

## OPEN ACCESS

## EDITED BY

Agnieszka Kuźniar,  
The John Paul II Catholic University of  
Lublin, Poland

## REVIEWED BY

Ees Ahmad,  
National Bureau of Agriculturally Important  
Microorganisms (ICAR), India  
Mariantonietta Colagiero,  
National Research Council (CNR), Italy

## \*CORRESPONDENCE

Wilton Mwema Mbinda  
✉ wilton.mbinda@gmail.com;  
✉ w.mbinda@pu.ac.ke

RECEIVED 24 February 2025

ACCEPTED 09 May 2025

PUBLISHED 30 May 2025

## CITATION

Muzami EM, Kitundu G, Muriithi OM,  
Kavoo AM, Gichuru VG and Mbinda WM  
(2025) Metagenomic insights to bacterial  
communities, functional traits, and soil health  
in banana smallholder agroecosystems of  
Kenya. *Front. Microbiol.* 16:1582271.  
doi: 10.3389/fmicb.2025.1582271

## COPYRIGHT

© 2025 Muzami, Kitundu, Muriithi, Kavoo,  
Gichuru and Mbinda. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License \(CC  
BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in  
other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Metagenomic insights to bacterial communities, functional traits, and soil health in banana smallholder agroecosystems of Kenya

Eugene Mwanza Muzami<sup>1</sup>, George Kitundu<sup>1</sup>,  
Oscar Mwaura Muriithi<sup>2</sup>, Agnes Mumo Kavoo<sup>3</sup>,  
Virginia Gathoni Gichuru<sup>2</sup> and Wilton Mwema Mbinda<sup>1,4\*</sup>

<sup>1</sup>Department of Biochemistry and Biotechnology, Pwani University, Kilifi, Kenya, <sup>2</sup>Department of Biological Sciences, Pwani University, Kilifi, Kenya, <sup>3</sup>Department of Horticulture, Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya, <sup>4</sup>Pwani University Biosciences Research Centre (PUBReC), Pwani University, Kilifi, Kenya

Microbes inhabiting the banana rhizosphere are varied and mediate essential functions that enhance plant growth and increase crop productivity. Their abundance in soil habitats is a potential indicator of soil health and quality. Despite the well-known benefits of rhizosphere microorganisms in banana cultivation, their genomic and functional diversity remains largely unexplored within smallholder agroecosystems. In this study, we characterized the community composition and functional potential of bacteria in banana rhizospheric soils from Gituamba, Mangu and Ngenda constituencies in Kiambu County, Kenya. Using Illumina Novaseq sequencing, we analyzed 16S rRNA gene amplicons and shotgun metagenomic profiles to explore these microbial communities. Variations of soil physicochemical parameters across the study sites were assessed. The parameters varied across the sites, with soils in Gituamba and Ngenda depicting better soil fertility characteristics than Mangu. Amplicon sequencing profiles revealed higher bacterial diversity in Gituamba compared to Mangu, while the single sample from Ngenda exhibited moderate diversity. The dominant phyla across the study sites were *Proteobacteria*, *Actinobacteria*, and *Cyanobacteria*. Functional profiling of 16S rRNA gene amplicons showed a higher enrichment in Gituamba compared to Mangu. Overall, the functional profiling indicated that predicted metabolic pathways across the study sites were linked to genes encoded by the members of the most abundant bacterial phyla in the soil environments, majorly contributing to beneficial roles for soil health and crop yield. This study offers methods to reveal the banana rhizosphere as a rich reservoir for potential microbes of agricultural and biotechnological significance, which can promote sustainable agriculture.

## KEYWORDS

bacteria, banana rhizosphere, metagenomics, soil physicochemical properties, sustainable agriculture

## Introduction

Banana (*Musa paradisiaca*) is a key crop in the global food industry, ranking among the top fruit crops in terms of both production and economic value across the world (Alemu, 2017). It is a vital staple food with high nutritional benefits attributed to its richness in potassium, carbohydrates and vitamin A (Muthee et al., 2019). Bananas serve as a staple food for both rural and urban populations in Kenya (Wahome et al., 2021). Banana cultivation is an important economic activity in Kenya. However, small-scale farmers' production does not meet the current market demand (Wahome et al., 2023). The low production is due to numerous challenges such as agronomic practices, poor soil quality and the impact of pests and diseases (Rossmann et al., 2012). Numerous studies suggest that factors impacting soil biodiversity, including changes in land use, depletion of soil organic matter, and land degradation, have contributed to the decline of both aboveground and belowground microbiome abundance (Dixit et al., 2024). These factors directly affect banana production and have led to increased attention to assess soil health, particularly bacterial composition in rhizosphere soil habitats. This reflects the understanding that soil bacteria are essential for improving the overall soil and plant production (Giri and Varma, 2020). A diverse array of microbes is crucial for soil functions, as these communities significantly contribute to sustainable agricultural practices. Therefore, soil microorganisms are recognized as mediators of numerous processes that enhance agricultural productivity.

Plants and their related microorganisms interact to create communities of genotypes known as holobionts (Sarkar et al., 2022). These microbes influence the health and functioning of the plant by modifying the supply of essential nutrients and resisting both biotic and abiotic stresses (Alawiye and Babalola, 2021). Plant roots not only grant mechanical support, but also facilitate absorption of water and nutrients, and exudate various chemical substances that enhance plant-microbiome interactions (Molefe et al., 2021). The rhizosphere is the narrow region that surrounds the roots of plants and houses numerous microorganisms that define variety of anabolic and catabolic activities within soils. Plants release organic compounds into the rhizosphere through root exudates, shaping complex metabolic interactions among organisms. These interactions can support plant growth, remain neutral, or even be detrimental, influencing the surrounding microbial community (Omotayo et al., 2021).

