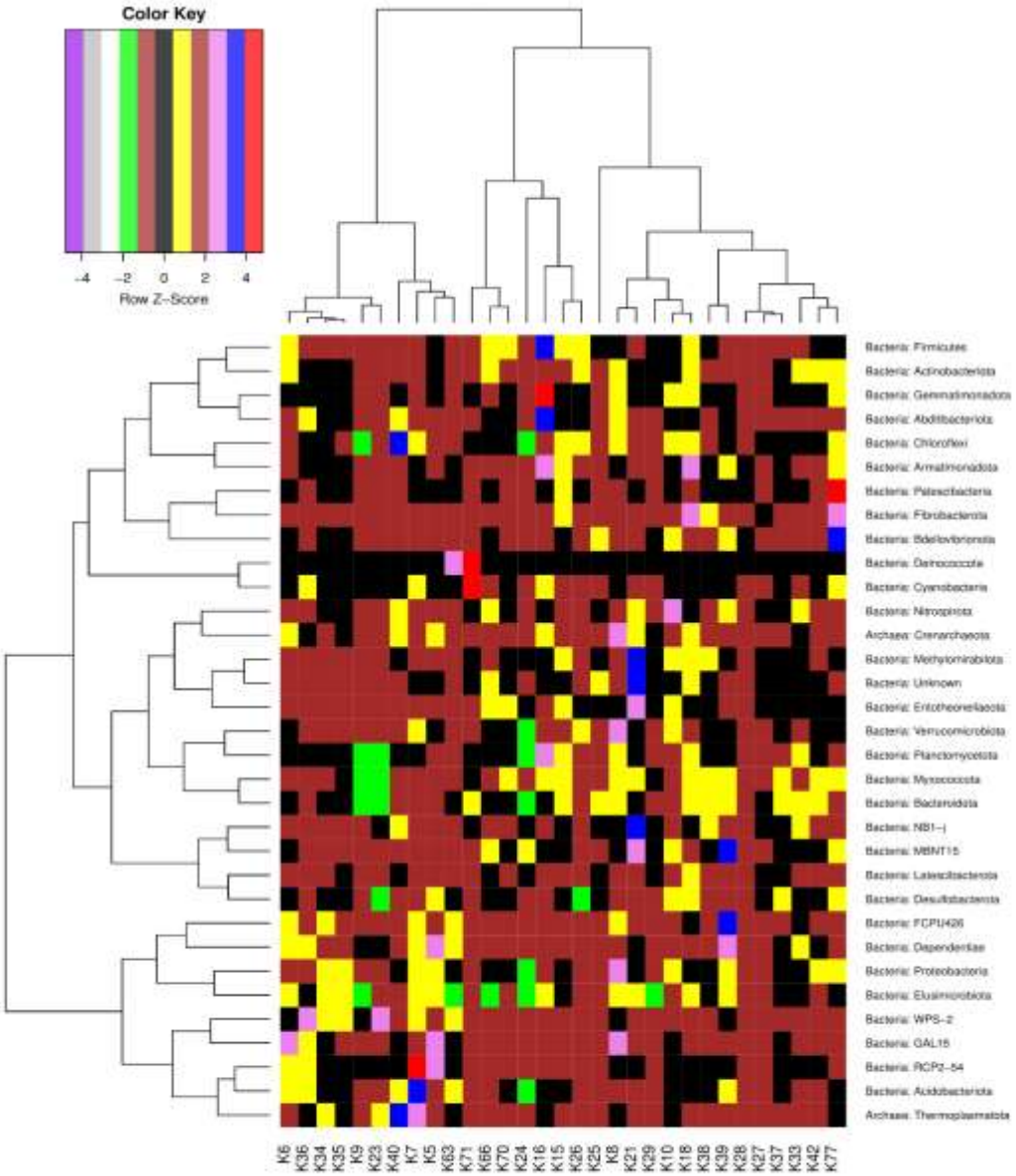


Figure 4.4: Hierarchical Clustering of the Most Predominant Prokaryotic Taxa at Phylum Level within the Selected Forests Ecosystem for Bacteria and Archaea. The codes represent the datasets of the sampling sites.

4.6 Alpha (α) Diversity of Soil Prokaryotic Communities

The prokaryotic taxa within each forest ecosystem were visualized using rarefaction curves. A steep slope that flattened to the right was observed in the rarefaction curves, indicating that a reasonable number of prokaryotic groups had been sequenced and more intensive sampling was likely to yield only a few additional species. The sampling curves tended to be asymptotic, denoting that prokaryotic communities were relatively deeply sampled (Figure 4.5).

Diversity indices (Chao1, Observed, Shannon and Simpson) within samples grouped as per the region showed that soils prokaryotes from Taita Taveta were highly diverse for Chao1, Observed and Shannon except Simpson (Figure4.6), which is as a result of one dominant genera (*Acidibacter)* that greatly outnumber other genera, leading to low diversity followed by western region that was highly variable on Shannon and Simpson but with low diversity on observed and chao1 because only one genera (RB41) that was present as the abundant which is as a result of low estimated genera richness within the community (Figure4.6). Nairobi and Mt. Kenya region had a moderate diversity alpha index whereas Aberdare Forest samples had the lowest diversity of prokaryotes and this could be as a result of having a few genera dominating the community, with little diversity overall (Figure 4.6). Analysis of sample alpha-diversity showed Western and Taita Taveta regions soils to have significantly different ($P = 0.0124603$) levels of Archaeal richness, while Western and Aberdare regions soils displayed significantly different Shannon diversity index ($P = 0.0199513$) (Figure 4.7 d, e and f). Although there were no significant differences between bacterial communities displayed within various forests ecosystems (Figure 4.7a, b, and c) soil samples under bamboo vegetation cover within Mt. Kenya and Aberdare regions displayed lower diversity than the other ecoregions.

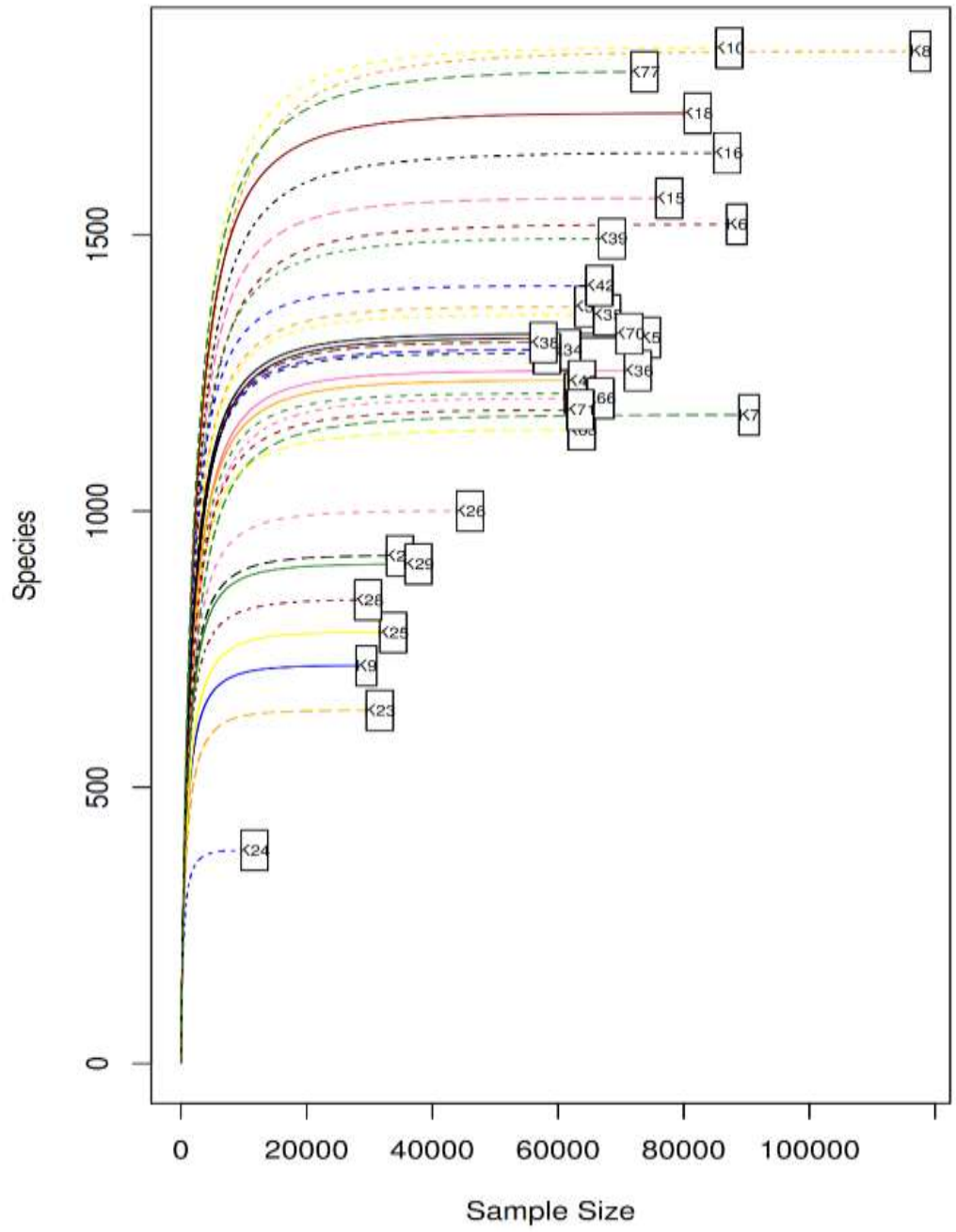


Figure 4.5: Rarefaction Curves Indicating Level of Sequence Coverage of Prokaryotic Communities within Soil Samples Collected from Selected Forests Ecosystems in Kenya.

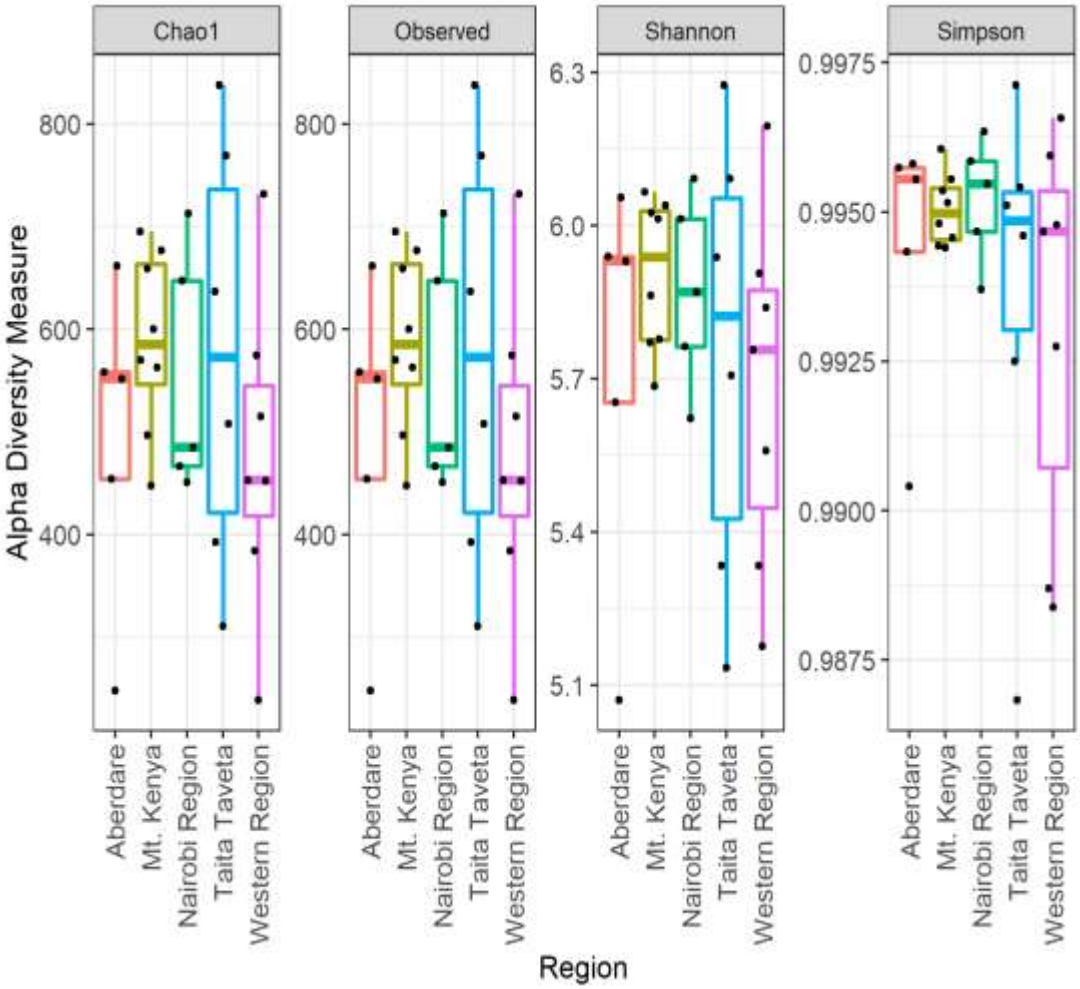


Figure 4.6: Alpha Diversity Indices (Chao1, Observed, Shannon and Simpson) of Prokaryotic Communities within Soil Samples Collected from Selected Forests Ecosystem in Kenya.

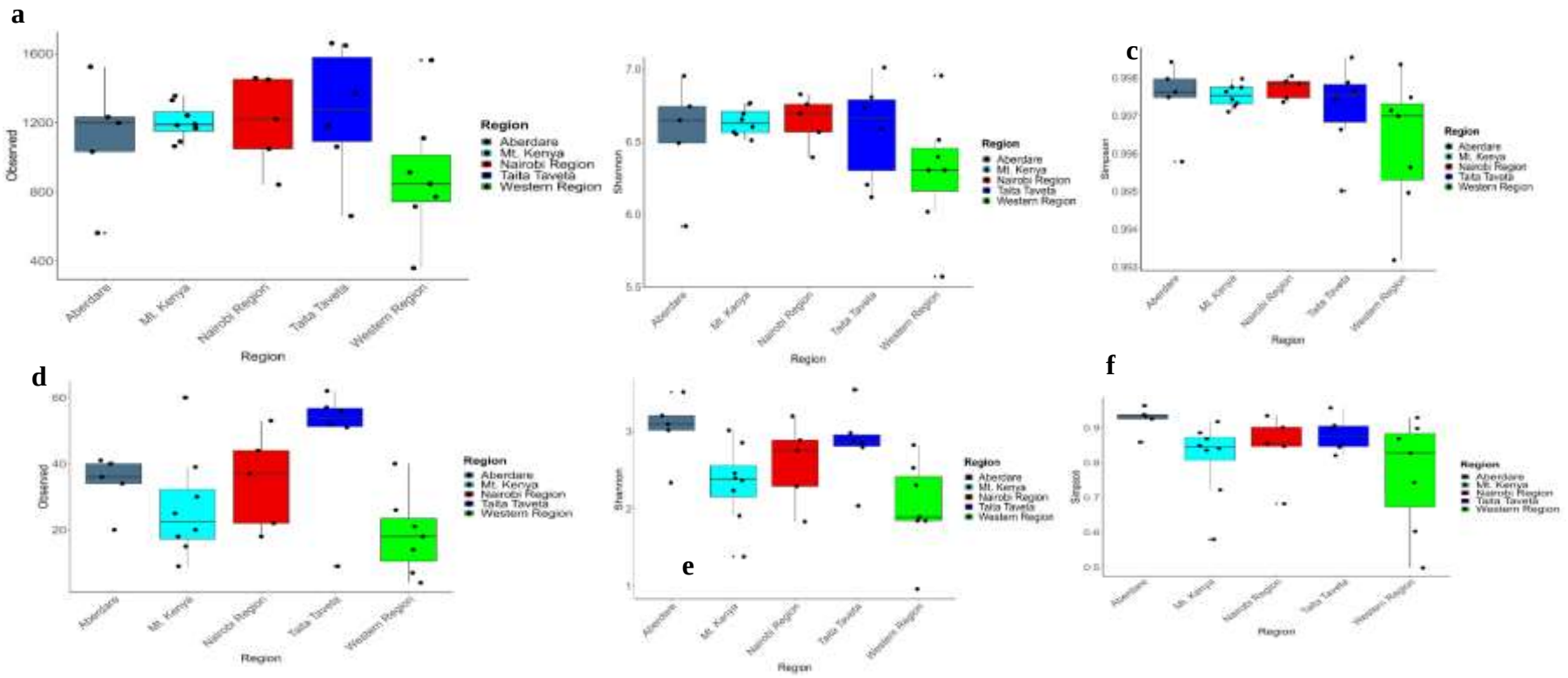


Figure 4.7: (a - f): Alpha Diversity of Soil Prokaryotic Communities. a, b and c Represent Diversity Indices (Observed, Shannon and Simpson) of Bacterial Communities while d, e and f Represent Diversity Indices (Observed, Shannon and Simpson) of Archaea Communities within Soil Samples Collected from the Five Regions in Kenya.

4.7 Beta Diversity of Prokaryotic Communities

Principal Coordinate Analysis (PCoA) using unweighted UniFrac analyses on ASVs and taxonomic composition showed highly significant differences (P= 0.006993) between samples collected from various forests ecosystem. The samples were grouped into five regions as shown in Figure 4.8c. There was an overlap of ellipses between the forest's ecosystem indicating that some sites harboured possibly related prokaryotic or because of habitat heterogeneity where variations in habitat features, such as microclimates, soil types, or topography, can create patches of suitable habitat for multiple species, resulting in overlapping ellipses. Beta-diversity analysis of soil samples from these regions showed a significant difference (P = 0.0010998) on bacterial and archaeal community structure (Bacteria R^2= 0.19; Archaea R^2 = 0.22) (Figure 4.8 a and b).

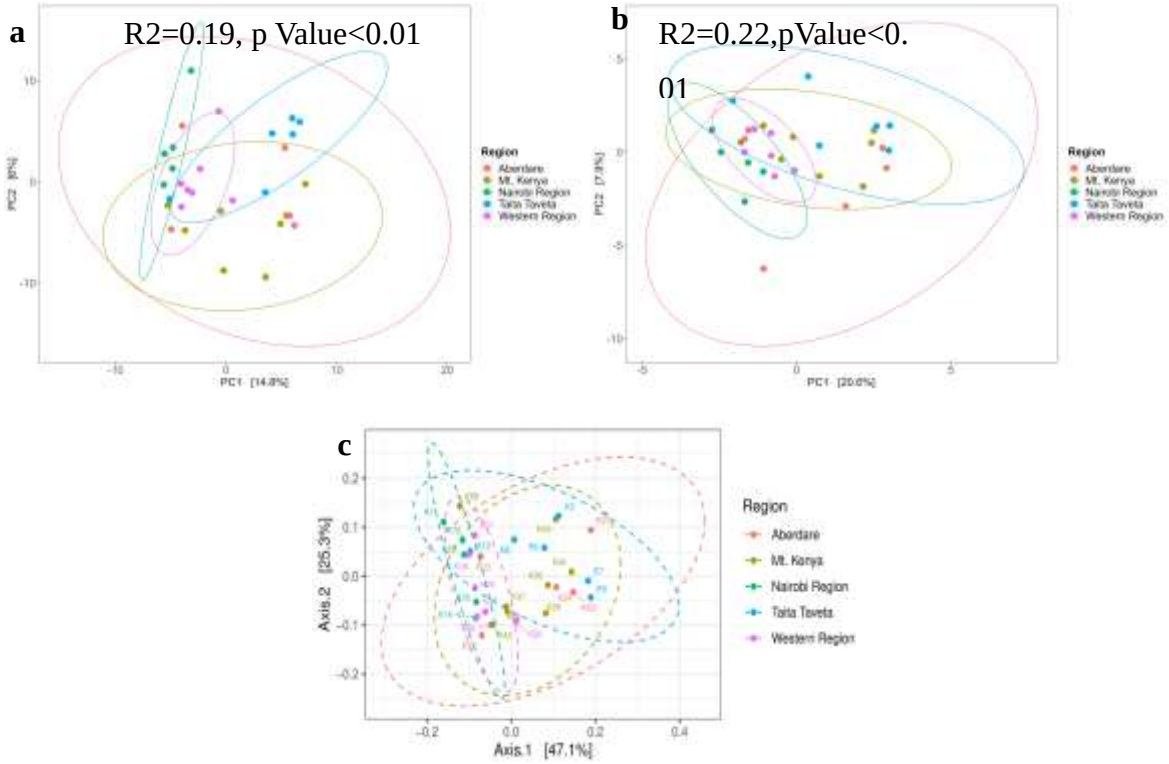


Figure 4.8(a, b and c): Principal Coordinate Analysis of Prokaryotic Diversity Based on Bray-Curtis Index within Soil Samples. **a** represents beta diversity of Bacterial community structure while **b** represents beta diversity of Archaea community structure within soil samples collected from the five regions in Kenya. **c**

Principal Coordinate Analysis of prokaryotic diversity based on weighted UniFrac analysis within soil samples collected from selected forests ecosystem in Kenya.

4.8 Environmental Drivers of Prokaryotic Communities

Principal component analysis (PCA) was performed based on soil physiochemical characteristics and prokaryotic taxa at order level to assess how environmental variables shaped soil prokaryotic community structure. Each of the soil physiochemical parameters were assessed on their ability to influence diversity positively or negatively within sampling sites. The parameters Na, Silt, Mg, Ca, K, pH and C were shown to have a positive diversity influence on both PC1 and PC2, while Al, P, Fe, Clay and Sand negatively influenced both PC1 and PC2. Manganese was displayed to have a positive influence on PC1 only (Figure 4.9).

A stepwise model building approach for constrained ordination models was used to assess the potential environmental drivers of the prokaryotic communities within forest ecosystems. Canonical correspondence analysis (CCA) ordination plots showed that bacterial and archaeal community structures were significantly affected by several soil physicochemical characteristics ($P = 0.001034$). Soil pH, Ca, K, Fe and %N were shown as key drivers of bacterial community structure, while Na, pH, Ca, P and %N were important factors in shaping archaeal community structure within forest soils (Figure 4.10 a and b). The significant effect of nitrogen to community structure is also consistent with the composition of soil microbiomes described in this study, which were dominated by taxa potentially involved in nitrogen fixation such as *Cyanobacteria* and *Nitrospirota* (Figure 4.1a).

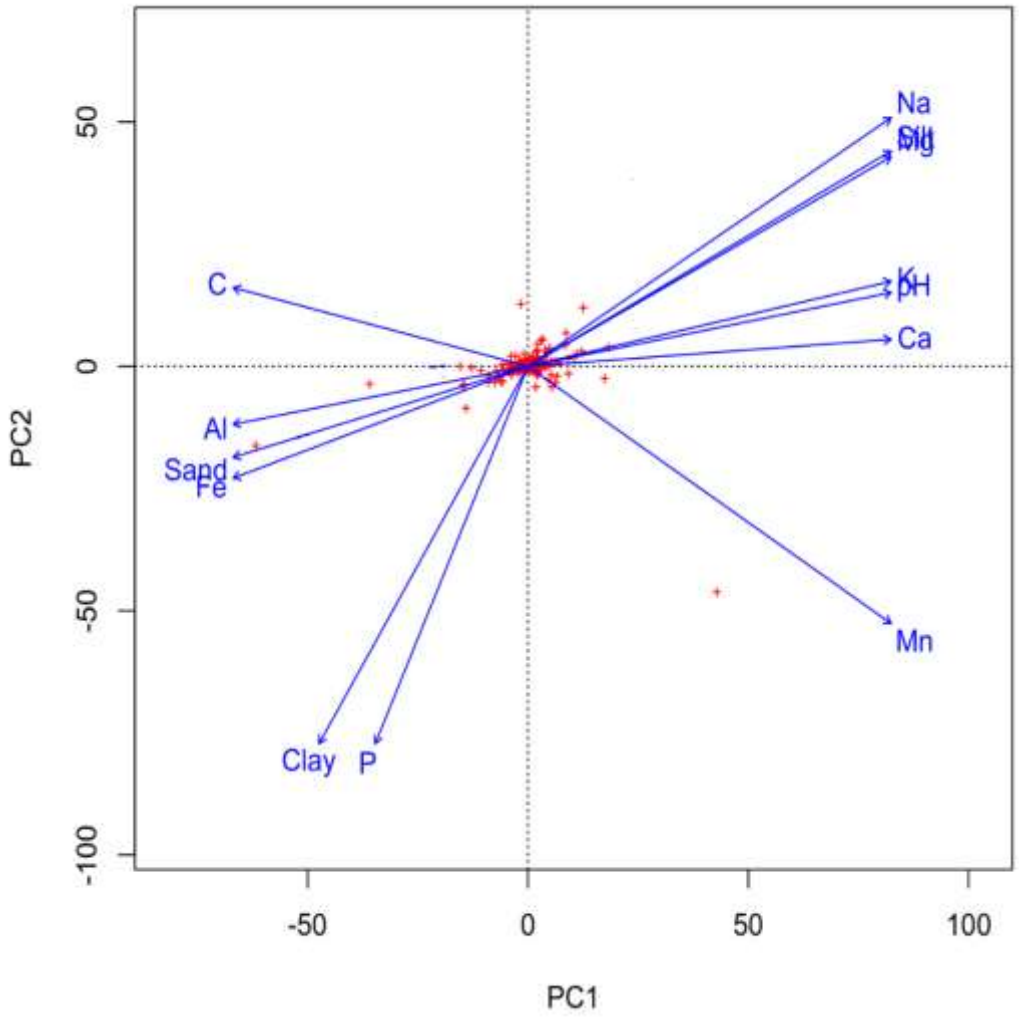


Figure 4.9: Principal Component Analysis of Soil Physiochemical Characteristics that Drive Prokaryotic Diversity within Selected Forest Ecosystems. The vectors points towards direction of increase in influence for a given soil characteristic on diversity at order level. The red crosses represent prokaryotic taxa distribution.

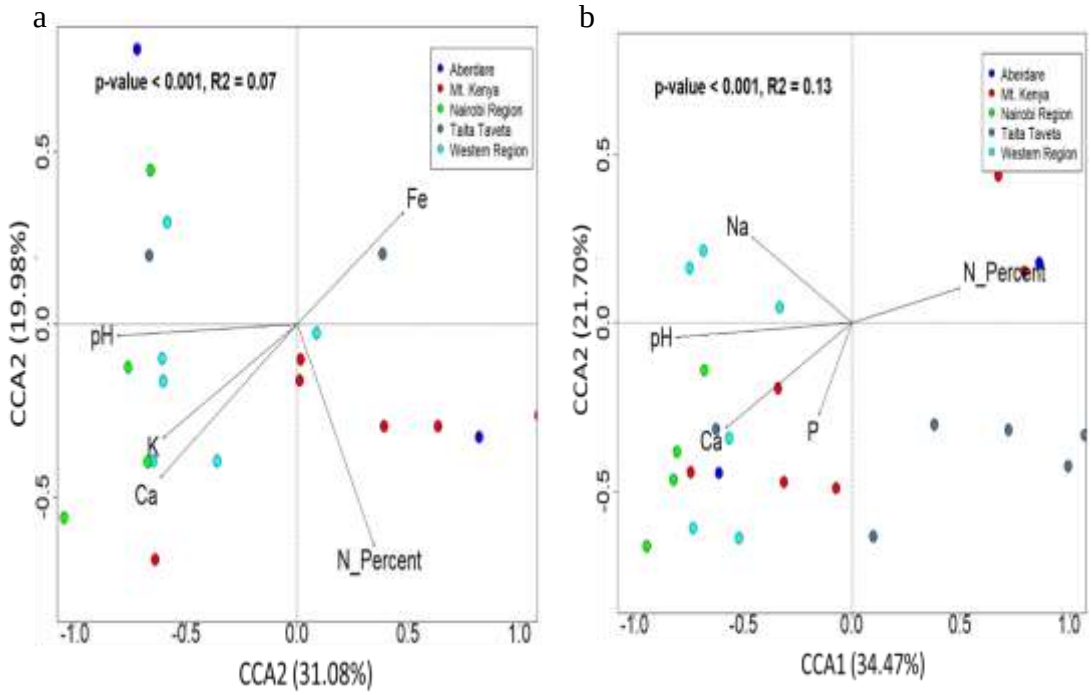


Figure 4. 10 (a and b): Canonical Correspondence Analysis (CCA) Plots Showing the Effect of Soil Physicochemical Characteristics and Plant Density Index on Bacterial and Archaeal Communities at 99% Significance. The percentage explained by various soil characteristics is expressed in the CCA1 and CCA2 axes and samples were color-coded on the plots according to forest regions. a represents CCA of Bacterial communities while b represents CCA of Archaea communities within soil samples collected from the five regions in Kenya.

4.9 Different Forest Soils in Kenya Have Unique Physicochemical Properties

In this study, 31 soil samples were obtained from forests ecosystems within the Taita Taveta, Nairobi, Western, Aberdare and Mt. Kenya ecoregions. Samples from the different ecoregions were shown to be significantly different (p-value = 0.001998, R^2=0.45) in terms of soil physicochemical properties, specifically in soil pH, soil texture, macro- and micro-nutrient composition and Enhanced Vegetation Index-2 (EVI2) (Figure 4.11 a and b, Appendix 2a and b). Taita Taveta forest soils were highly distinct from those of the Nairobi, Aberdare and Western regions (Figure 4.11 b). Conversely Nairobi and Western region soils exhibited the least variability

(Figure 4.11 b). Several soil physicochemical properties were found to be correlated, and thus could be considered as interdependent (Figure 4.11 c).

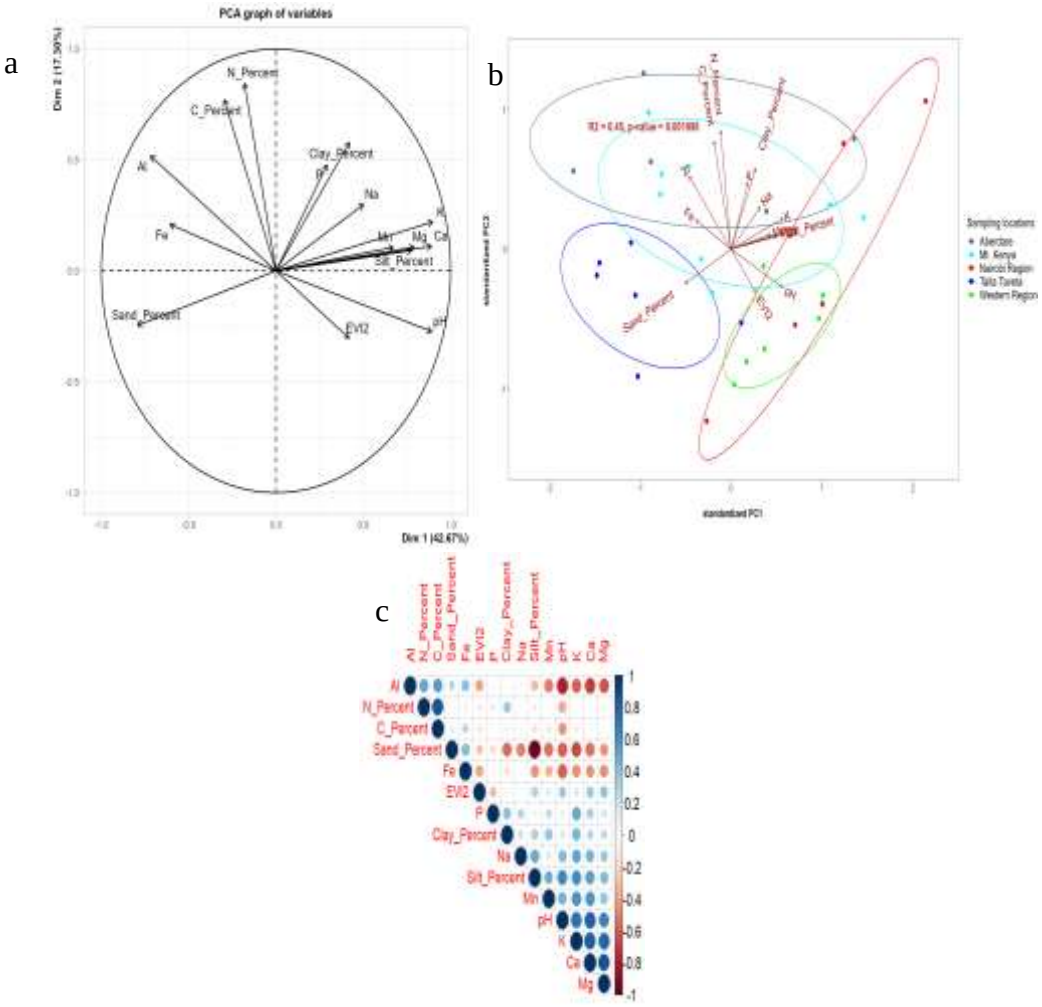


Figure 4.11(a-c): (a) Principal Component Analysis (PCA) Biplot of Forest Soil Samples According to their Physicochemical Properties. The influence of each variable on sample distribution is represented by the arrows radiating from the center of the PCA plot. (b): The sample clusters corresponding to different ecoregions are highlighted within ellipses of the same color. (c) Pearson correlation plot between measured soil physicochemical properties. Positive and negative correlations are displayed in blue and red shades, respectively, while the size and intensity of matrix circles is proportional to correlation coefficient between variables.

4.10 Microbial Interactions

The relationship between various soil microbiomes was presented in an interaction network based on the correlation of the top twenty most abundant prokaryotic taxa (Figure 4.12). Eight classes were presented as the most associating prokaryotic taxa within the co-expression network. The thick edges (with a distance of 0.0) represent a close interaction between the prokaryotic classes. The nodes in green and blue colors represent highly ubiquitous and co-occurring groups of *Alphaproteobacteria* and *Gammaproteobacteria* taxa at class level. The edges with boldness ranging from 0.0 to 0.6 represent association co-dependence of different classes, where 0.0 and 0.6 show high and insignificance co-dependence, respectively. Among bacterial communities, *Candidatus udaeobacter*, an uncultured genus of the class V*errumicrobiae* and *Pseudolabrys* genus from the family of *Nitrobacteraceae*, class *Alphaproteobacteria* were shown to have the most interactions with other prokaryotic genera (Figure 4.10). It was noted that *uncultured* prokaryotic taxa had high interaction levels with other organisms within the association network compared to the few known co-occurring genera; *Bradyrhizobium*, *Pseudolabrys*, *Acidibacter*, *Puia* and *RB41* involved in major networks in the forest's environment (Figure 4.12).