Soil hosts a large community of microbes per gram of soil which include fungi, bacteria, viruses and protozoa in the rhizosphere. Among rhizobiomes, bacteria are the most abundant and serve as key markers of soil health and productivity concerning their swift responses to alternating environmental changes (Enagbonma et al., 2020). Rhizospheric bacterial communities promote growth through various mechanisms, both directly and indirectly. Rhizosphere bacteria promote nitrogen fixation, serve as biological control agents against pathogens in the

soil, synthesize phytohormones and generate compounds like siderophores, antibiotics, cyanides, and ammonia (Alawiye and Babalola, 2019). The distribution of soil bacteria in various agroecosystems is primarily influenced by soil parameters such as soil texture, organic matter, organic carbon, soil nutrients and pH. Given that rhizospheric bacteria are highly responsive to alterations in the soil environment, shifts in their community structure indicate variations in soil biological activity and overall soil quality (He et al., 2023).

Although numerous studies have explored the banana root microbiome (Birt et al., 2022; Kaushal et al., 2020a,b), the relationship between soil microorganisms and soil health, particularly within smallholder banana farming systems, remains poorly understood. Identifying key microorganisms that shape the rhizosphere and understanding their functional traits are essential for advancing knowledge of soil-plant interactions in these agroecosystems. Moreover, shifts in the abundance and functional capabilities of soil bacterial communities under different farming conditions are still not well-characterized. To address these gaps, meta-omic approaches utilizing next-generation sequencing technologies are increasingly applied to analyze environmental metagenomes, allowing for the identification of microbial taxa and the discovery of functionally important genes and bioactive compounds (Nwachukwu and Babalola, 2022). In this study, we employed Illumina NovaSeq high throughput 16S rRNA gene amplicon and shotgun metagenomic sequencing to characterize the taxonomic composition and functional potential of bacterial communities in rhizosphere soils from smallholder banana farms in Kiambu County, Kenya. We also assessed variations in soil physicochemical properties across the study sites. We hypothesized that differences in soil physicochemical properties among sites would be associated with distinct bacterial community structures and functional profiles in the banana rhizosphere.