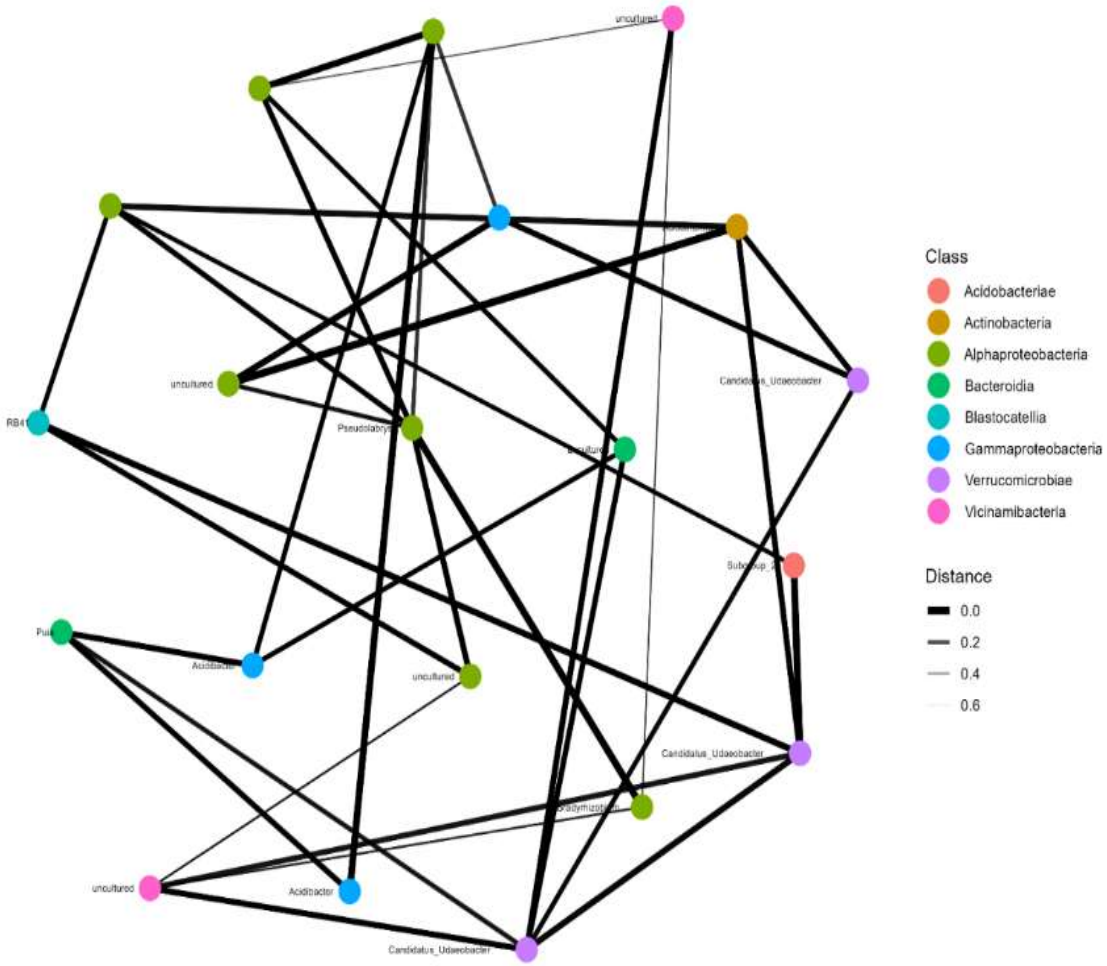


Figure 4.12: Microbial Interaction Network within Soil Samples Collected from Selected Forests Ecosystem. Each Node is represented by an ASVs Indicating Individual Species.

4.11 The "Uniqueness" of Kenyan Forest Microbiomes

In order to address the question on whether Kenyan forest soils harbor unique microbiome compositions, the phylogenetic datasets used in this study (Table 4.3) were compared with datasets on forest soil microbiomes from other countries across the globe (Table 3.2)**.** Comparisons of the beta-diversity scores between these datasets, based on Bray-Curtis index (Figure 4.13), revealed community structures of forest soil microbiomes, to some extent, being distinguished by the country of origin (R^2 = 0.63; p-value = 0.0098). The significant compositional differences between

national datasets were reflected in the LDA comparison results, which identified 177 taxa differently distributed across the datasets (Appendix 3).

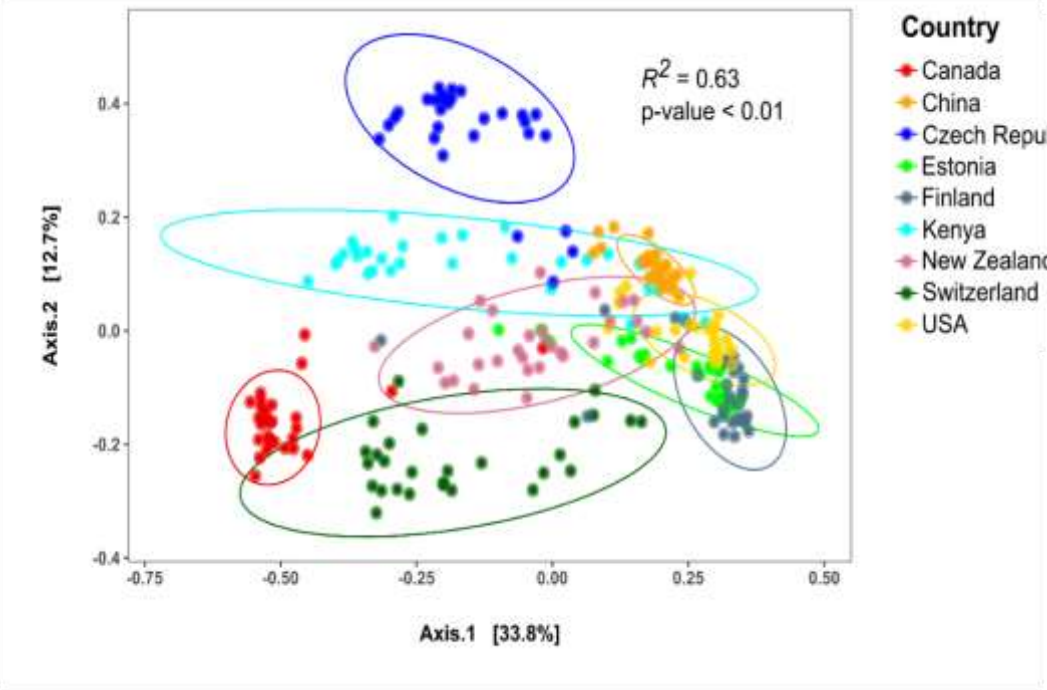


Figure 4.13: Principal Coordinate Analysis (Pcoa) Ordination of the Bray-Curtis Distance between Samples from Different Country Datasets.

Samples are colored according to country origin. The PERMANOVA significance results on differences in beta-diversity according to country of origin are displayed in the plot. Ellipses around the clusters of samples represent the predicted distribution of points within each country group at a 95% confidence interval.

CHAPTER FIVE

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

5.1 Discussion

Studies on forest soil microbiomes have enabled researchers to recognize and appreciate the vast extent of microbial diversity, thanks to the recent advancements in metagenomic studies. These studies reveal components within the interactive assemblage of microorganisms that are often undetectable using standard culturing techniques (Bull, 2004; Handelsman, 2004).

In this study, Sample K66 collected from Aberdare Forest colonized by *Podocarpus* trees as vegetation cover (Table 3.1) clustered separately from other soil samples. This could be attributed to the fact that *Podocarpus* species are extremely allergenic, and have an Ogren Plant Allergy Scale rating of 10 out of 10 (Ogren, 2015). *Podocarpus* leaves, stems, bark and pollen are cytotoxic and exposure to the pollen could produce symptoms that mimic the cytotoxic side effects of chemotherapy. Some species of *Podocarpus* have been used in systems of traditional medicine for conditions such as fevers, coughs, arthritis, sexually transmitted diseases, and canine distemper (Abdillahi *et al.*, 2011; Ogren, 2015).

The Mt. Kenya region has the highest number of abundant ASVs at 10,542, followed by Taita Taveta with 8,397, Western with 6,857, Nairobi with 6,622, and Aberdare with the lowest number at 6,243. The abundant ASVs at Mt. Kenya and Taita Taveta regions are integral to soil dynamics, providing essential ecosystem services that support plant growth and soil health. However, while most contributions are beneficial, careful management and monitoring are required to mitigate any negative impacts from pathogenic or harmful microbial members like bamboo trees that has an inherent natural barricade against bacteria that can lead to negative influence.

Sample K23 obtained from soils under bamboo vegetation cover within Aberdare Forest, displayed a low number of ASVs (Table 4.3). This could be attributed to the fact that bamboo trees has an inherent natural barricade against bacteria and

therefore, most varieties of bacteria and bugs that attempt to thrive on bamboo plant are eradicated naturally on contact (Afrin *et al.*, 2012).These results, suggests that bamboo grows so rapidly in nature due to its antimicrobial properties resulting from steroids, saponins, tannins, polyphones and alkaloids which might be responsible for broad antibiotic activity (Tao *et al.*, 2018). Sample K10 followed by K8 obtained from a montane vegetation, with thick ground cover, had high number of ASVs (Table 4.3) that could be attributed to forest ecosystem with a broad range of bacterial micro habitats such as foliage, wood of living trees, bark surface, ground vegetation, roots and rhizosphere, litter, soil, deadwood, rock surfaces, invertebrates, wetland or the atmosphere of which, each has its own specific features such as nutrients availability or temporal dynamic and specific drivers that affect bacterial abundance positively (Lladó *et al.*, 2017). Forest ecosystems with indigenous vegetation cover have been reported to improve soil quality thus increasing soil water retention and abundance of certain microbial species (Backes, 2001). These represent innate microbial consortium that inhabits soil and the surface of all living things that has potential to biodegrade, bio compost, bioleaching, nitrogen fix, improve soil fertility and production of plant growth hormone (Kumar & Gopal, 2015).

5.1.1 Abundance of Prokaryotic Taxa Within Selected Forests Ecosystem

The bacterial groups were dominant taxa within forests ecosystem. *Proteobacteria* was the most abundant phylum across all samples, with a high relative abundance of up to 44.98% whereas *Crenarchaeota,* an archaeal phylum was represented across all forest soil samples. Other key phyla included *Acidobacteriota, Actinobacteriota, Bacteroidota*, *Planctomycetota, Verrucomicrobiota* and *Chloroflexi*. Dominant orders across all samples included *Rhizobiales, Chitinophagales, Burkholderiales* and *Vicinamibacterales* (Figure 4.1).

Members of *Acidobacteriota* phylum have been reported as ubiquitous and dominant in soil ecosystems, mainly thriving on acidic soils (Lauber *et al.*, 2009).This is evidenced in sample K7 collected from Susu plantation with Eucalyptus plantation and wattle trees in highly acidic environment. These have been speculated to carry

out dynamic roles in ecological processes such as regulation of biogeochemical cycles, decomposition of biopolymers, exopolysaccharide secretion, and plant growth promotion (Kalam *et al.*, 2020). In this study, *Acidobacteriota* members included the family *Acidobacteriae Vicinamibacterales* across all samples, as characterized by highly acidic pH of 3.75, 3.42, 3.92 and 3.90 in samples K7, K23, K26 and K40 samples respectively (Table 4.1). *Bacteroidota* orders such as *Chitinophagales* found within the forest soils are known to degrade complex organic matter, such as chitin and cellulose (Rosenberg, 2014) while others have shown β-glucosidase activity (Bailey *et al.*, 2013).

Alphaproteobacteria, Gammaproteobacteria and *Betaproteobacteria* whose orders included *Rhizobiales, Xanthomonadales* and *Rhodospirillales* are well known for nitrogen fixation, mineralization and denitrification (Werner & Newton, 2005). *Actinobacteriota* members such as *Frankiales* and *Thermoleophilia* revealed in this study are known to produce biological active secondary metabolites of industrial importance like antibiotics, antifungal agents, antioxidants and insecticides (Chen *et al.*, 2019).

These results were similar to a previous findings where by five bacterial phyla; *Acidobacteria, Actinobacteria, Proteobacteria, Bacteroidetes* and *Firmicutes* were the most abundant taxa in most forest soils (Lauber *et al.*, 2009). Archaeal community was dominated by *Thaumarchaeota*, the genus *Nitrososphaera*, whereas some sequences belonged to *Euryarchaeota* and *Crenarchaeota* phyla (Siles & Margesin, 2016). Members of these phyla have been reported as major contributors to soil biogeochemical processes within various ecosystems (Karanja *et al.*, 2020). Different microbial phylotypes perform varied functions within the forest ecosystem that lead to a shift in diversity and abundance due to available nutrient sources and their effect on the soil health. The pattern of major biomes, soil type, geology, topography and vegetation are determinants of soil evolution within the forests. Macroclimate variations within the Taita Taveta and Western region imply that each forest harbors some degree of unique soil microbial genetic resource.

Although most of the previous research on forest soil ecology has focused on fungi, especially saprotrophic and mycorrhizal as the most abundant and diverse communities within forest ecosystems (Baldrian, 2017; Uroz *et al.*, 2016), there has been an increased appreciation of the role of bacteria as the other major significant, though less explored vital part of microbial community in forest soils (Baldrian Petr 2017). Previous research findings designate bacteria to commonly harbor genes encoding plant cell wall-degrading enzymes and contribute significantly to decomposition of organic matter (Berlemont & Martiny, 2015; López-Mondéjar *et al.*, 2015). The bacterial phyla unveiled in the current study are similar to those reported in, numerous bacterial strains previously isolated from forest soils and described as capable of degrading cellulose, belonged to *Proteobacteria*, *Actinobacteria*, *Bacteroidetes* and *Firmicutes* phyla (Eichorst & Kuske, 2012; Štursová *et al.*, 2012). In other studies, members of *Bacteroidetes* and *Proteobacteria* phyla isolated from forest soils had Ton B-dependent receptors previously identified as plant carbohydrate scavengers within these bacterial organisms (Eichlerová *et al.*, 2015).

The soil samples under investigation in this study were collected from 0-20 cm depth, a soil profile characteristically comprising the organic horizon that results from decomposition of litter-derived organic matter, thus demonstrating a mixture of processed, plant-derived organic matter and soil components (López-Mondéjar *et al.*, 2015). Previous studies indicate that bacterial communities are horizon specific when studied along a sharp vertical stratification acquired from decomposition of litter-derived organic matter and weathering of the mineral matrix, although a high level of taxon overlap has been observed (López-Mondéjar *et al.*, 2015). For instance, 40 to 50% of all sequences obtained from the organic and mineral horizons in Quercus Forest were dominated by *Acidobacteria*; together with members of *Actinobacteria*, *Proteobacteria* and *Bacteroidetes* (López-Mondéjar *et al.*, 2015). In other separate studies, *Acidobacteria*, *Proteobacteria*, and *Actinobacteria* were the most abundant in organic horizon which have been proposed to preferentially utilize easily accessible carbon substrates (Fierer & Jackson, 2006).

5.1.2 Taxonomic Composition of Soil Microbiomes Across Kenyan Forest Biomes

Members of *Actinobacteriota phylum* such as *Frankiales* and *Streptomycetales* are known as nitrogen-fixing bacteria and may produce biologically active secondary metabolites (Chen *et al.*, 2019). The dominant bacterial phyla from the current study were consistent with other studies within two forests sites, where bacterial ASVs were assigned to 44 phyla, ten of which were; (*Proteobacteria*, *Acidobacteria*, *Verrucomicrobia*, *Firmicutes*, *Actinobacteria*, *Bacteroidetes*, *Planctomycetes*, *Chloroflexi*, WD272, and *Gemmatimonadetes*) comprised more than 90% of the relative abundance in each library (Wei *et al.*, 2018). Our results on bacterial abundance were also consistent with several previous studies where *Proteobacteria*, *Acidobacteria*, *Verrucomicrobia*, *Firmicutes*, *Actinobacteria*, *Bacteroidetes*, *Planctomycetes*, *Chloroflexi* were the most abundant phyla (Bastian *et al.*, 2009; Janssen, 2006; Miyashita, 2015).

In particular, members of *Proteobacteria* and *Acidobacteriota* phylum are ubiquitous and dominant in` soil ecosystems (Lauber *et al.*, 2009). Members of these Phyla, such as *Anaeromyxobacter*, *Bradyrhizobium*, *Azospirillum*, *Ralstonia*, *Burkholderia*, *Brevundimonas* *Rhodopseudomonas* (Proteobacteria), *Mycobacterium*, *Nocardia*, A*mycolatopsis* *Thermobispora*, *Pseudonocardia*, *Brachybacterium*, *Frankia*, *Conexibacter* (Actinobacteria), *Streptococcus*, *Lactococcus*, and *Enterococcus* (Firmicutes) carry out various key ecological processes such as regulation of biogeochemical cycles, decomposition of biopolymers, exopolysaccharide secretion and plant growth promotion (Kalam *et al.*, 2020) .

The order *Chitinophagales* that were represented across all samples contains members that are known to degrade complex organic matter, such as chitin and cellulose (Rosenberg, 2014). The orders *Rhizobiales*, *Xanthomonadales and Rhodospirillales* found in this study are also well known for nitrogen fixation, mineralization and denitrification activities (Delmotte *et al.*, 2009; Ivanova *et al.*, 2000; Verginer *et al.*, 2010; Werner & Newton, 2005).

Nitrososphaeria is a class from the Archaeal phylum, which includes many ammonia-oxidizing taxa that have been identified previously in forest soil microbiomes (Stieglmeier *et al.*, 2014; Tourna *et al.*, 2011; Wang *et al.*, 2014; Wu *et al.*, 2017). These results are in agreement with previous studies where archaeal communities in forest biomes were dominated by *Nitrososphaera*, (Saghaï *et al.*, 2022; Siles & Margesin, 2016). Members of *Nitrososphaera* have been described as major contributors to soil biogeochemical processes such as carbon, methane, nitrogen and, sulfur cycle within many ecosystems (Dubey *et al.*, 2015; Karanja *et al.*, 2020; Offre *et al.*, 2013).

5.1.3 Alpha- and Beta- Diversity Analysis of Soil Prokaryotic Communities

Soils collected from the Taita Taveta region (Vuria and Ngangao) had the highest number of observed prokaryotic taxa (Table 4.3). These forests are characterized by a montane climate vegetation with thick ground cover (Pellikka *et al.*, 2018). The high number of ASVs could be attributed to a broad range of bacterial micro-habitats associated with high nutrient availability besides other specific microbial diversity drivers such as plant density and vegetation index that positively influenced bacterial abundance (Lladó *et al.*, 2017).

We identified several differentially abundant archaeal taxa within the three regions (Aberdare, Nairobi and Taita Taveta) using LEfSe algorithm each having a taxon. The genus *Acidibacter* was over-represented in Taita forest soils, possibly due to the low soil pH (Table 4.1) observed in this region. *IMCC26256* was over-represented in western region. *Burkholderia-Caballeronia-Paraburkholderia taxa,* which typically have a very broad ecological diversity and metabolic versatility (Hussain *et al.*, 2019; Uroz *et al.*, 2016) were the most abundant in Aberdare Forest soils, *RB41* in Mt. Kenya whereas *Rhodovastum* was the most abundant *in* Mt. Kenya region soil samples.

However, there was high prokaryotic variability observed within each region, an indication of distinct microhabitats and microclimates in each forest region covered in beta-diversity. Notably, the microbial composition of samples from the Taita Taveta region showed a lower degree of overlap with other regions (Figure 4.11b),

which mimics the soil chemistry differences observed between the regions. Taita Hills comprise the northernmost part of the Precambrian Eastern-Arc Mountain range, known for its rich biodiversity (Platts *et al.*, 2011) and recognized as one of the world's 25 biodiversity hot-spots (Myers *et al.*, 2000).

The highly significant ($P < 0.01$) richness and Shannon diversity index values for samples from Western region forests could be attributed to the tropical nature of forests within this region such as sample K21 (Kakamega forest) which is considered an important biodiversity reservoir and the only remaining Guinea-Congolian tropical rain forest in Kenya (Lung & Schaab, 2004). Kakamega forest is the largest moist lowland forest ecosystem in Kenya, and has similar characteristics to Central Africa equatorial forests (Pellikka *et al.*, 2018). In other studies, shifts in bacterial richness and diversity as well as community structure that comprised of *Proteobacteria*, *Acidobacteria*, *Actinobacteria*, and *Bacteroidetes* significantly correlated with several environmental and soil chemical factors, especially soil (Siles & Margesin, 2016).

5.1.4 Factors in the Environment that Shape Soil Microbial Communities in Kenyan Forest Ecosystems

Eight soil physicochemical parameters were identified as the most significant drivers of microbial community structure across the forest's ecosystem gradient. Silt was shown as one of the key drivers of prokaryotic diversity within forest soils (Figure 4.11a)**.** This could be attributed to the small nature of particles where filamentous microbial groups have access to adequate surfaces to colonize and stabilize thus adjusting the soil microclimate to support growth (Weber *et al.*, 2016). In addition to Na, the other important driver of the bacterial community composition within the forest soils were pH, K, C, Ca, Silt and Mg which were positively modulated the composition of prokaryotic communities (both PC1 and PC2), whereas Sand, P, Clay, Fe and Al exhibited a negative influence on microbial diversity. Fe concentration and soil texture are the major factors in shaping bacterial community structures in some soils (Maundu & Tengnäs, 2005). Soil pH possibly affected the thermodynamics and kinetics of microbial respiration, thus shaping the microbial

communities' composition and function. In a study to determine impacts of soil texture and associated properties to the diversity and composition of soil microbial communities, genetic potential for degradation of organic compounds was higher in finer textured soils than the coarse-textured soils (Xia *et al.*, 2020). The significance of soil texture, texture-associated pores and nutrient resource locality in regulating microbial diversity and community composition were highlighted by identifying sand, silt or clay preferred microbial taxa and characterizing mineral particle-dependent genetic potential of organic carbon degradation and nitrogen cycling (Curd *et al.*, 2018; Lilly & Lin, 2004; Plante *et al.*, 2006; Vos *et al.*, 2013). Soil texture was revealed as the second most important factor after soil pH in shaping the soil microbial community, with relative abundances of some filamentous bacteria; *Actinobacteria* and *Chloroflexi* significantly increasing with silt and/or clay content (Hemkemeyer *et al.*, 2018; Karimi *et al.*, 2017; Sessitsch *et al.*, 2001; Xia *et al.*, 2020) .

Manganese had a positive influence on PC1 only. Results of a previous study demonstrated that the dominant tree species has a unique variable effect on soil chemical and microbial biomass magnitudes where larger microbial biomass concentrations revealed under *Podocarpus* experimental plots were attributed to bacterial taxa (Yohannes *et al.*, 2020). However, several soil physicochemical properties were correlated, and thus could be considered as interdependent as per Pearson correlation plot of measured soil physicochemical properties (Figure 4.8c)**.** For instance, the vegetation index (EVI2) was positively correlated with all the measured soil nutrients, apart from phosphorus. This was similar to the previous study where nutrient-rich forest soils are associated with high density plant growth (Rocha & Shaver, 2009).

The soil samples used in this study were collected from 0-5 cm depth, which is within the 0-20 cm soil profile characteristically comprising the organic horizon that results from decomposition of litter-derived organic matter and representing a nutrient-rich mixture of processed, plant-derived organic matter (López-Mondéjar *et al.*, 2015). Low titratable phosphorus concentrations were possibly due to presence of a high content of Al and Fe, which form oxides that fix phosphorus at the low

pH's associated with these soils (Kanyanjua *et al.*, 2002; Neina, 2019). In this study, the pH was positively correlated with EVI2+ but negatively correlated with C and N content (Figure 4.8c). This result contradicts a previous study which concluded that at higher soil pH levels, the mineralizable fractions of C and N increased, possibly due to the direct effect of pH on microbial populations and their activities (Neina, 2019).

5.1.5 The Distinctiveness of Microbial Communities in Kenyan Forests

Kenyan dataset formed a distinct group with some degree of overlap with soil microbiomes from China, Czech Republic, New Zealand and the USA (Figure 4.10). This overlap could be a result of common plant cover between the sampled areas in the different countries. Some forests in Kenya are known to harbor globally distributed plant species such as bamboo (*A. alpina*), indigenous plant species found within forests with highest floral diversity such as (*Coffea fadenii*, *Juniperus procera* – African pencil cedar, *Podocarpus falcatus, latifolius*, *Tabernaemontana stapfiana, Ocotea usambarensis*, *Macaranga conglomerata*, and *Psychotria petit*. Forests harboring moderate floral diversity included *Podocarpus*, *Dombeya*, *Croton megalocarpus*, whereas dryland species included Acacia species such as *A. tortilis*, *A. melifera*, *A. abyssinica, and A. polyacantha*. Plantation species included *Eucalyptus grandis, E. saligna, E. camaldulensis* and *E. urophylla* (VanInsberghe *et al.*, 2015). It is also worth noting that the Kenyan dataset exhibited the highest variability of beta-diversity scores, which reflect the variety of ecoregions sampled in this study.

From the LDA comparison results of which the significant compositional differences between national datasets were reflected, 178 taxa were distributed across the dataset of which fourteen of these were dominated in Kenyan forest soils, including the Archeal genus *Nitrososphaera*. Other over-represented genera with a potential ecological relevance to Kenya forest soils included *Bradyrhizobium*, which is positively associated with soil health (VanInsberghe *et al.*, 2015). and *Chitinophaga* which is a group of chitinolytic *Myxobacteria* known to control fungal populations in the soil (Mehlich, 1984). Several of the taxa over-represented in Kenyan soil dataset

belonged to uncultured groups of bacteria, including members of uncultured genera *TK-10* and *Ellin606*, an indication that Kenyan forest soils harbor a catalogue of novel taxa. During development of bio-conservation strategies in these forest regions, consideration of these distinct microbiomes with unique taxa should be taken into account.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

1. Taxonomic classification of the obtained ASVs were assigned to 1 Kingdom, 2 Phyla, 120 Classes, 280 Orders, 450 Families, 873 Genera and 2313 Species within selected forests ecosystem.
2. Western and Taita Taveta regions had significantly different (P= 0.0124603) levels of Archaeal richness, Western and Aberdare regions soil displayed Archaeal Shannon diversity index (P=0.0199513) but there were no significant differences between bacterial communities displayed within various forests ecosystems.
3. Samples from the different ecoregions showed significant differences (p-value=0.001998, R^2=0.45) in soil physiochemical properties, specifically in soil pH, soil texture, macro- and micro-nutrient composition and Enhanced Vegetation Index-2.
4. The Kenyan dataset formed a distinct group with some degree of overlap with soil microbiomes from China, the Czech Republic, New Zealand and the USA, which could be a result of common plant cover between the sampled areas in the selected countries.

6.2 Recommendations

Based on the findings from this study, the following key recommendations are emphasized:

1. Employment of a combination of alpha diversity metrics, such as species richness and evenness, to quantify the within-sample diversity of prokaryotic communities in each forest ecosystem and utilize beta diversity measures, such as Bray-Curti's dissimilarity or UniFrac distances, to assess the between-sample diversity and compare community compositions among different forest sites.

2. Elucidation of the underlying ecological mechanisms driving microbial community dynamics in Kenyan forests ecosystems.
3. Collaboration with researchers from other countries to provide valuable insights into global patterns of microbial diversity and ecosystem functioning across diverse forests ecosystem worldwide.

REFERENCES

Abdillahi, H., Finnie, J., & Van Staden, J. (2011). Anti-inflammatory, antioxidant, anti-tyrosinase and phenolic contents of four Podocarpus species used in traditional medicine in South Africa. *Journal of ethnopharmacology*, *136*(3), 496-503.

Abera, T. A., Vuorinne, I., Munyao, M., Pellikka, P. K., & Heiskanen, J. (2022). Land cover map for multifunctional landscapes of taita taveta county, kenya, based on sentinel-1 radar, sentinel-2 optical, and topoclimatic data. *Data*, *7*(3), 36.

Afrin, T., Tsuzuki, T., Kanwar, R., & Wang, X. (2012). The origin of the antibacterial property of bamboo. *Journal of the Textile Institute*, *103*(8), 844-849.

Aleklett, K., Kiers, E. T., Ohlsson, P., Shimizu, T. S., Caldas, V. E., & Hammer, E. C. (2018). Build your own soil: exploring microfluidics to create microbial habitat structures. *The ISME journal*, *12*(2), 312-319.

Ali, N., Abbas, S. A. A. A., Sharif, L., Shafiq, M., Kamran, Z., Haseeb, M., & Shahid, M. A. (2024). Microbial extracellular polymeric substance and impacts on soil aggregation. In *Bacterial Secondary Metabolites* (pp. 221-237): Elsevier.

Anderson, M. J. (2005). Permutational multivariate analysis of variance. *Department of Statistics, University of Auckland, Auckland*, *26*, 32-46.

Anyona, S., & Rop, B. (2022). *Environmental impacts of artisanal and small-scale mining in Taita Taveta County*. Paper presented at the Proceedings of the Sustainable Research and Innovation Conference.

Armstrong, R. A., & Hilton, A. (2004). The use of analysis of variance (ANOVA) in applied microbiology. *microbiology, 2004*, 18 - 21.

Aşçı, F., Bahadır, M., & Akkuş, G. U. (2015). Study on the impact of elements in water on the diversity of water mites (Acari, Hydrachnidia) species. *Advances in Bioscience and Biotechnology, 6*(04), 259.

Babalola, O. O. J. B. l. (2010). Beneficial bacteria of agricultural importance. *Biotechnology letters, 32*(11), 1559-1570.

Bach, E. M., Williams, R. J., Hargreaves, S. K., Yang, F., & Hofmockel, K. S. (2018). Greatest soil microbial diversity found in micro-habitats. *Soil biology and Biochemistry, 118*, 217-226.

Backes, M. M. (2001). The role of indigenous trees for the conservation of biocultural diversity in traditional agroforestry land use systems: the Bungoma case study: in-situ conservation of indigenous tree species. *J Agroforestry Systems, 52*(2), 119-132.

Bailey, V. L., Fansler, S. J., Stegen, J. C., & McCue, L. A. (2013). Linking microbial community structure to β-glucosidic function in soil aggregates. *The ISME journal, 7*(10), 2044-2053.

Baldrian, P. (2017). Forest microbiome: diversity, complexity and dynamics. *J FEMS Microbiology reviews, 41*(2), 109-130.

Baldrian, P., Kolařík, M., Štursová, M., Kopecký, J., Valášková, V., Větrovský, T., ... & Voříšková, J. (2012). Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *The ISME journal, 6*(2), 248-258.

Bastian, F., Alabouvette, C., & Saiz-Jimenez, C. (2009). Bacteria and free-living amoeba in the Lascaux Cave. *Research in Microbiology, 160*(1), 38-40.

Berlemont, R., & Martiny, A. C. (2015). Genomic potential for polysaccharide deconstruction in bacteria. *J Applied environmental microbiology, 81*(4), 1513-1519.

Bickel, S., Chen, X., Papritz, A., & Or, D. (2019). A hierarchy of environmental covariates control the global biogeography of soil bacterial richness. *Scientific Reports, 9*(1), 12129.

Biederman, L. A., & Harpole, W. S. (2013). Biochar and its effects on plant productivity and nutrient cycling: a meta-analysis. *GCB bioenergy*, *5*(2), 202-214.

Bleher, B., Uster, D., & Bergsdorf, T. (2006). Assessment of threat status and management effectiveness in Kakamega Forest, Kenya. *Forest diversity and management*, 99-117.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., . . . & Asnicar, F. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology, 37*(8), 852-857.

Bouyoucos, G. J. (1962). Hydrometer method improved for making particle size analyses of soils 1. *Agronomy journal*, *54*(5), 464-465.

Brady, N. C., Weil, R. R., & Weil, R. R. (2008). *The nature and properties of soils* (Vol. 13): Upper Saddle River, NJ.: Prentice Hall.

Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecological monographs, 27*(4), 326-349.

Bremner, J. M. (1996). Nitrogen-total. *Methods of soil analysis: Part 3 Chemical methods, 5*, 1085-1121.

Bull, A. T. (2004). *Microbial diversity and bioprospecting*. (pp. xix+-496).

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods*, *13*(7), 581-583.

Cao, Y., Dong, Q., Wang, D., Zhang, P., Liu, Y., & Niu, C. (2022). microbiomeMarker: an R/Bioconductor package for microbiome marker identification and visualization. *Bioinformatics, 38*(16), 4027-4029.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., . . . & Bauer, M. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *J The ISME journal, 6*(8), 1621-1624.

Chao, A., Gotelli, N. J., Hsieh, T., Sander, E. L., Ma, K., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological monographs, 84*(1), 45-67.

Chen, P., Zhang, C., Ju, X., Xiong, Y., Xing, K., & Qin, S. (2019). Community composition and metabolic potential of endophytic Actinobacteria from coastal salt marsh plants in Jiangsu, China. *J Frontiers in microbiology, 10*, 1063.

Chu, D. (2018). *Effects of heavy metals on soil microbial community*. Paper presented at the IOP Conference Series: Earth and environmental science.

Cowan, D. A., Lebre, P. H., Amon, C., Becker, R. W., Boga, H. I., Boulange, A., . . . Zeze, A. (2022). Biogeographical survey of soil microbiomes across sub-Saharan Africa: structure, drivers, and predicted climate-driven changes. *Microbiome, 10*(1), 131. doi:10.1186/s40168-022-01297-w

Curd, E. E., Martiny, J. B., Li, H., & Smith, T. B. J. E. (2018). Bacterial diversity is positively correlated with soil heterogeneity. *9*(1), e02079.

Cycoń, M., Mrozik, A., & Piotrowska-Seget, Z. (2019). Antibiotics in the soil environment—degradation and their impact on microbial activity and diversity. *Frontiers in Microbiology, 10*, 338.

Dacal, M., García-Palacios, P., Asensio, S., Wang, J., Singh, B. K., & Maestre, F. T. (2022). Climate change legacies contrastingly affect the resistance and resilience of soil microbial communities and multifunctionality to extreme drought. *Functional Ecology, 36*(4), 908-920.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., . . . & Vorholt, J. A. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proceedings of the National Academy of Sciences, 106*(38), 16428-16433.

Didan, A. (2018). *VIIRS/NPP Vegetation Indices 16-Day L3 Global 500m SIN Grid V001*. Oak Ridge, TN, USA: NASA EOSDIS Land Processes DAAC.

Ding, G.-C., Piceno, Y. M., Heuer, H., Weinert, N., Dohrmann, A. B., Carrillo, A., . . . & Smalla, K. J. P. o. (2013). Changes of soil bacterial diversity as a consequence of agricultural land use in a semi-arid ecosystem. *PLOS one, 8*(3), e59497.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., . . . & Leitão, P. J. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*(1), 27-46.

Dubey, G., Kollah, B., Ahirwar, U., Sneh, T., & Mohanty, S. (2015). Significance of Archaea in terrestrial biogeochemical cycles and global climate change. *African Journal of Microbiology Research, 9*, 201-208. doi:10.5897/AJMR2014.7036

Dunivin, T. K., & Shade, A. J. F. m. e. (2018). Community structure explains antibiotic resistance gene dynamics over a temperature gradient in soil. *FEMS microbiology ecology, 94*(3), fiy016.

Eichlerová, I., Homolka, L., Žifčáková, L., Lisá, L., Dobiášová, P., & Baldrian, P. (2015). Enzymatic systems involved in decomposition reflects the ecology and taxonomy of saprotrophic fungi. *J Applied microbiology, 13*, 10-22.