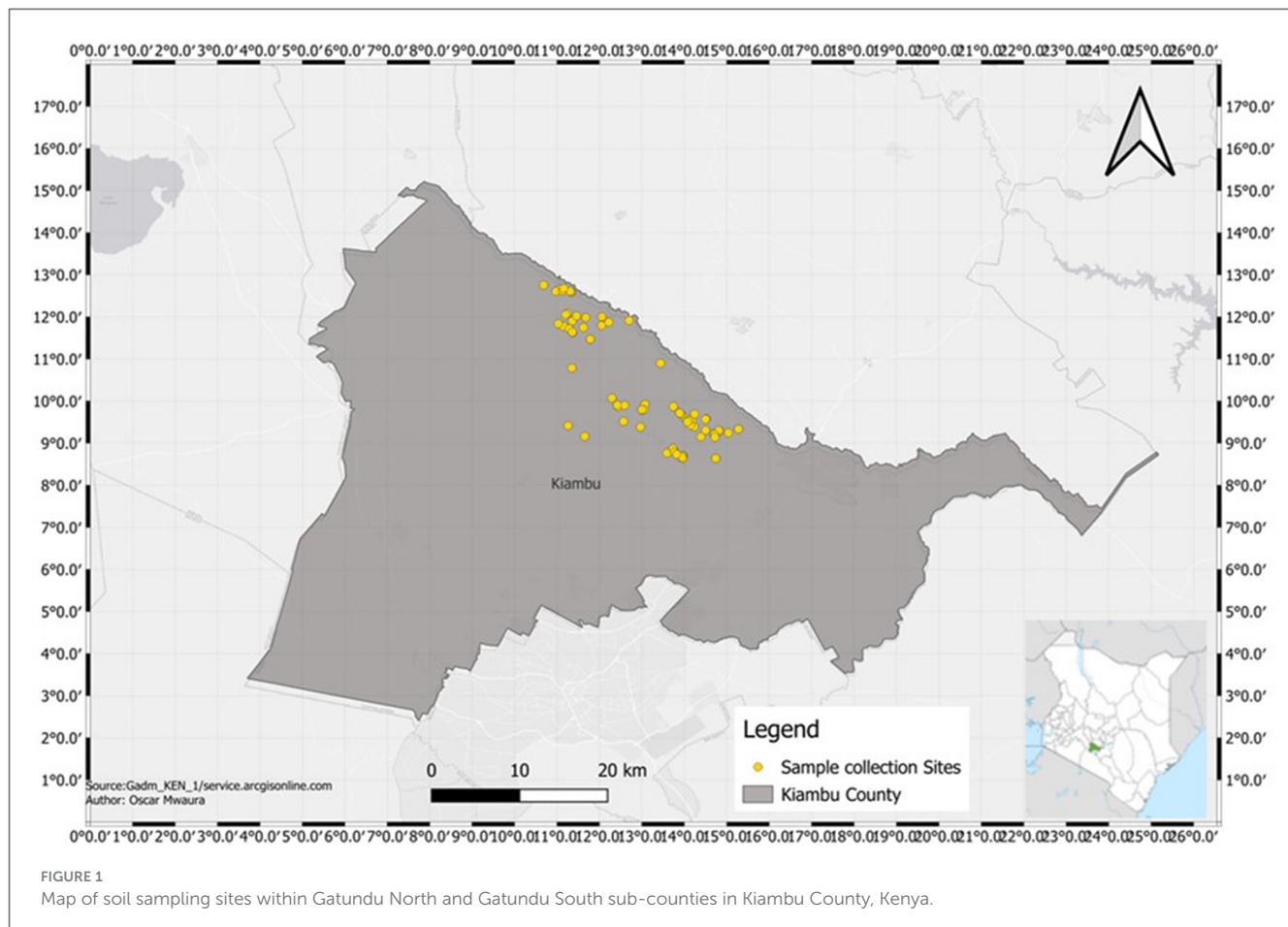
## Materials and methods

### Study site

The study was conducted in Kiambu County (36°49'0.0"E and 1°10'0.0"S) located in Kenya's central region (Figure 1) and covers a total area of 2,543.5 km<sup>2</sup> (Kutwa et al., 2016). It has four broad topographical zones: upper highland, lower highland, upper midland and lower midland zones (Njiru, 2019). Three study sites were investigated within the lower highland zone: Gituamba and Mangu wards in Gatundu North sub-County and Ngenda ward in Gatundu South sub-County. The County is characterized by an annual average rainfall of 1,200 mm and an annual mean temperature of 26°C with temperatures ranging from 7°C in the upper highland areas to 34°C in the lower midland zone. The altitude of the County lies between 1,500 and 1,800 m above sea level and is generally a tea and dairy zone with agricultural activities like maize, horticultural crops and sheep farming. The average relative humidity ranges from 54% in the dry months and 300% in the wet months of March up to August. A bi-modal type of rainfall: short rains (mid-October to November) and long rains (mid-March to May) are experienced in the County. The major categories of soils in the County are high-level upland soils, plateau soils and volcanic

---

**Abbreviations:** ABC, ATP-binding cassette; ECF, extracytoplasmic function; KOs, KEGG orthologs; OM, organic matter; SQM, sequence quality monitoring; TOC, total organic carbon; TPM, transcripts per million.



footbridge soils. These soils are of varying fertility levels with soils from high-level uplands, which are from volcanic rocks, being very fertile. Coffee and tea are the main cash crops in the county while the main food crops grown in the county include maize, beans, pineapples, potatoes, vegetables and bananas. In Kiambu County, banana production is primarily practiced in high-rainfall areas that have been under continuous cultivation for extended periods, often resulting in decreased yields (Nzioka, 2009).

## Collection of soil samples

Sampling was conducted according to the described method by Wahome et al. (2023) in October 2022. Sampling occurred in smallholder banana farms, where most farmers practice mixed cropping systems, typically integrating bananas with other crops such as maize, beans, and vegetables. Across all farms, banana plants sampled were mature, characterized by fruiting bunches, ensuring consistency in the plant growth stage and minimizing variability due to plant age or vegetative state. Rhizospheric soil samples were collected at a depth of 15–20 cm at ~100 m spacing from each collection point using a sterile hand shovel applying the cluster sampling method. Soil sampling was carried out across and diagonally from 26 points per sampling site in each study site. A total of 78 soil samples of 200 g each were mixed and packed in sterile Ziplock bags in a dry iced box before they were transported

and stored at  $-20^{\circ}\text{C}$  at the Pwani University Biosciences Research Center (PUBReC) lab for further analysis. One portion (half) of each of the 78 soil samples was air-dried at room temperature for 1 week and sieved through a 100-mesh sifter to remove stones and visible plant fragments. The soils were stored at  $4^{\circ}\text{C}$  for further physicochemical analysis at the soil mechanics laboratory at Jomo Kenyatta University of Agriculture and Technology. The remaining halves of the soil samples were pooled into three composite samples representing each study site and stored at  $-80^{\circ}\text{C}$  for soil genomic DNA extraction at Pwani University Biosciences Research Center (PUBReC).

## Soil physicochemical analysis

The measured soil parameters mainly included: the soil pH, soil water electrical conductivity, total organic carbon, organic matter, nitrogen, potassium and phosphorus content. Soil pH was measured using a soil pH meter (FE28, Mettler Toledo, Zurich, Switzerland) with a soil–water ratio of 1:2.5 (w/v) (Han et al., 2023). The electrical conductivity (EC) of the soil was measured using a portable analyzer (YSI, USA) (Tan et al., 2023). Total organic carbon (TOC) and available nitrogen (N) were determined by the sulphuric acid/potassium dichromate ( $\text{H}_2\text{SO}_4\text{-K}_2\text{Cr}_2\text{O}_7$ ) wet oxidation and the Kjeldahl methods, respectively (Li et al., 2022). TOC was analyzed using Shimadzu TOC-L CPH while N

was measured using Lachat QuikChem 8500 equipment. Available phosphorus (P) was analyzed by inductively coupled ICP-OES Optima 8000 plasma-optical emission spectroscopy (Perkin Elmer, Massachusetts, USA) (Li et al., 2022). The soil organic matter (OM) was measured using the potassium dichromate external heating method (Tan et al., 2023) while available potassium (K) was extracted using the flame photometric method via digestion with ammonium acetate (NH<sub>4</sub>OAC) using the Cole-Parmer (USA) Flame Photometer 1265 (Han et al., 2023).

## DNA extraction and sequencing

Total soil DNA was extracted by the cetyltrimethylammonium bromide (CTAB) method. The purity and quality of the genomic DNA samples were checked using 1.2% agarose gel electrophoresis, and the concentrations were measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher, Massachusetts, USA) before shipping the samples to Macrogen, South Korea for 16S rRNA gene amplicon and shotgun metagenomic sequencing using the Illumina Novaseq 6000 sequencing platform. The first two DNA samples (Gituamba and Mangu) were subjected to amplicon sequencing where the hypervariable region (V3–V4) was amplified using the 16S amplicon rRNA primer pair 515F (5'-GTGYCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACN VGGGTWCTAAT-3') while the third DNA sample (Ngenda) was subjected to shotgun metagenomic sequencing. For amplicon sequencing, libraries were created, and each sample underwent a single-step PCR with 35 cycles using the HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA). The PCR conditions were as follows: an initial denaturation at 95°C for 10 min, then 35 cycles comprising denaturation at 95°C for 30 s, annealing at 53°C for 40 s, and extension at 72°C for 1 min, with a final extension at 72°C for 10 min. After PCR, all amplicons from the samples were purified using SPRI beads. For shotgun metagenomic sequencing, library preparation followed the protocol for the Nextera DNA Flex kit (Illumina). Briefly, the sample began with 50 ng of DNA, which was fragmented and then tagged with Illumina sequencing adapters. Amplification of the libraries involved six PCR cycles. After amplification, library concentration was determined with the Qubit® dsDNA HS Assay Kit (Life Technologies), and average fragment size was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies). The libraries were then pooled at a final concentration of 0.7 nM in equimolar amounts and sequenced in a paired-end format for 300 cycles on the NovaSeq 6000 platform (Illumina). The paired-end raw reads for all three samples were stored in FASTQ format.

## Sequence analysis, taxonomic and functional classification

### Taxonomic classification of 16S rRNA profiles

Quantitative Insights into Microbial Ecology (QIIME v2024.2) was used to process the raw amplicon sequences (Bolyen et al., 2019). Raw data quality control was done using FASTQC v0.11.9 (Chen et al., 2017). Raw sequence data was imported into QIIME

as a manifest file using the q2-tools import plugin followed by quality filtering that involved adapter removal from the paired-end demultiplexed sequences using the q2-Cutadapt trim-paired plugin. Subsequently, the reads were denoised using the Divisive Amplicon Denoising Algorithm (DADA2). The q2-dada2 denoise-paired plugin was used to filter the paired-end reads based on the quality score and length of the sequences as well as the removal of chimeras and dereplicated sequences. This was followed by the calculation of denoising statistics that produced files containing a summary of denoising results, representative sequences, and the feature count table. Further, the q2-feature classifier and assigned taxonomy plugin were used to classify the representative sequences. First, a machine learning approach involving a Scikit-learn multinomial naïve Bayes classifier was trained on an integrated reference database: Greengenes, SILVA and RDP database (GSR-DB) (Molano et al., 2024). The classifier was trained on our target hypervariable region of interest, the V3–V4 of 16S rRNA sequences. Secondly, the classified representative sequences were assigned taxonomy using the trained classifier. Taxonomy barplots were created using the q2-taxa barplot plugin. A *de novo* phylogeny-based approach was used to create a phylogenetic tree using the MAFFT-FastTree method in which a multiple sequence alignment with MAFFT was followed by building the maximum-likelihood rooted tree with FastTree using the q2-phylogeny plugin. Alpha diversity metrics (Observed features, Shannon index, and phylogenetic distance) were calculated after samples were subsampled without replacement (rarefied) to 225,541 sequences per sample using the q2-diversity alpha-rarefaction plugin. Data was presented as a direct visualization of QIIME2 artifacts on QIIME2 View website (<https://view.qiime2.org>).

### Functional prediction of 16S rRNA profiles

Predictive functional profiling of bacterial communities was conducted using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States2 (PICRUSt2) v2.5.2 (Douglas et al., 2020). Representative sequences were aligned using HMMER followed by placement into a reference tree utilizing EPA-NG (Barbera et al., 2018) and gappa (Czech et al., 2020). The number of 16S rRNA gene copies was normalized and gene families were inferred with Castor, a tool for hidden state prediction (Ye and Doak, 2009). These predicted gene families were subsequently mapped to MetaCyc pathways using MinPath (Ye and Doak, 2009).

### Taxonomic and functional classification of shotgun metagenomic sequencing profiles

Shotgun sequencing data was analyzed with the SqueezeMeta program v1.6.3, a fully automated pipeline for metagenomics that covered all steps of the analysis (Tamames and Puente-Sánchez, 2019). Adapter removal, quality filtering and trimming of the reads were done by Trimmomatic (Bolger et al., 2014). The high-quality reads were assembled into contigs using MEGAHIT v1.2.9 (Li et al., 2015) via the sequential mode. Prinseq was used to remove short contigs (<200 bp) and determine the contig statistics (Schmieder and Edwards, 2011). The contigs were subjected to gene prediction using Prodigal software v2.6.3, which was employed to retrieve the corresponding amino acid sequences (Hyatt et al., 2010). Diamond

v2.1.10 (Buchfink et al., 2021) was used to search for similarity between the NCBI nr database (Sayers et al., 2019). Taxonomic assignments of the genes were implemented using the lowest common ancestor (LCA) algorithm of the hits for each query gene searched against the reference database, NCBI nr. For functional assignment, Diamond was also used to compare gene sequences against the KEGG Orthology database (Kanehisa et al., 2017).

## Statistical analyses

All statistical analyses were conducted in R v4.3.3 (Kelley et al., 2008). A one-way analysis of variance (ANOVA) was used to evaluate the differences in soil physicochemical properties across the study sites and Tukey's pairwise comparison was used for the means at a significant level ( $p$ -value < 0.05). Spearman's correlation coefficient method was employed to assess the correlation between soil pH and other soil physicochemical properties. The diversity within the bacterial communities in the amplicon sequence data was depicted qualitatively using observed features, Shannon and phylogenetic distance (PD) matrices which were visualized as alpha rarefaction plots. Vegan package in R (Dixon, 2003) was used to evaluate the alpha diversity of the bacterial community in the shotgun metagenomic sequence data using the Shannon and Simpson indexes. The sequencing quality monitoring (SQM) tools package in R (Puente-Sánchez et al., 2020) was used to analyze both taxonomic and functional profiling data generated from the SqueezeMeta pipeline. Default parameters were used for all software unless otherwise specified.