Eichorst, S. A., & Kuske, C. R. (2012). Identification of cellulose-responsive bacterial and fungal communities in geographically and edaphically different soils by using stable isotope probing. *J Applied environmental microbiology, 78*(7), 2316-2327.

Eldridge, D. J., Delgado-Baquerizo, M., Travers, S. K., Val, J., Oliver, I., Hamonts, K., & Singh, B. K. J. E. (2017). Competition drives the response of soil microbial diversity to increased grazing by vertebrate herbivores. *Ecology, 98*(7), 1922-1931.

Everard, M., Johnston, P., Santillo, D., & Staddon, C. (2020). The role of ecosystems in mitigation and management of Covid-19 and other zoonoses. *Environmental science & policy, 111*, 7-17.

Fierer, N., Bradford, M. A., & Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology, 88*(6), 1354-1364.

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *J Proceedings of the National Academy of Sciences, 103*(3), 626-631.

Filimon, M. N., Voia, S. O., Popescu, R., Dumitrescu, G., Ciochina, L. P., Mituletu, M., & Vlad, D. C. J. R. B. L. (2015). The effect of some insecticides on soil microorganisms based on enzymatic and bacteriological analyses. *Romanian Biotechnological Letters, 20*(3), 10439-10447.

Griffiths, B. S., & Philippot, L. (2013). Insights into the resistance and resilience of the soil microbial community. *FEMS Microbiology reviews, 37*(2), 112-129.

Han, X.-m., Wang, R.-q., Jian, L., Wang, M.-c., Juan, Z., & Guo, W.-h. (2007). Effects of vegetation type on soil microbial community structure and catabolic diversity assessed by polyphasic methods in North China. *Journal of environmental Sciences, 19*(10), 1228-1234.

Handelsman, J. (2004). Metagenomics: application of genomics to uncultured microorganisms. *J Microbiology molecular biology reviews*, *68*(4), 669-685.

Hardoim, P. R., Van Overbeek, L. S., Berg, G., Pirttilä, A. M., Compant, S., Campisano, A., ... & Sessitsch, A. (2015). The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiology and molecular biology reviews*, *79*(3), 293-320.

He, R., Yang, K., Li, Z., Schädler, M., Yang, W., Wu, F., ... & Xu, Z. (2017). Effects of forest conversion on soil microbial communities depend on soil layer on the eastern Tibetan Plateau of China. *PloS one*, *12*(10), e0186053.

Hemkemeyer, M., Dohrmann, A. B., Christensen, B. T., & Tebbe, C. C. (2018). Bacterial preferences for specific soil particle size fractions revealed by community analyses. *Frontiers in Microbiology*, *9*, 149.

Hollander, M., & Wolfe, D. A. (1973). *Nonparametric statistical Methods*, New York: John W iley & Sons. In: Inc.

Hussain, S., Min, Z., Xiuxiu, Z., Khan, M. H., Lifeng, L., & Hui, C. (2019). Significance of Fe (II) and environmental factors on carbon-fixing bacterial community in two paddy soils. *Ecotoxicology and Environmental Safety*, *182*, 109456.

Ivanova, I., Kabadjova, P., Pantev, A., Danova, S., & Dousset, X. (2000). Detection, purification and partial characterization of a novel bacteriocin substance produced by Lactococcus lactis subsp. lactis B14 isolated from boza-Bulgarian traditional cereal beverage. *Biocatalysis*, *41*(6), 47-53.

Jacobsen, K. L., Niewolny, K. L., Schroeder-Moreno, M. S., Van Horn, M., Harmon, A. H., Fanslow, Y. H. C., ... & Parr, D. (2012). Sustainable agriculture undergraduate degree programs: A land-grant university mission. *Journal of Agriculture, Food Systems, and Community Development*, *2*(3), 13-26.

Jakobsson, H. E., Abrahamsson, T. R., Jenmalm, M. C., Harris, K., Quince, C., Jernberg, C., ... & Andersson, A. F. (2014). Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section. *Gut*, *63*(4), 559-566.

Janssen, P. H. (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Applied and environmental microbiology*, *72*(3), 1719-1728.

Jari, O., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., & McGlinn, D. (2020). Peter R. *Minchin, RB O'Hara, Gavin L. Simpson, Peter Solymos, M., Henry H. Stevens, Eduard Szoecs and Helene Wagner Oksanen Jari, & Oksanen Jari. vegan: Community Ecology Package. Retrieved from https://CRAN. R-project. org/package= vegan*.

Jolliffe, I. T. (2002). *Principal component analysis for special types of data*. New York: Springer.

Kalam, S., Basu, A., Ahmad, I., Sayyed, R., El-Enshasy, H. A., Dailin, D. J., & Suriani, N. L. (2020). Recent understanding of soil acidobacteria and their ecological significance: a critical review. *J Frontiers in microbiology*, 2712.

Kanyanjua, S., Ireri, L., Wambua, S., & Nandwa, S. (2002). *Acidic soils in Kenya: Constraints and remedial options.* Nairobi: KARI.

Karanja, E. N., Fliessbach, A., Adamtey, N., Kambura, A. K., Musyoka, M., Fiaboe, K., & Mwirichia, R. (2020). Diversity and structure of prokaryotic communities within organic and conventional farming systems in central highlands of Kenya. *J PloS one*, *15*(8), e0236574.

Karimi, B., Maron, P. A., Chemidlin-Prevost Boure, N., Bernard, N., Gilbert, D., & Ranjard, L. (2017). Microbial diversity and ecological networks as indicators of environmental quality. *Environmental Chemistry Letters*, *15*, 265-281.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, *30*(4), 772-780.

Kilonzi, F. M., & Ota, T. (2019). Ecosystem service preferences across multilevel stakeholders in co-managed forests: Case of Aberdare protected forest ecosystem in Kenya. *One Ecosystem*, *4*, e36768.

Kittredge, J. (2015). Soil Carbon Restoration: Can Biology do the Job?. *NOFA/Mass*.

Kong, A. Y., Six, J., Bryant, D. C., Denison, R. F., & Van Kessel, C. (2005). The relationship between carbon input, aggregation, and soil organic carbon stabilization in sustainable cropping systems. *Soil science society of America journal*, *69*(4), 1078-1085.

Krashevska, V., Klarner, B., Widyastuti, R., Maraun, M., & Scheu, S. (2015). Impact of tropical lowland rainforest conversion into rubber and oil palm plantations on soil microbial communities. *Biology and fertility of soils*, *51*, 697-705.

Krause, T., & Nielsen, M. R. (2019). Not seeing the forest for the trees: the oversight of defaunation in REDD+ and global forest governance. *Forests*, *10*(4), 344.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, *47*(260), 583-621.

Kumar, B. L., & Gopal, D. S. (2015). Effective role of indigenous microorganisms for sustainable environment. *J 3 Biotech*, *5*(6), 867-876.

Lakshmanan, V., Selvaraj, G., & Bais, H. P. (2014). Functional soil microbiome: belowground solutions to an aboveground problem. *Plant physiology*, *166*(2), 689-700.

Lauber, C. L., Hamady, M., Knight, R., & Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *J Applied environmental microbiology*, *75*(15), 5111-5120.

Legendre, P., Oksanen, J., & ter Braak, C. J. (2011). Testing the significance of canonical axes in redundancy analysis. *Methods in Ecology and Evolution, 2*(3), 269-277.

Li, H., Ye, D., Wang, X., Settles, M. L., Wang, J., Hao, Z., ... & Ma, Z. (2014). Soil bacterial communities of different natural forest types in Northeast China. *Plant and Soil*, *383*, 203-216.

Lilly, A., & Lin, H. (2004). Using soil morphological attributes and soil structure in pedotransfer functions. *Developments in soil science*, *30*, 115-141.

Lladó, S., López-Mondéjar, R., & Baldrian, P. (2017). Forest soil bacteria: diversity, involvement in ecosystem processes, and response to global change. *J Microbiology Molecular Biology Reviews, 81*(2).

López-Mondéjar, R., Voříšková, J., Větrovský, T., & Baldrian, P. (2015). The bacterial community inhabiting temperate deciduous forests is vertically stratified and undergoes seasonal dynamics. *J Soil Biology Biochemistry, 87*, 43-50.

Lung, T., & Schaab, G. (2004). Change-detection in Western Kenya: the documentation of fragmentation and disturbance for Kakamega forest and associated forest areas by means of remotelysensed imagery. *ISPRS archives*, *35*(Part B).

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50-60.

Maundu, P., & Tengnäs, B. (2005). Useful trees and shrubs for Kenya. *ICRAF Technical handbook series, 40*.

McArdle, B. H., & Anderson, M. J. (2001). Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology, 82*(1), 290-297.

McCauley, A., Jones, C., & Jacobsen, J. (2009). Soil pH and organic matter. *Nutrient management module, 8*(2), 1-12.

McLaren, D. G., Shah, V., Wisniewski, T., Ghislain, L., Liu, C., Zhang, H., & Saldanha, S. A. (2021). High-throughput mass spectrometry for hit identification: current landscape and future perspectives. *SLAS DISCOVERY: Advancing the Science of Drug Discovery*, *26*(2), 168-191.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one, 8*(4), e61217.

Meena, M., Yadav, G., Sonigra, P., Nagda, A., Mehta, T., Swapnil, P., . . . Kumar, S. (2023). Multifarious responses of forest soil microbial community toward climate change. *Microbial ecology*, *86*(1), 49-74.

Mehlich, A. (1984). Mehlich 3 soil test extractant: A modification of Mehlich 2 extractant. *J Communications in soil science plant analysis, 15*(12), 1409-1416.

Miyashita, N. T. (2015). Contrasting soil bacterial community structure between the phyla Acidobacteria and Proteobacteria in tropical Southeast Asian and temperate Japanese forests. *Genes & genetic systems, 90*(2), 61-77.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature, 403*(6772), 853-858.

Neina, D. (2019). The role of soil pH in plant nutrition and soil remediation. *Applied and Environmental Soil Science, 2019.*

Niemelä, T., & Pellikka, P. (2004). *Zonation and characteristics of the vegetation of Mt. Kenya.* Paper presented at the Taita Hills and Kenya, 2004-seminar, reports and journal of a field excursion to Kenya.

NJERI, W. F., Githaiga, J., & MWALA, A. K. (2018). The effects of fires on plants and wildlife species diversity and soil physical and chemical properties at Aberdare Ranges, Kenya. *Asian Journal of Forestry*, *2*(1), 25-38.

Nordman, E., & Mutinda, J. (2016). Biodiversity and wind energy in Kenya: Revealing landscape and wind turbine perceptions in the world's wildlife capital. *Energy Research & Social Science, 19*, 108-118.

Offre, P., Spang, A., & Schleper, C. (2013). Archaea in biogeochemical cycles. *Annual review of microbiology, 67*, 437-457.

Ogren, T. L. (2015). *The allergy-fighting garden: stop asthma and allergies with smart landscaping*, New York: Ten Speed Press.

Otieno, N. E., & Analo, C. (2012). Local indigenous knowledge about some medicinal plants in and around Kakamega forest in western Kenya. *F1000Research, 1*.

Pellikka, P., Heikinheimo, V., Hietanen, J., Schäfer, E., Siljander, M., & Heiskanen, J. (2018). Impact of land cover change on aboveground carbon stocks in Afromontane landscape in Kenya. *Applied Geography*, *94*, 178-189.

Pellikka, P. K., Lötjönen, M., Siljander, M., & Lens, L. (2009). Airborne remote sensing of spatiotemporal change (1955–2004) in indigenous and exotic forest cover in the Taita Hills, Kenya. *International Journal of Applied Earth Observation and Geoinformation, 11*(4), 221-232.

Philippot, L., Griffiths, B. S., & Langenheder, S. (2021). Microbial community resilience across ecosystems and multiple disturbances. *Microbiology and Molecular Biology Reviews, 85*(2), 10.1128/mmbr. 00026-00020.

Plante, A. F., Conant, R. T., Stewart, C. E., Paustian, K., & Six, J. (2006). Impact of soil texture on the distribution of soil organic matter in physical and chemical fractions. *Soil Science Society of America Journal*, *70*(1), 287-296.

Platts, P. J., Burgess, N. D., Gereau, R. E., Lovett, J. C., Marshall, A. R., McClean, C. J., . . . & Marchant, R. (2011). Delimiting tropical mountain ecoregions for conservation. *Environmental Conservation*, *38*(3), 312-324.

Prasad, S., Malav, L. C., Choudhary, J., Kannojiya, S., Kundu, M., Kumar, S., & Yadav, A. N. (2021). Soil microbiomes for healthy nutrient recycling. *Current trends in microbial biotechnology for sustainable agriculture*, 1-21.

Prevost-Boure, N. C., Maron, P. A., Ranjard, L., Nowak, V., Dufrene, E., Damesin, C., ... & Lata, J. C. (2011). Seasonal dynamics of the bacterial community in forest soils under different quantities of leaf litter. *Applied Soil Ecology*, *47*(1), 14-23.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., . . . & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research, 41*(D1), D590-D596.

Rocha, A. V., & Shaver, G. R. (2009). Advantages of a two band EVI calculated from solar and photosynthetically active radiation fluxes. *Agricultural and Forest Meteorology, 149*(9), 1560-1563.

Romaniuk, R., Lidia, G., Alejandro, C., Norberto, B., & Paolo, N. (2010). *A comparison of indexing methods to evaluate quality of soils subjected to*

different erosion: the role of soil microbiological properties. Paper presented at the EGU General Assembly Conference Abstracts.

Rosenberg, E. (2014). The family chitinophagaceae. *The prokaryotes*, 493-495.

Royston, J. P. (1982). An extension of Shapiro and Wilk's W test for normality to large samples. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 31*(2), 115-124.

Rozewicki, J., Li, S., Amada, K. M., Standley, D. M., & Katoh, K. (2019). MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic acids research, 47*(W1), W5-W10.

RStudio, T. (2020). RStudio: integrated development for R. *Rstudio Team, PBC, Boston, retrieved from http://www. rstudio. com*.

Saghaï, A., Banjeree, S., Degrune, F., Edlinger, A., García-Palacios, P., Garland, G., . . . Pescador, D. S. (2022). Diversity of archaea and niche preferences among putative ammonia-oxidizing Nitrososphaeria dominating across European arable soils. *Environmental Microbiology, 24*(1), 341-356.

Sahu, N., Vasu, D., Sahu, A., Lal, N., & Singh, S. (2017). Strength of microbes in nutrient cycling: a key to soil health. *Agriculturally Important Microbes for Sustainable Agriculture: Volume I: Plant-soil-microbe nexus*, 69-86.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome biology, 12*(6), 1-18.

Semedo, M., Song, B., Sparrer, T., & Phillips, R. L. (2018). Antibiotic effects on microbial communities responsible for denitrification and N2O production in grassland soils. *Frontiers in microbiology, 9*, 392571.

Sessitsch, A., Weilharter, A., Gerzabek, M. H., Kirchmann, H., & Kandeler, E. (2001). Microbial population structures in soil particle size fractions of a

long-term fertilizer field experiment. *Applied and environmental microbiology*, *67*(9), 4215-4224.

Shannon, C. E. (1949). *The mathematical theory of communication, by CE Shannon (and recent contributions to the mathematical theory of communication), W. Weaver,* Champaign, IL, USA: University of illinois Pres.

Shridhar, B. S. (2012). Nitrogen fixing microorganisms. *Int J Microbiol Res*, *3*(1), 46-52.

Siles, J. A., & Margesin, R. (2016). Abundance and diversity of bacterial, archaeal, and fungal communities along an altitudinal gradient in alpine forest soils: what are the driving factors? *%J Microbial ecology*, *72*(1), 207-220.

Stief, P., Kamp, A., Thamdrup, B., & Glud, R. N. (2016). Anaerobic nitrogen turnover by sinking diatom aggregates at varying ambient oxygen levels. *Frontiers in Microbiology*, *7*, 98.

Stieglmeier, M., Klingl, A., Alves, R. J., Simon, K.-M. R., Melcher, M., Leisch, N., & Schleper, C. (2014). Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt 8), 2738.

Štursová, M., Žifčáková, L., Leigh, M. B., Burgess, R., & Baldrian, P. (2012). Cellulose utilization in forest litter and soil: identification of bacterial and fungal decomposers. *J FEMS microbiology ecology*, *80*(3), 735-746.

Sundh, I., Wilcks, A., & Goettel, M. S. (2012). *Beneficial microorganisms in agriculture, food and the environment: safety assessment and regulation*: Cananda: CABI.

Sylvia, D. M., Fuhrmann, J. J., Hartel, P. G., & Zuberer, D. A. (2005). *Principles and applications of soil microbiology,* New Delhi*:* Pearson.

Tabish, M., Azim, S., Hussain, M. A., Rehman, S. U., Sarwar, T., & Ishqi, H. M. (2013). Bioinformatics approaches in studying microbial diversity. *Management of Microbial Resources in the Environment*, 119-140.

Tao, C., Wu, J., Liu, Y., Liu, M., Yang, R., & Lv, Z. (2018). Antimicrobial activities of bamboo (Phyllostachys heterocycla cv. Pubescens) leaf essential oil and its major components. *J European Food Research Technology, 244*(5), 881-891.

Team, R. C. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2019. 3. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one, 8*, e61217.

Team, R. C. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2012. In.

Thies, J. E., & Rillig, M. C. (2012). Characteristics of biochar: biological properties. In *Biochar for environmental management* (pp. 117-138), London: Routledge.

Totsche, K. U., Amelung, W., Gerzabek, M. H., Guggenberger, G., Klumpp, E., Knief, C., . . . Prechtel, A. (2018). Microaggregates in soils. *Journal of Plant Nutrition and Soil Science, 181*(1), 104-136.