## Results

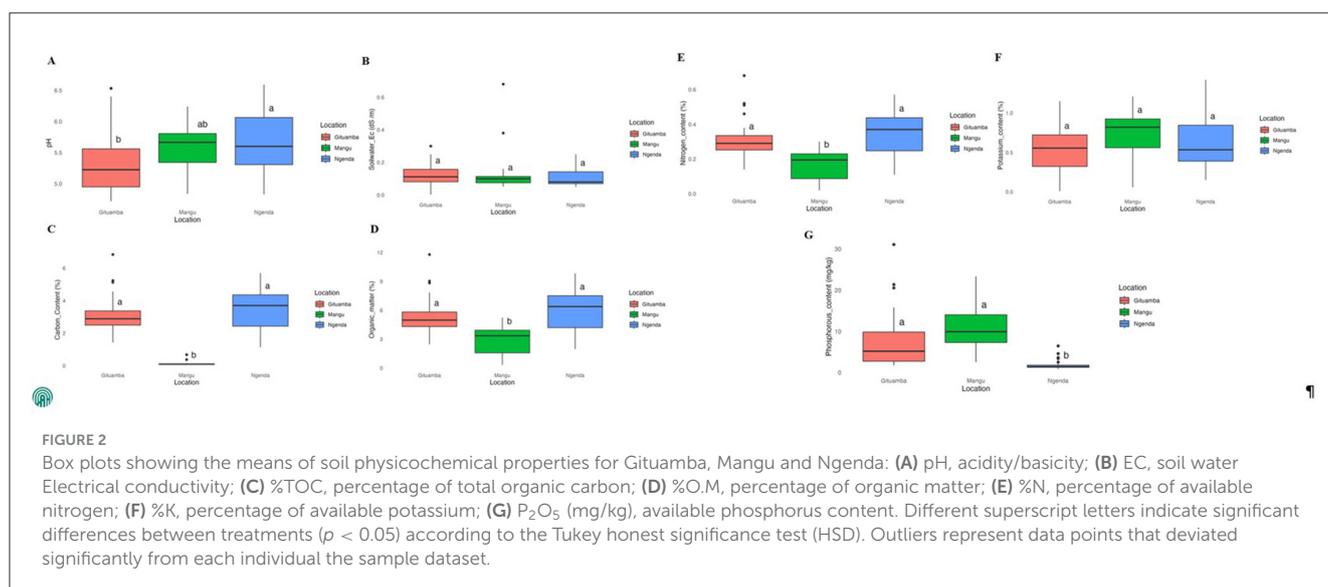
### Soil physicochemical properties analysis

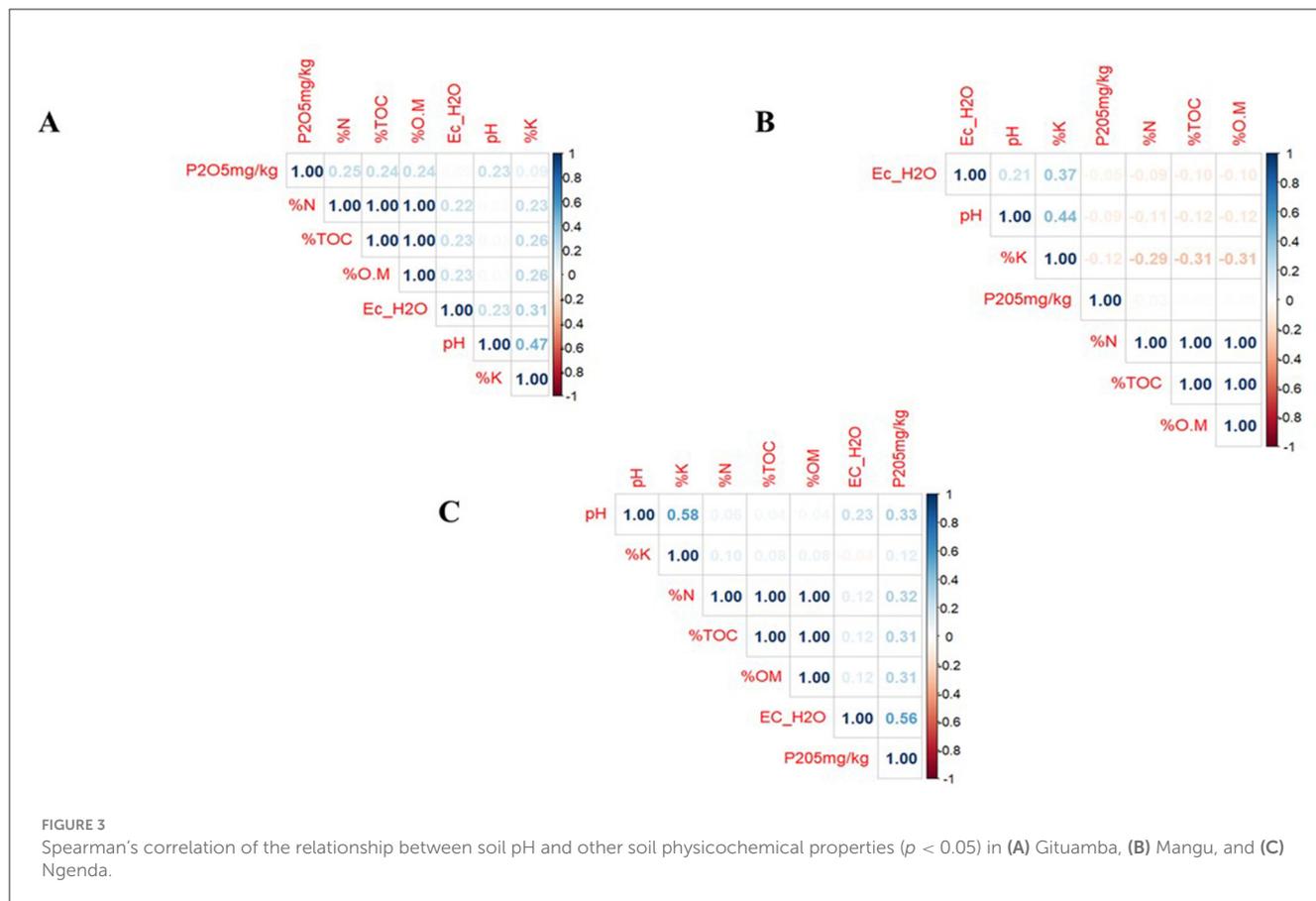
Results of the means of soil physicochemical properties of the composite samples revealed that the soil pH was slightly acidic across the sites. Soils in Gituamba had the lowest pH (5.32

$\pm 0.47$ ) while Ngenda recorded the highest pH (5.67  $\pm 0.51$ ; Supplementary material 1). There was no significant difference in pH between Mangu and other study sites (Figure 2A). Mangu soils contained a slightly higher EC value (0.13  $\pm 0.13$  dS/m) while Ngenda had the lowest value (0.11  $\pm 0.01$  dS/m). The differences in EC across the three sites were not significant (Figure 2B). The contents of TOC were highest in Ngenda soils (3.49  $\pm 1.30\%$ ), while Mangu soils exhibited the lowest content (1.77  $\pm 0.87\%$ ; Figure 2C). Similarly, Ngenda soils had the highest content of OM (6.02  $\pm 2.24\%$ ) while Mangu soils had a significantly lower content (3.05  $\pm 1.50\%$ ) as compared to Gituamba and Ngenda respectively (Figure 2D). The highest content of available N was contained in Ngenda soils (0.35  $\pm 0.13\%$ ) while Mangu had a significantly lower available N (0.18  $\pm 0.09\%$ ) compared to Gituamba and Ngenda respectively (Figure 2E). The levels of available K content were relatively similar and insignificant across all the sites with the highest content recorded in Mangu soils (0.64  $\pm 0.33\%$ ; Figure 2F). A higher amount of available P content was contained in Mangu (10.56  $\pm 5.69$  mg/kg) while the lowest content was recorded in Ngenda soils (1.89  $\pm 1.29$  mg/kg; Figure 2G).

### Spearman's correlation analysis

The relationship between soil pH and most of the soil parameters [EC, %TOC, %O.M, %N and P (mg/kg)] in Gituamba soils exhibited weak positive correlations indicating that changes in pH did not greatly affect other parameters. Available potassium levels showed a moderate correlation to soil pH suggesting that an increase in pH resulted in a noticeable increase in potassium content in Gituamba (Figure 3A). In Mangu soils, there was a weak negative correlation between pH and other soil parameters [%TOC, %O.M, %N, and P (mg/kg)] while a weak positive correlation was recorded between pH and EC. Like Gituamba soils, pH positively correlated with available potassium content (Figure 3B). The soil pH in Ngenda soils had weak positive correlations with other soil parameters (EC, %TOC, %O.M, and %N) as well as moderate





positive correlations with available potassium and phosphorus were observed (Figure 3C).

## Analysis of sequencing data

A total of 2,769,898 16S rRNA sequence reads were obtained from the two samples. Sample 1 (Gituamba) had 1,887,650 reads while sample 2 (Mangu) had the least number 882,248 reads. The percentage of reads that passed quality filtering for the two samples was 90.32 and 89.09%, respectively (Supplementary material 2). A total of 67,136,628 shotgun metagenomic sequence reads were obtained from sample 3 (Ngenda) of which 27,584,941 (41.09%) reads were mapped and used for taxonomic and functional annotation. After assembly, 2,504,874 contigs were identified. The longest contig was 125,240 bp, while the shortest contig was 128 bp. A total of 3,205,501 open reading frames (ORFs) were predicted from Prodigal of which 1,508,761 were annotated against the KEGG Orthology database (Supplementary material 3).

## Bacterial diversity analysis

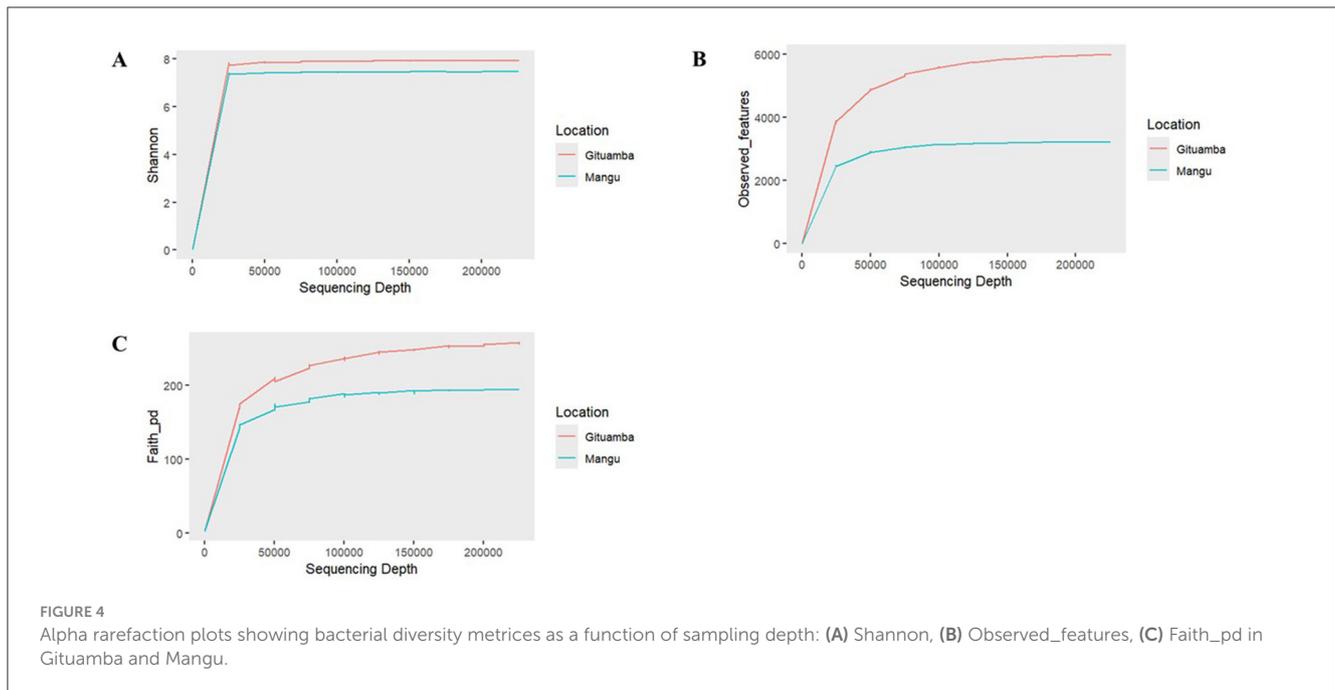
Based on Shannon diversity metrics, a higher bacterial diversity was recorded in Gituamba than Mangu (Figure 4A). The rarefaction plot of observed features revealed that Gituamba exhibited a higher number of observed features compared to

Mangu, indicating a richer species composition (Figure 4B). The rarefaction plot based on phylogenetic distance (PD) produced similar results to Shannon and observed features, indicating greater phylogenetic distances among bacterial communities in Gituamba than Mangu (Figure 4C). Due to the limited number of amplicon samples ( $n = 1$  per site), statistical comparison using  $t$ -tests was not appropriate. The lack of replication prevented reliable estimation of within-group variability. As a result, we focused on qualitative and descriptive analyses to highlight differences between sites. In the sample analyzed from Ngenda, the Shannon index (1.77) and Simpson index (0.49) (Supplementary material 4) indicated moderate microbial diversity. Nonetheless, given the lack of sample replication, these results are limited to the individual sample and should not be generalized to characterize the overall bacterial community at the site.