Tourna, M., Stieglmeier, M., Spang, A., Könneke, M., Schintlmeister, A., Urich, T., . . . Richter, A. (2011). Nitrososphaera viennensis, an ammonia oxidizing archaeon from soil. *Proceedings of the National Academy of Sciences, 108*(20), 8420-8425.

Uroz, S., Buee, M., Deveau, A., Mieszkin, S., & Martin, F. (2016). Ecology of the forest microbiome: highlights of temperate and boreal ecosystems. *J Soil Biology Biochemistry, 103*, 471-488.

Van Der Heijden, M. G., Bardgett, R. D., & Van Straalen, N. M. (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology letters*, *11*(3), 296-310.

VanInsberghe, D., Maas, K. R., Cardenas, E., Strachan, C. R., Hallam, S. J., & Mohn, W. W. (2015). Non-symbiotic Bradyrhizobium ecotypes dominate North American forest soils. *The ISME journal*, *9*(11), 2435-2441.

Venables, W., & Ripley, B. (2002). *Modern applied statistics with S*, (4th edn). New York: Springer.

Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, T., ... & Young, I. M. (2016). Modeling soil processes: Review, key challenges, and new perspectives. *Vadose zone journal*, *15*(5), vzj2015-09.

Verginer, M., Siegmund, B., Cardinale, M., Müller, H., Choi, Y., Míguez, C. B., . . . Berg, G. (2010). Monitoring the plant epiphyte Methylobacterium extorquens DSM 21961 by real-time PCR and its influence on the strawberry flavor. *FEMS microbiology ecology*, *74*(1), 136-145.

Vermue, A., Philippot, L., Munier-Jolain, N., Hénault, C., & Nicolardot, B. (2013). Influence of integrated weed management system on N-cycling microbial communities and N 2 O emissions. *Plant and soil*, *373*, 501-514.

Vos, M., Wolf, A. B., Jennings, S. J., & Kowalchuk, G. A. (2013). Micro-scale determinants of bacterial diversity in soil. *FEMS microbiology reviews*, *37*(6), 936-954.

Wambugu, E. W. (2018). *Value of ecosystem services and socio-economic factors that enhance community participation in forest management in Aberdare forest, Kenya.* Nakuru: Egerton University,

Wang, B., Zheng, Y., Huang, R., Zhou, X., Wang, D., He, Y., & Jia, Z. (2014). Active ammonia oxidizers in an acidic soil are phylogenetically closely

related to neutrophilic archaeon. *Applied and environmental microbiology, 80*(5), 1684-1691.

Wang, M., Noor, S., Huan, R., Liu, C., Li, J., Shi, Q., . . . He, H. (2020). Comparison of the diversity of cultured and total bacterial communities in marine sediment using culture-dependent and sequencing methods. *PeerJ, 8*, e10060.

Weand, M. P., Arthur, M. A., Lovett, G. M., McCulley, R. L., & Weathers, K. C. (2010). Effects of tree species and N additions on forest floor microbial communities and extracellular enzyme activities. *Soil Biology and Biochemistry, 42*(12), 2161-2173.

Weber, B., Bowker, M., Zhang, Y., & Belnap, J. (2016). Natural recovery of biological soil crusts after disturbance. *Biological soil crusts: an organizing principle in drylands*, 479-498.

Wei, H., Peng, C., Yang, B., Song, H., Li, Q., Jiang, L., . . . Liu, S. (2018). Contrasting soil bacterial community, diversity, and function in two forests in China. *Frontiers in Microbiology, 9*, 1693.

Werner, D., & Newton, W. E. (2005). *Nitrogen fixation in agriculture, forestry, ecology, and the environment* (Vol. 4), Germany: Springer Science & Business Media.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biom. Bull., 1, 80–83. In. *statistics: Methodology and distribution* (pp. 196-202). New York: Springer

Wu, R.-N., Meng, H., Wang, Y.-F., Lan, W., & Gu, J.-D. (2017). A more comprehensive community of ammonia-oxidizing archaea (AOA) revealed by genomic DNA and RNA analyses of amoA gene in subtropical acidic forest soils. *Microbial ecology, 74*(4), 910-922.

Xia, Q., Rufty, T., & Shi, W. (2020). Soil microbial diversity and composition: Links to soil texture and associated properties. *J Soil Biology Biochemistry*, *149*, 107953.

Yohannes, Y., Shibistova, O., & Guggenberger, G. (2020). Comparative effects of dominant forest tree species on soil characteristics and microbial biomass. *J International Letters of Natural Sciences, 78*.

Zechmeister-Boltenstern, S., Michel, K., & Pfeffer, M. (2011). Soil microbial community structure in European forests in relation to forest type and atmospheric nitrogen deposition. *Plant and soil, 343*, 37-50.

Zhang, H., Gao, Z., Shi, M., Fang, S., Xu, H., Cui, Y., & Liu, J. J. W. (2019). Study of the Effects of Land Use on Hydrochemistry and Soil Microbial Diversity. *Water*, *11*(3), 466.

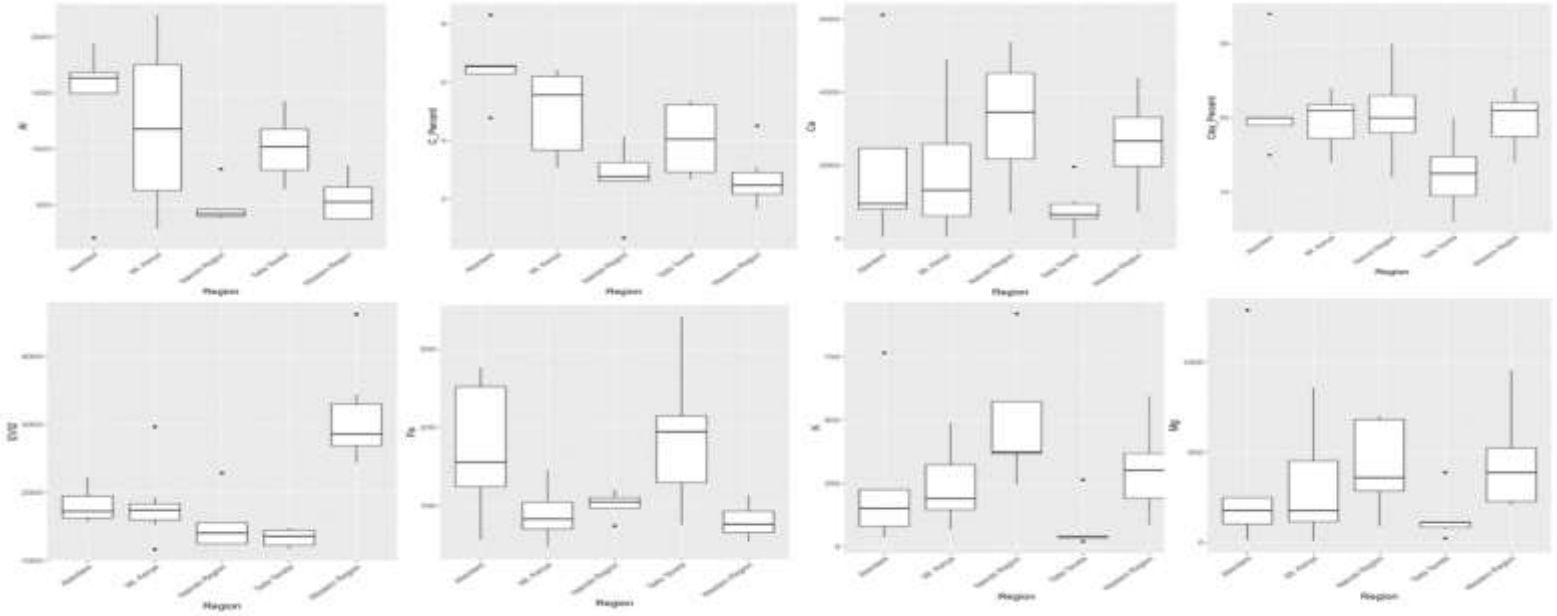
APPENDICES

Appendix I: Summaries of Phylum Level

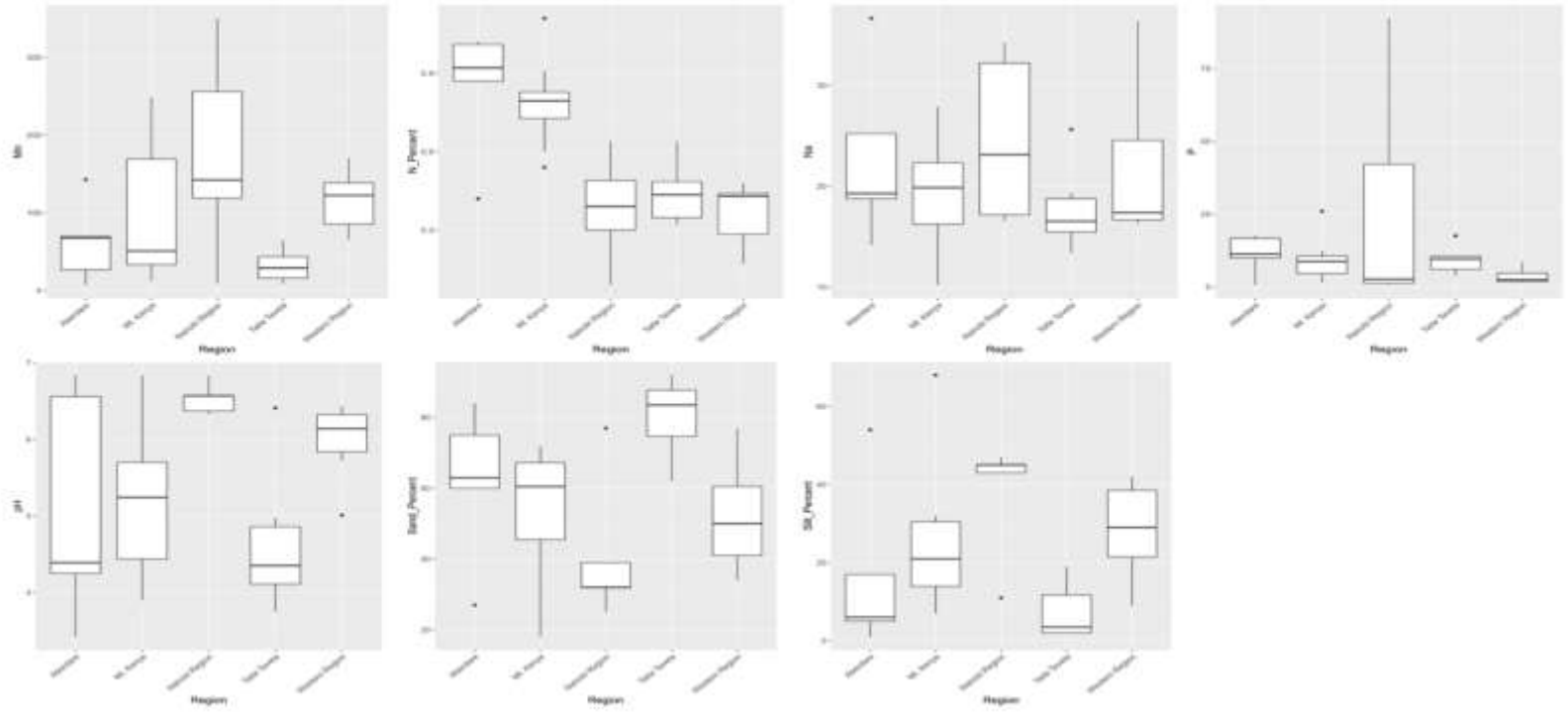
	K1	K1	K1	K1	K2	K2	K2	K2	K2	K2	K2	K2	K3	K3	K3	K3	K3	K3	K4	K4			K6	K6			K7	K7	K7		
Kingdom Taxa	0	5	6	8	1	3	4	5	6	7	8	9	3	4	5	6	7	8	0	2	K5	K6	3	6		K7	0	1	7	K8	K9
ArchaeaUnknown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	43	0	0
	887	132	747	723	769	189	25	15	11	62	22	32	444	159	246	247	161	146		617		620	564	181	108	201	894	102		131	
Archaea:Crenarchaeota	5	6	8	4	2	6	0	92	59	8	6	47	7	6	4	1	6	2	459	3	416	0	4	5	90	1	1	73	266	05	170
Archaea:Euryarchaeota	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	202	0	0	0	0	0	0	0	0
Archaea:Iainarchaeota	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	12	4	0
Archaea:Micrarchaeota	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0	0	0	0	0	0	0	0	0	0	2	0	0
Archaea:Nanoarchaeota	38	20	0	33	8	0	0	0	0	0	0	0	0	0	0	0	0	3	0	20	4	0	0	0	3	0	0	0	547	24	0
Archaea:Thermoplasmatot																				135						101					
a	175	43	329	34	12	371	0	0	0	0	0	23	0	405	276	159	0	0	42	0	0	842	795	345	0	2	40	42	266	259	17
								28		16		18																			
BacteriaUnknown	175	172	143	326	760	44	4	2	11	1	40	5	196	71	7	24	155	179	70	55	96	185	53	20	279	137	154	147	408	396	50
Bacteria:Abditibacteriota	4	7	44	8	0	0	0	0	7	0	0	0	0	5	4	15	0	0	3	17	0	0	0	0	3	0	9	24	20	14	0
	161	141	150	155	100	957	22	52	74	63	58	71	936	152	167	202	111	122	180	209	101	238	225	233	100	369	113	100	107	241	891
Bacteria:Acidobacteriota	41	25	03	24	37	1	94	56	91	47	64	45	7	99	95	54	52	30	11	06	05	74	73	34	04	13	98	28	97	16	9
	847	130	149	907	586	268	33	27	99	45	40	60	949	602	672	649	610	420	458	312	116	592	974	356	103	523	137	122	925	110	314
Bacteria:Actinobacteriota	2	84	38	5	5	4	00	88	14	72	51	88	3	8	7	4	5	7	0	4	60	2	6	2	04	6	00	44	1	51	9
									16																						
Bacteria:Armatimonadota	54	106	179	223	20	26	4	14	4	28	36	15	0	51	64	67	55	51	90	136	21	6	30	54	0	40	27	20	108	57	13
	103	845	692	901	621		48	87	40	50	32	45	907	540	472	462	873	845	745	287	880	240	554	660	578	271	628	744	101	838	
Bacteria:Bacteroidota	64	9	7	9	9	678	9	51	45	49	51	50	7	2	0	1	1	3	0	6	3	0	5	6	4	2	3	4	16	3	631
Bacteria:Bdellovibrionota	31	11	61	0	2	0	0	24	0	0	11	10	4	11	20	0	0	0	31	0	0	0	22	0	19	43	0	50	89	0	0
	258	252	346	265	119	138	49	94	28	21	13	10	169	189	365	186	244	142	228	532	174	107	142		182	258	190	209	263	263	
Bacteria:Chloroflexi	8	0	4	8	6	4	7	9	17	66	83	93	3	1	7	0	0	9	4	7	3	5	9	901	1	1	0	1	5	1	446
Bacteria:Cyanobacteria	103	18	196	96	7	37	0	0	36	0	18	14	35	87	54	225	57	100	49	93	41	79	55	125	32	212	62	659	168	98	39
Bacteria:Deinococcota	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	4	0	0	0
Bacteria:Dependentiae	0	0	0	0	12	9	0	0	0	0	0	0	23	0	0	23	16	0	49	3	9	54	30	27	0	27	0	0	0	45	6
Bacteria:Desulfobacterota	141	112	32	154	235	0	62	76	0	83	44	61	80	87	94	43	149	204	233	31	121	132	120	95	113	34	99	27	133	226	21
Bacteria:Elusimicrobiota	128	77	82	107	85	35	0	25	30	29	31	0	74	111	108	71	50	59	105	148	37	89	92	0	0	97	39	21	77	118	6
Bacteria:Entotheonellaeot								14				10																			
a	209	212	344	137	423	0	79	9	99	75	0	3	106	11	23	14	119	66	26	19	62	11	0	7	245	0	236	293	95	128	8