## Bacterial community composition

### Bacterial community composition based on 16S rRNA profiles

Classified representative amplicon sequence variants (ASVs) were assigned to 50 bacterial phyla, 117 classes, 247 orders, 508 families, 1,440 genera, and 2,748 species (Supplementary material 5). There were 11 dominant bacterial phyla in Gituamba and Mangu soils respectively. The most dominant phylum was *Proteobacteria* (59.7 and 58%) followed



by *Actinobacteria* (13.5 and 12.4%), and *Cyanobacteria* (8.2 and 11.9%; [Figure 5A](#)). At the genus level, 24 prominent genera were observed in both samples. The most dominant were *Pantoea* (43.7 and 30.8%) and *Brasilonema* (13.6 and 19.4%; [Figure 5B](#)).

### Bacterial community composition based on shotgun metagenomic sequencing profiles

Assembly-based metagenome taxonomic profiling revealed that Ngenda soils had four superkingdoms: Bacteria, Archaea, Eukaryota, and Viruses. Bacteria was the most abundant taxa containing 98% of the total soil taxa. The bacterial community was composed of 102 distinct phyla, 238 classes, 435 orders, 808 families, 1,772 genera, and 2,791 species ([Supplementary material 6](#)). At the phylum level, *Actinobacteria* (37%) was the most abundant followed by *Pseudomonadota* (34%), *Acidobacteria* (7%), and *Chloroflexota* (3%; [Figure 6A](#)). At the genus level, *Bradyrhizobium* (4.0%) was the most abundant followed by *Actinomycetes* (2.0%; [Figure 6B](#)).

## Metagenomic functional prediction

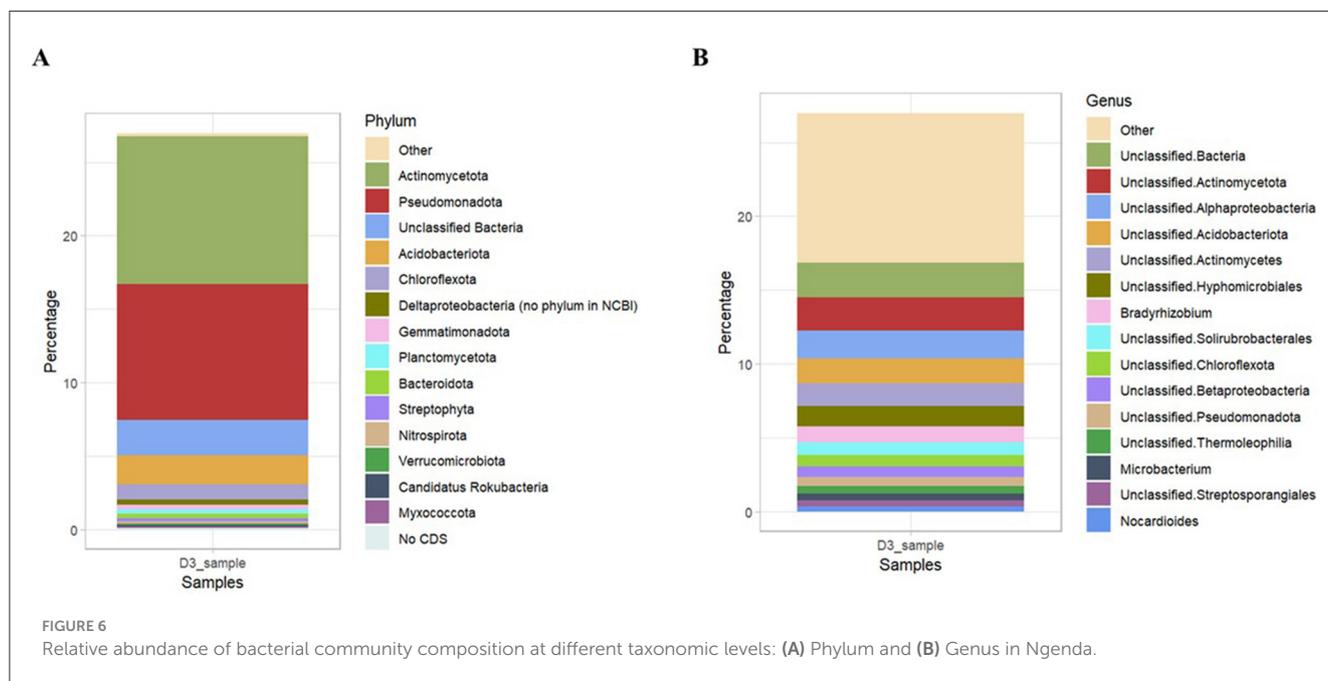
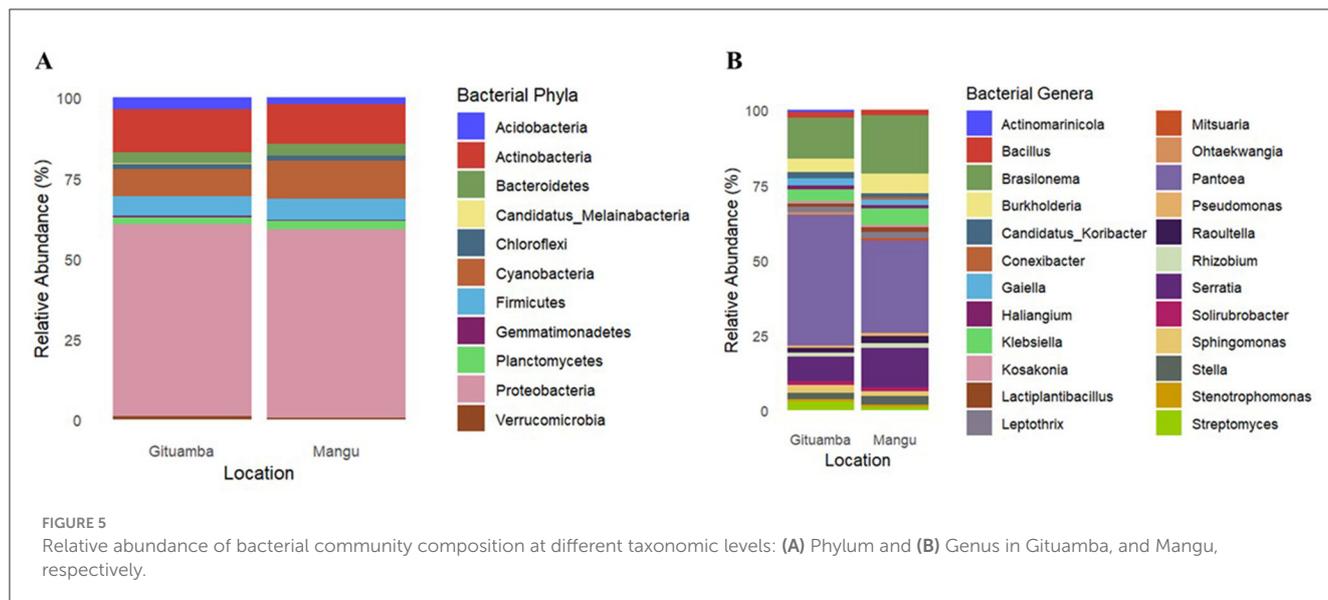
### Metagenomic functional prediction based on 16S rRNA profiles