Bacteria:FCPU426	0	0	0	5	0	0	0	0	0	0	0	0	2	10	0	0	0	0	28	19	0	0	9	13	0	12	0	0	0	10	14
Bacteria:Fibrobacterota	0	11	16	23	0	0	0	0	0	3	0	0	16	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	23	0	0
			205					45	80	10	16	32																119			
Bacteria:Firmicutes	268	935	1	918	226	48	19	0	9	7	3	7	212	154	230	81	147	571	98	153	275	308	865	203	668	208	729	0	368	325	126
Bacteria:GAL15	7	0	0	0	0	41	0	0	0	0	0	0	0	16	43	18	8	0	0	14	0	50	50	12	0	0	0	0	0	49	0
Bacteria:Gemmatimonado	108		360					18	80	28	22	43																		130	
ta	3	871	8	909	372	129	93	7	2	5	6	7	452	462	559	467	400	453	406	590	518	734	726	235	316	243	630	590	929	9	115
Bacteria:Latescibacterota	100	89	0	45	80	0	29	20	0	26	0	28	0	7	31	0	33	97	97	20	4	0	0	0	36	0	10	0	0	82	0
Bacteria:MBNT15	37	6	9	0	71	0	35	0	0	0	0	6	15	0	0	0	8	8	77	3	19	0	14	0	23	0	11	0	32	15	0
Bacteria:Methylomirabilot					161		23	21		22		38																			
a	761	575	424	710	0	11	6	7	59	9	16	5	328	73	150	0	329	504	217	318	111	83	49	35	994	17	408	185	379	947	123
	209	168	162	150	141		32	65	79	78	77	90	192		103		142	138	178		152				220		151	127	179	149	
Bacteria:Myxococcota	7	1	2	2	1	17	4	5	9	0	2	8	7	804	8	344	3	6	3	521	5	385	779	614	7	299	2	5	6	1	176
Bacteria: NB1-j	98	21	0	19	141	11	14	15	0	32	0	11	60	6	5	0	24	59	83	43	0	7	0	0	70	0	4	13	6	16	0
								25		13		10																			
Bacteria:Nitrospirota	812	64	285	260	432	0	31	5	43	6	69	4	403	144	199	24	268	643	330	336	22	53	57	97	466	22	275	236	26	518	35
Bacteria:Patescibacteria	55	74	27	138	34	0	0	5	51	0	38	7	42	25	18	11	27	29	36	5	2	16	15	28	21	39	0	2	263	134	0
	723	634	889	597	363		33	15	26	20	13	18	482	290	279	314	263	355	451	276	435	270	419	243	333	312	402	343	819	620	
Bacteria:Planctomycetota	6	2	5	9	9	771	9	62	97	37	68	21	7	6	0	4	6	4	7	7	0	3	7	6	7	5	2	1	8	2	716
	218	171	126	192	178	113	31	86	90	93	92	93	194	224	246	274	178	167	234	162	224	218	288	202	160	238	161	106	220	336	132
Bacteria:Proteobacteria	53	59	15	18	95	80	81	82	33	84	50	56	55	00	15	57	20	04	44	80	19	20	22	93	19	43	69	72	82	99	78
																124						292	121			409					
Bacteria:RCP2-54	122	45	0	80	131	728	4	21	30	7	48	7	97	625	377	7	124	137	239	412	80	9	5	830	35	8	29	0	0	508	353
Bacteria:SAR324_clade																															
(Marine_group_B)	0	0	0	0	0	0	0	0	0	7	0	0	0	4	0	0	0	0	0	13	0	0	0	11	14	0	0	0	0	0	0
Bacteria:Spirochaetota	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	16	16	0	0	0	0	11	0	0	0	0	4	0	0
Bacteria:Sumerlaeota	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
Bacteria:Verrucomicrobio	518	951	814	773	453		41	18	58	26	26	18	236	206	226	243	427	495	357	127	411	380	493	149	305	694	467	261	459	115	
ta	1	2	9	2	4	519	4	07	18	77	89	85	8	3	9	5	0	8	6	0	5	6	8	4	5	2	3	1	5	16	897
						125										107															
Bacteria:WPS-2	18	9	0	32	2	6	0	0	39	0	39	0	0	665	424	9	57	66	122	761	7	873	355	555	0	412	0	0	0	165	213
EukaryotaUnknown	25	0	0	0	5	0	0	0	22	0	0	0	10	0	0	0	3	9	24	0	16	18	0	2	0	26	0	0	0	12	0
Eukaryota:Annelida	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0
Eukaryota:Basidiomycota	0	0	0	0	0	0	0	0	5	3	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	4	0	0	0	0	0
Eukaryota:Mucoromycota	5	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Eukaryota:Nematozoa	0	16	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0
Eukaryota:Phragmoplasto																															
phyta	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	0	0	0
UnassignedUnknown	0	0	0	4	0	0	0	0	0	2	16	0	0	17	28	36	6	2	7	8	0	10	8	10	0	43	0	0	7	4	0

Appendix IIa: Boxplots showing the differences in soil physicochemical properties of samples collected from the selected ecoregions in Kenya (Al, C, Ca, Clay, ENV2, Fe, K and Mg).



Appendix II b: Boxplots Showing the Differences in Soil Physicochemical Properties of Samples Collected from the Selected Ecoregions in Kenya (Mn, N, Na, P, pH, Sand and Silt).



Appendix III: Combined Sequence Datasets Read Counts Obtained from Selected Forests Around the Globe

	Feature	Enrich_group	ef_lda	pvalue	padj
marker1	Vicinamibacterales_uncultured_uncultured	Canada	4.343203	3.95E-32	3.95E-32
marker2	67-14	Canada	4.216335	1.91E-32	1.91E-32
marker3	Vicinamibacteraceae	Canada	4.106769	5.25E-30	5.25E-30
marker4	KD4-96	Canada	4.075852	2.61E-28	2.61E-28
marker5	Gemmatimonadaceae_uncultured	Canada	3.996026	2.39E-26	2.39E-26
marker6	Bacillus	Canada	3.993573	6.43E-27	6.43E-27
marker7	MB-A2-108	Canada	3.915554	3.04E-29	3.04E-29
marker8	Methyloligellaceae_uncultured	Canada	3.847186	4.16E-28	4.16E-28
marker9	Rubrobacter	Canada	3.846084	5.64E-47	5.64E-47
marker10	Solirubrobacter	Canada	3.750874	2.32E-30	2.32E-30
marker11	Nitrospira	Canada	3.734969	3.07E-34	3.07E-34
marker12	Pseudomonas	Canada	3.697968	6.71E-23	6.71E-23
marker13	Latescibacterota	Canada	3.696536	1.24E-28	1.24E-28
marker14	Gaiella	Canada	3.679635	8.75E-12	8.75E-12
marker15	Dongia	Canada	3.604339	2.69E-31	2.69E-31
marker16	Subgroup_17	Canada	3.584598	6.64E-23	6.64E-23
marker17	Subgroup_5	Canada	3.533124	1.86E-18	1.86E-18
marker18	Gitt-GS-136	Canada	3.52995	3.78E-22	3.78E-22
marker19	Sutterellaceae_uncultured	Canada	3.514275	7.23E-27	7.23E-27
marker20	Subgroup_11	Canada	3.491535	5.23E-26	5.23E-26
marker21	NB1-j	Canada	3.459975	1.83E-25	1.83E-25
marker22	Subgroup_22	Canada	3.426526	1.68E-28	1.68E-28
marker23	S085	Canada	3.415034	3.47E-16	3.47E-16
marker24	Steroidobacteraceae_uncultured	Canada	3.403029	2.15E-21	2.15E-21
marker25	Lysobacter	Canada	3.337233	5.84E-28	5.84E-28
marker26	Saprospiraceae_uncultured	Canada	3.326663	4.07E-23	4.07E-23

marker27	PLTA13	Canada	3.226748	2.53E-29	2.53E-29
marker28	Microvirga	Canada	3.208227	1.03E-31	1.03E-31
marker29	CCD24	Canada	3.160052	2.49E-17	2.49E-17
marker30	Actinomarinales_uncultured_uncultured	Canada	3.158015	2.70E-17	2.70E-17
marker31	Subgroup_10	Canada	3.10942	3.98E-17	3.98E-17
marker32	Candidatus_Alysiosphaera	Canada	3.102652	4.73E-19	4.73E-19
marker33	Entotheonellaceae	Canada	3.058928	4.41E-21	4.41E-21
marker34	Methyloceanibacter	Canada	3.045932	9.05E-10	9.05E-10
marker35	Burkholderia-Caballeronia-Paraburkholderia	China	4.446735	9.98E-37	9.98E-37
marker36	Elsterales_uncultured_uncultured	China	4.31735	6.83E-36	6.83E-36
marker37	Acidibacter	China	4.178201	1.61E-32	1.61E-32
marker38	Granulicella	China	3.842374	2.95E-29	2.95E-29
marker39	Candidatus_Koribacter	China	3.76237	4.03E-31	4.03E-31
marker40	Edaphobacter	China	3.545482	2.95E-28	2.95E-28
marker41	Dyella	China	3.441753	2.27E-37	2.27E-37
marker42	Phenylobacterium	China	3.35892	1.49E-20	1.49E-20
marker43	Caulobacteraceae_uncultured	China	3.353503	4.91E-31	4.91E-31
marker44	A21b	China	3.333883	2.43E-16	2.43E-16
marker45	Bacteroides	China	3.105294	6.63E-49	6.63E-49
marker46	Inquilinus	China	3.029197	1.75E-27	1.75E-27
marker47	Pedosphaera	China	3.016275	5.10E-24	5.10E-24
marker48	Blautia	China	3.004475	4.80E-53	4.80E-53
marker49	Candidatus_Udaeobacter	Czech_Republic	4.387196	5.04E-26	5.04E-26
marker50	Mycobacterium	Czech_Republic	4.093294	2.02E-39	2.02E-39
marker51	Sphingomonas	Czech_Republic	3.990824	4.23E-29	4.23E-29
marker52	Mucilaginibacter	Czech_Republic	3.946459	9.24E-30	9.24E-30
marker53	Mitochondria	Czech_Republic	3.915616	3.14E-35	3.14E-35
marker54	Nocardioides	Czech_Republic	3.826761	2.09E-33	2.09E-33
marker55	Puia	Czech_Republic	3.726668	9.46E-39	9.46E-39

marker56	Conexibacter Allorhizobium-Neorhizobium-	Czech_Republic	3.702955	6.75E-36	6.75E-36
marker57	Pararhizobium-Rhizobium	Czech_Republic	3.644512	1.52E-32	1.52E-32
marker58	Reyranella	Czech_Republic	3.61205	9.07E-24	9.07E-24
marker59	Ferruginibacter	Czech_Republic	3.524717	9.15E-22	9.15E-22
marker60	Isosphaeraceae_uncultured	Czech_Republic	3.498812	5.53E-31	5.53E-31
marker61	Massilia	Czech_Republic	3.454473	1.08E-34	1.08E-34
marker62	Pedobacter	Czech_Republic	3.452585	9.76E-18	9.76E-18
marker63	Pseudonocardia	Czech_Republic	3.433021	1.13E-27	1.13E-27
marker64	Actinoplanes	Czech_Republic	3.417413	7.61E-30	7.61E-30
marker65	Streptomyces	Czech_Republic	3.412552	3.91E-32	3.91E-32
marker66	Marmoricola	Czech_Republic	3.327408	5.52E-30	5.52E-30
marker67	Nakamurella	Czech_Republic	3.326777	1.39E-30	1.39E-30
marker68	Jatrophihabitans	Czech_Republic	3.321674	7.42E-25	7.42E-25
marker69	Luteibacter	Czech_Republic	3.320501	2.09E-24	2.09E-24
marker70	Kineosporia	Czech_Republic	3.297478	2.23E-26	2.23E-26
marker71	CL500-29_marine_group	Czech_Republic	3.286898	1.96E-20	1.96E-20
marker72	Devosia	Czech_Republic	3.27417	5.78E-33	5.78E-33
marker73	LWQ8	Czech_Republic	3.197647	1.08E-29	1.08E-29
marker74	Dokdonella	Czech_Republic	3.160625	4.36E-30	4.36E-30
marker75	Singulisphaera	Czech_Republic	3.12242	2.88E-27	2.88E-27
marker76	Caulobacter	Czech_Republic	3.098057	1.61E-23	1.61E-23
marker77	Sandaracinaceae_uncultured	Czech_Republic	3.002953	2.74E-16	2.74E-16
marker78	Acidobacteriales_uncultured_uncultured	Estonia	4.733093	9.73E-40	9.73E-40
marker79	Candidatus_Solibacter	Estonia	4.215271	7.94E-37	7.94E-37
marker80	Bryobacter	Estonia	4.193981	3.12E-28	3.12E-28
marker81	Occallatibacter	Estonia	3.949981	1.08E-40	1.08E-40
marker82	Micropepsaceae_uncultured	Estonia	3.944111	2.99E-35	2.99E-35

	Acidimicrobiia_uncultured_uncultured				
marker83	_uncultured	Estonia	3.770643	2.15E-32	2.15E-32
marker84	Pirellulaceae_uncultured	Estonia	3.595737	2.01E-31	2.01E-31
marker85	JG36-TzT-191	Estonia	3.558833	1.06E-39	1.06E-39
marker86	Pedosphaeraceae	Estonia	3.489879	3.38E-24	3.38E-24
marker87	Rhodanobacter	Estonia	3.369643	2.23E-20	2.23E-20
marker88	Subgroup_13	Estonia	3.252949	3.34E-36	3.34E-36
marker89	Opitutaceae_uncultured	Estonia	3.226034	7.91E-39	7.91E-39
marker90	Chloroplast	Estonia	3.181594	3.33E-24	3.33E-24
marker91	Lineage_IIa	Estonia	3.123337	7.40E-22	7.40E-22
marker92	Acidobacteriaceae_(Subgroup_1)_uncultured	Estonia	3.121794	7.17E-25	7.17E-25
marker93	Edaphobaculum	Estonia	3.059994	2.16E-25	2.16E-25
marker94	FCPU426	Estonia	3.011947	5.03E-36	5.03E-36
marker95	Subgroup_2	Finland_forest	4.84833	9.71E-37	9.71E-37
marker96	Aquisphaera	Finland_forest	4.259652	1.85E-46	1.85E-46
marker97	RCP2-54	Finland_forest	4.11165	2.93E-33	2.93E-33
marker98	Roseiarcus	Finland_forest	3.985683	7.88E-32	7.88E-32
marker99	Acetobacteraceae_uncultured	Finland_forest	3.82747	2.71E-41	2.71E-41
marker100	Acidipila	Finland_forest	3.802661	1.46E-34	1.46E-34
marker101	Group_1.1c	Finland_forest	3.533119	4.83E-23	4.83E-23
marker102	Paenibacillus	Finland_forest	3.347716	7.37E-28	7.37E-28
marker103	0319-6G20	Finland_forest	3.283848	7.89E-40	7.89E-40
marker104	Simkaniaceae_uncultured	Finland_forest	3.272134	9.16E-36	9.16E-36
marker105	Gemmata	Finland_forest	3.205191	9.69E-30	9.69E-30
marker106	Tundrisphaera	Finland_forest	3.179691	1.78E-31	1.78E-31
marker107	Aquicella	Finland_forest	3.178234	4.17E-32	4.17E-32
marker108	Legionella	Finland_forest	3.16242	6.30E-38	6.30E-38
marker109	JG30a-KF-32	Finland_forest	3.124962	3.14E-12	3.14E-12
marker110	Gimesiaceae_uncultured	Finland_forest	3.045269	7.34E-22	7.34E-22

marker111	Nitrososphaeraceae	Kenya	4.019443	2.21E-33	2.21E-33
marker112	Bradyrhizobium	Kenya	3.992296	3.80E-39	3.80E-39
marker113	RB41	Kenya	3.99105	8.93E-28	8.93E-28
marker114	Chitinophagaceae_uncultured	Kenya	3.612074	1.29E-39	1.29E-39
marker115	Ellin6067	Kenya	3.551251	9.53E-14	9.53E-14
marker116	Pir4_lineage	Kenya	3.456272	4.21E-27	4.21E-27
marker117	Terrimonas	Kenya	3.434053	3.15E-22	3.15E-22
marker118	Roseiflexaceae_uncultured	Kenya	3.370352	2.57E-21	2.57E-21
marker119	Steroidobacter	Kenya	3.292553	1.25E-28	1.25E-28
marker120	Pirellula	Kenya	3.26089	1.48E-22	1.48E-22
marker121	TK10	Kenya	3.23013	7.20E-20	7.20E-20
marker122	Chryseolinea	Kenya	3.085649	2.71E-24	2.71E-24
marker123	Chitinophaga	Kenya	3.017387	5.47E-21	5.47E-21
marker124	Niastella	Kenya	3.00643	2.10E-35	2.10E-35
marker125	Xanthobacteraceae_uncultured	New Zealand	4.275655	3.56E-30	3.56E-30
marker126	Candidatus_Xiphinematobacter	New Zealand	4.095831	5.17E-36	5.17E-36
marker127	HSB_OF53-F07	New Zealand	4.059941	3.96E-25	3.96E-25
marker128	IMCC26256	New Zealand	3.877842	6.20E-21	6.20E-21
marker129	Haliangium	New Zealand	3.586306	1.75E-16	1.75E-16
marker130	JG30-KF-AS9	New Zealand	3.502766	3.98E-34	3.98E-34
marker131	Ktedonobacteraceae_uncultured	New Zealand	3.501946	3.02E-21	3.02E-21
marker132	Ktedonobacter	New Zealand	3.334167	9.43E-13	9.43E-13
marker133	24-Nov	New Zealand	3.243162	2.40E-21	2.40E-21
marker134	cvE6	New Zealand	3.17313	2.31E-33	2.31E-33
marker135	Ellin516	New Zealand	3.159073	5.51E-23	5.51E-23
marker136	Hyphomicrobium	New Zealand	3.103506	1.23E-19	1.23E-19
marker137	CPla-3_termite_group	New Zealand	3.095726	3.47E-34	3.47E-34
marker138	Anaeromyxobacter	New Zealand	3.064348	1.41E-27	1.41E-27
marker139	Rhodomicrobium	New Zealand	3.053495	2.19E-24	2.19E-24

marker140	AKYH767	New Zealand	3.032064	2.92E-16	2.92E-16
marker141	SWB02	New Zealand	3.01557	9.20E-18	9.20E-18
marker142	Pedosphaeraceae_uncultured	Swiss_forest	4.24903	3.09E-44	3.09E-44
marker143	WD2101_soil_group	Swiss_forest	4.201785	9.21E-42	9.21E-42
marker144	ADurb.Bin063-1	Swiss_forest	4.005691	1.58E-25	1.58E-25
marker145	Anaerolineaceae_uncultured	Swiss_forest	3.896905	1.36E-16	1.36E-16
marker146	Rokubacteriales	Swiss_forest	3.887235	1.08E-30	1.08E-30
marker147	SC-I-84	Swiss_forest	3.826563	4.39E-23	4.39E-23
marker148	TRA3-20	Swiss_forest	3.791144	1.39E-18	1.39E-18
marker149	Candidatus_Nomurabacteria	Swiss_forest	3.76399	7.07E-48	7.07E-48
marker150	S-BQ2-57_soil_group	Swiss_forest	3.653539	4.33E-43	4.33E-43
marker151	Chthoniobacter	Swiss_forest	3.620861	1.57E-21	1.57E-21
marker152	Subgroup_7	Swiss_forest	3.594516	1.99E-23	1.99E-23
marker153	Candidatus_Kaiserbacteria	Swiss_forest	3.51445	1.82E-44	1.82E-44
marker154	MBNT15	Swiss_forest	3.484571	1.53E-20	1.53E-20
marker155	A4b	Swiss_forest	3.447831	9.40E-26	9.40E-26
marker156	Alphaproteobacteria_uncultured_uncultured_uncultured	Swiss_forest	3.423257	1.92E-22	1.92E-22
marker157	SBR1031	Swiss_forest	3.383359	6.51E-20	6.51E-20
marker158	Planctomycetales_uncultured_uncultured	Swiss_forest	3.364045	6.33E-21	6.33E-21
marker159	Anaerolinea	Swiss_forest	3.350888	5.78E-19	5.78E-19
marker160	KF-JG30-B3	Swiss_forest	3.303144	1.24E-09	1.24E-09
marker161	Parcubacteria	Swiss_forest	3.300209	6.96E-33	6.96E-33
marker162	CCM11a	Swiss forest	3.200941	2.94E-28	2.94E-28
marker163	RBG-13-54-9	Swiss_forest	3.153114	4.06E-17	4.06E-17
marker164	Candidatus_Zambryskibacteria	Swiss_forest	3.147815	1.01E-16	1.01E-16
marker165	OM190	Swiss_forest	3.117235	2.82E-22	2.82E-22
marker166	Candidatus_Adlerbacteria	Swiss forest	3.044237	4.38E-23	4.38E-23
marker167	Thermodesulfovibrionia_uncultured_uncultured_uncultured	Swiss forest	3.022322	2.80E-11	2.80E-11
marker168	Gemmataceae_uncultured	USA	4.364538	2.22E-40	2.22E-40

marker169	WD260	USA	4.177342	1.71E-32	1.71E-32
marker170	Pajaroellobacter	USA	3.733413	9.34E-35	9.34E-35
marker171	Solirubrobacteraceae_uncultured	USA	3.698867	2.57E-37	2.57E-37
marker172	AD3	USA	3.564567	1.50E-19	1.50E-19
marker173	WPS-2	USA	3.553667	2.76E-21	2.76E-21
marker174	Obscuribacteraceae	USA	3.537224	2.22E-32	2.22E-32
marker175	Methylacidiphilaceae_uncultured	USA	3.504631	2.83E-33	2.83E-33
marker176	Acidocella	USA	3.28184	1.37E-23	1.37E-23
marker177	Diplorickettsiaceae_uncultured	USA	3.281512	2.31E-35	2.31E-35

Appendix IV: Prior Information Consent

PRIOR INFORMED CONSENT (PIC) FOR THE SURVEY, SAMPLING AND COLLECTION OF SOIL SAMPLES IN DIFFERENT AGRO-ECOLOGICAL ZONES - IN KENYA

THIS Prior Informed Consent AGREEMENT is made and entered into this 30 day of July 2017, by and between:

The Provider:

1. Kenya Wildlife Service (KWS)

The Users:

1. Taita Taveta University (TTU)
2. University of Pretoria (UP)

WITNESSETH:

Access to biological resources and associated knowledge is governed by various Multilateral Environmental agreements (MEAS) including but not limited to the Convention on Biological Diversity (CBD), Nagoya Protocol, International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA), Convention on International Trade on Endangered Species (CITES), and World intellectual Property Organization Treaties (WIPO) among others.

That parties to CBD, CITES, Nagoya Protocol, ITPGRFA are obliged to comply with the provisions of the conventions through their domestic legislations, and that access to biological resources/genetic resources is subject to Prior Informed Consent (PIC) and Mutually Agreed Terms (MAT) between the provider and user.

Realizing that Kenya's Constitution 2010, Article 2(5) and 16) state that the general rules of international law and any treaty or convention ratified by Kenya shall form part of Kenyan law.

Recognizing that access to Kenya's Biological/genetic resources is subject to Access permit from National Environment Management Authority (NEMA), research clearance from National Commission for Science, Technology and Innovation (NACOSTI), Prior Informed Consent (PIC), Mutually Agreed Terms (MAT) and Material Transfer Agreement (MTA) from resources provider, in case of wildlife biological/genetic resources the resource provider is Kenya Wildlife Service, EMCA 2015, and ABS legal notice No. 160 of 1999 and Wildlife (Conservation Management) Act 2013.

TTU

Recognizing that access to South African resources are subject to the Convention on Biological Diversity (CBD), the Bioprospecting, Access and Benefit-Sharing (BABS: Act No. 10 of 2004) and the Nagoya Protocol (January 2013).

Whereas KWS is the Competent Authority on matters of wildlife where it manages and conserves the country's wildlife under the Wildlife Conservation and Management Act 2013.

Whereas KWS Grants PIC on wildlife user rights and it also administers Material Transfer Agreements on all wildlife specimens on behalf of the government and where possible PIC on wildlife is jointly granted between the Wildlife Management Authority, landowners and local communities;

Whereas TTU is an institution duly established in terms of the laws of the Republic of Kenya, under Universities Act, 2012; No. 42 of 2012 whose mission is to produce leaders and professionals in engineering, science and entrepreneurship through knowledge creation, dissemination and application for socio-economic development.

Whereas, the Principal Investigators, including Prof Hamadi Iddi Boga of the TTU of Kenya and Prof Don Cowan) of the (UP) of South Africa have interest in surveying, sampling and collecting specimens and samples for -microbial biodiversity studies.

Whereas various consultations between the resources provider (KWS) and Principal Investigators have been carried out via emails and phone calls.

NOW THEREFORE, IT IS HEREBY AGREED by the parties as follows:

i. Implementation of the Project titled "**The African Soil Microbiology project: A 'low resolution' microbial community survey of soils across sub-Saharan Africa**" is relevant for addressing key biological research gaps in Kenya and that the collaboration is a key component of Research and Development (RD). The project will throw light on key areas including status of soil health through detection of crop, animal and human pathogens within the soils and soil fertility.
ii. Knowledge of crop, animal and human pathogens in soil will generate vital data and information that will be useful in understanding ecological functions and biodiversity conservation across the various agro-ecological zones in Kenya.

HHB

SIGNED FOR THE PROVIDING INSTITUTION – KENYA WILDLIFE SERVICE

1. Director General
Kenya Wildlife Service
Postal Address: P.O. Box 40241
Nairobi, Kenya
Tel: +254-20-6000800

Deputy Director
Biodiversity Research & Monitoring
KENYA WILDLIFE SERVICE
P.O. Box 40241 - 00100, NAIROBI.

Name DR. FRANCIS GAKUYA

Signature ..... Date 30/07/2017

SIGNED FOR THE RECIPIENT

2.
Vice Chancellor
Taita Taveta University
P.O. Box 635-80300, Voi
Tel: + 254733926733
hamadibwega@ttu.ac.ke

Name Prof. Hamadi Boga

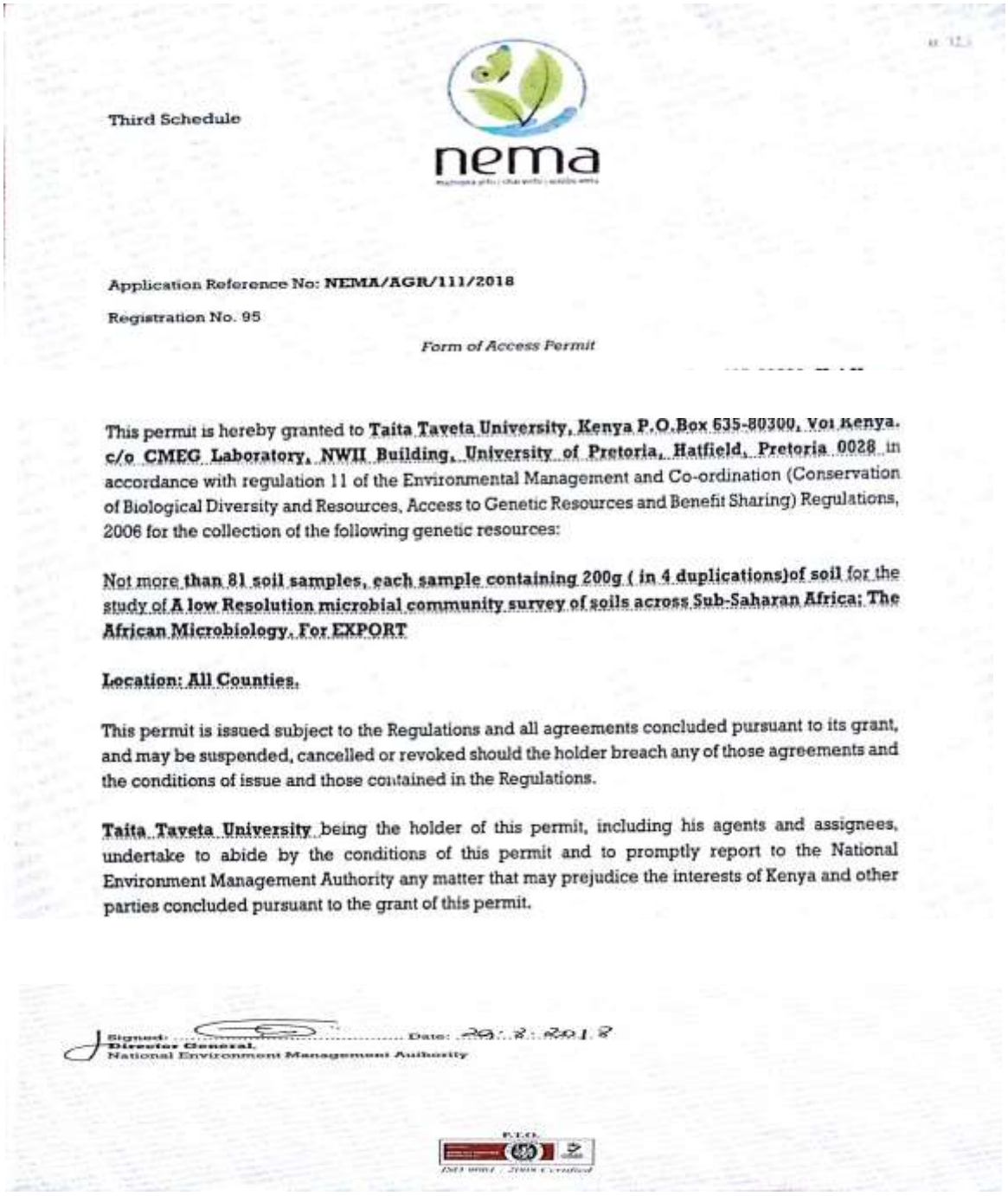
Signature ..... Date 29/7/17

b)
Department of Genetics
University of Pretoria
Hatfield 0002, Pretoria
Tel: + 27 12 420 5873
Fax: + 27 12 420 6870
Email: dorcas.mar@up.ac.za

Name Dorcas Arthur Cowan

Signature ..... Date 17/07/2017

Appendix V: NEMA Access Permit to Genetic Resources



Third Schedule

nema

Application Reference No: **NEMA/AGR/111/2018**

Registration No. 95

Form of Access Permit

This permit is hereby granted to **Taita Taveta University, Kenya P.O.Box 635-80300, Voi Kenya. c/o CMEG Laboratory, NWII Building, University of Pretoria, Hatfield, Pretoria 0028** in accordance with regulation 11 of the Environmental Management and Co-ordination (Conservation of Biological Diversity and Resources, Access to Genetic Resources and Benefit Sharing) Regulations, 2006 for the collection of the following genetic resources:

Not more than 81 soil samples, each sample containing 200g (in 4 duplications)of soil for the study of A low Resolution microbial community survey of soils across Sub-Saharan Africa; The African Microbiology. For EXPORT

Location: All Counties.

This permit is issued subject to the Regulations and all agreements concluded pursuant to its grant, and may be suspended, cancelled or revoked should the holder breach any of those agreements and the conditions of issue and those contained in the Regulations.

Taita Taveta University being the holder of this permit, including his agents and assignees, undertake to abide by the conditions of this permit and to promptly report to the National Environment Management Authority any matter that may prejudice the interests of Kenya and other parties concluded pursuant to the grant of this permit.

Signed: ..... Date: 29.2.2018
Director General,
National Environment Management Authority

Appendix VI: KWS Authorization Letter and Renewal Letter

KENYA WILDLIFE SERVICE

ISO 9001:2008 Certified

KWS/BRM/5001

20 August 2018

Dr. Anne Kelly Kambura
Taita Taveta University
P.O.Box 635-00800
VOI
e-mail: annnderitu@gmail.com
mobile: 0721235147

Dear Dr. Kambura,

RESEARCH EXTENSION PERIOD

We acknowledge receipt of your application requesting for research extension period to continue conducting research on a project titled: **'The African Soil Microbiology Project: A 'low resolution' microbial community survey across sub-Saharan Africa**. The study will continue generating vital data and information that will be useful in understanding ecological functions and biodiversity conservation across the various agro-ecological zones in Kenya.

We are pleased to inform you that you have been granted research extension to continue conducting the study from **September 2018 – August 2019** upon payment to KWS research extension fees of **ksh.6000**. During the research extension period, you will continue abiding by the set KWS regulations and guidelines regarding the carrying out of research in and outside protected areas. You will also be required to continue working closely with our Senior Scientist in-charge of the Conservation Area where the soil sampling activities will be carried out.

You will submit a copy of your research findings to the KWS Director, Biodiversity Research and Planning on completion of the study.

Yours Sincerely

PATRICK OMONDI, PhD, OGW
Ag. DIRECTOR
BIODIVERSITY RESEARCH AND PLANNING

Copy to:

- Senior Scientist, WCA, NCA, CRCA, MCA, ECA, CCA, TCA, SCA

P.O Box 40241-00100, Nairobi, Kenya.
Tel: +254-020-2379407/8/9-15. Mobile: +254-735 663 421, +254-726 610 508/9.
Email: kws@kws.go.ke Website:www.kws.go.ke

KENYA WILDLIFE SERVICE

ISO 9001:2008 Certified

KWS/BRM/5001

27 June 2017

Prof. Hamadi Iddi Boga
Taita Taveta University
P.O. Box 635-00800
VOI
e-mail: hamadiboga@yahoo.com

Dear Prof. Boga

PERMISSION TO CONDUCT TO CONDUCT RESEARCH ON MICROBIAL COMMUNITIES IN KENYAN SOILS

We acknowledge receipt of your letter dated 21st June 2017 requesting for permission to conduct research on a project titled: '**The African Soil Microbiology Project: A 'low' resolution microbial community survey of soils across Sub-Saharan Africa**'. The study will generate vital data and information that will be useful in understanding ecological functions and biodiversity conservation across the various agro-ecological zones in Kenya.

We are pleased to inform you that your team of researchers that includes **Dr. Anne Kelly Kambura** and **Mr. Vitalis Wafula Wekesa** has been granted permission to conduct the study from **July 2017 – June 2018** upon payment to KWS institutional research fees of **Ksh.12000** (Project research). However, you will obtain an Access Permit to Genetic Resources from the National Environment Management Authority (NEMA) and Research Clearance from the National Commission for Science Technology and Innovation (NACOSTI). You will also be required to obtain Prior Informed consent (PIC) and a Materials Transfer Agreement (MTA) from the genetic resource provider and a samples export permit from KWS in the event the samples will be shipped to South Africa for analysis.

We look forward to the successful implementation of the project and fruitful output from it.

Yours Sincerely.

SAMUEL M. KASIKI, PhD, OGW
DEPUTY DIRECTOR
BIODIVERSITY RESEARCH AND MONITORING

Copy to:
- Director-General NEMA
- Senior Scientist, Bio-prospecting/Biotechnology

P.O Box 40241-00100, Nairobi, Kenya. Tel: +254-20-2609233, +254-20-2609234
Wireless: +254-020-2379407-15. Mobile: +254-735 663 421, +254-726 610 508/9. Fax: +254-020-2661923
Email: kws@kws.go.ke Website:www.kws.go.ke

Appendix VII: NACOSTI Authorization letter

NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY AND INNOVATION

Telephone: +254-20-2213471,
2241349,3310571,2219420
Fax: +254-20-318245,318249
Email: dg@nacosti.go.ke
Website : www.nacosti.go.ke
When replying please quote

NACOSTI, Upper Kabete
Off Waiyaki Way
P.O. Box 30623-00100
NAIROBI-KENYA

Ref. No. **NACOSTI/P/18/02166/18443**

Date: **5th April, 2018**

Anne Kelly Kambura
Taita Taveta University College
P.O. Box 635-80300
VOI.

RE: RESEARCH AUTHORIZATION

Following your application for authority to carry out research on ***"A 'Low Resolution' microbial community survey of soils across Sub-Saharan Africa,"*** I am pleased to inform you that you have been authorized to undertake research in **all Counties** for the period ending **28th March, 2019.**

You are advised to report to **the County Commissioners and the County Directors of Education, all Counties** before embarking on the research project.

Kindly note that, as an applicant who has been licensed under the Science, Technology and Innovation Act, 2013 to conduct research in Kenya, you shall deposit **a copy** of the final research report to the Commission within **one year** of completion. The soft copy of the same should be submitted through the Online Research Information System.

BONIFACE WANYAMA
FOR: DIRECTOR-GENERAL/CEO

Copy to:

The County Commissioners
All Counties.

The County Directors of Education
All Counties.