A total of 7,634 KEGG orthologs (enzymes) predicted by PICRUSt2 were consolidated into 435 Meta-Cyc pathways. There were 15 most abundant functional pathways in both Gituamba and Mangu. PWY-3781 [Aerobic respiration I (cytochrome c)] was the most abundant followed by PWY-7111 (pyruvate fermentation to isobutanol), PWY-5101 (L-Isoleucine biosynthesis II), ILEUSYN-PWY (L-Isoleucine biosynthesis I from threonine) and VALSYN-PWY (L-Valine biosynthesis; [Figure 7](#)). Analysis of pathway differentiation by location demonstrated that the mean abundance

of predicted pathways was significantly elevated in Gituamba relative to Mangu.

### Metagenomic functional prediction based on shotgun metagenomic sequencing profiles

A total of 9,585 KEGG orthologs (KOs)/Enzymes were predicted among which 15 of them were the most abundant KOs [higher number of TPM (transcripts per million)] within the bacterial community metagenome. The most abundant ortholog in the metagenome was: RNA polymerase sigma-70 factor, ECF subfamily (KEGG ID: K03088). Other abundant KOs included: eukaryotic-like serine/threonine-protein-kinase [EC:2.7.11.1] (KEGG ID: K12132) followed by 3-oxoacyl- [acyl-carrier protein] reductase [EC:1.1.1.100] (KEGG ID: K00059), putative ABC transport system permease protein (KEGG ID: K02004), ABC- 2 type transport system ATP- binding protein (KEGG ID:K01990), acyl-COA dehydrogenase [EC:1.3.8.7] (KEGG ID: K00249) and aerobic carbon-monoxide dehydrogenase large subunit [EC:1.2.5.3] (KEGG ID: K03520; [Figure 8](#)). The KOs predicted within the metagenome were associated with six high-level functions: metabolism, genetic information processing, cellular processes, environmental information processing, organismal systems and human diseases ([Supplementary material 7](#)). The key lower-level functions included amino acid, lipid, sulfur, energy, nucleotide, vitamins and carbohydrate metabolism. For genetic information processing, the prominent low-level functions were replication, translation, transcription, folding, sorting, and degradation. Cellular processes within the metagenome included cell growth and death, cell motility, cellular community and transport and catabolism while the key functions predicted under environmental information processing were signal transduction and membrane transport. The most abundant bacterial genera with open reading frames (ORFs) containing



annotation of the six high-level functions were *Bradyrhizobium*, *Alphaproteobacteria*, *Betaproteobacteria*, *Pseudomonadota*, *Solirubacterales*, *Acidobacteria*, and *Hyphomicrobiales* (Figure 9).

## Discussion

Rhizosphere-inhabiting microbes can enhance plant growth and health (Kong and Liu, 2022). Their diversity in soil habitats is a potential indicator of soil health and quality (Bargali, 2024). However, changes in the diversity and functional potential of soil microbes within banana farming systems are still not well-understood. We used Illumina Novaseq, a high-throughput sequencing approach, to examine the bacterial communities in the rhizosphere of banana plants grown in small-scale farms in

Kiambu, focusing on their composition and functional potential. Also, the variation of the soil physicochemical parameters within our study sites was assessed.

Physicochemical analysis revealed that soil pH was slightly acidic across all study sites, with non-significant differences ( $p$ -value > 0.05) between Mangu and the other locations. The observed pH values in Mangu and Ngenda fall within the optimal range of 5.5–7.5 for banana cultivation, as reported by Nyamamba et al. (2020). Soil pH plays a crucial role in regulating chemical reactions and nutrient availability in the soil, directly impacting plant growth (Cordero et al., 2020). In addition, it is the primary factor that determines the microbial community structure in natural environmental systems (Kang et al., 2013). Our results align with the research conducted by Swafo and Dlamini (2022), who observed slightly acidic soils for banana production upon

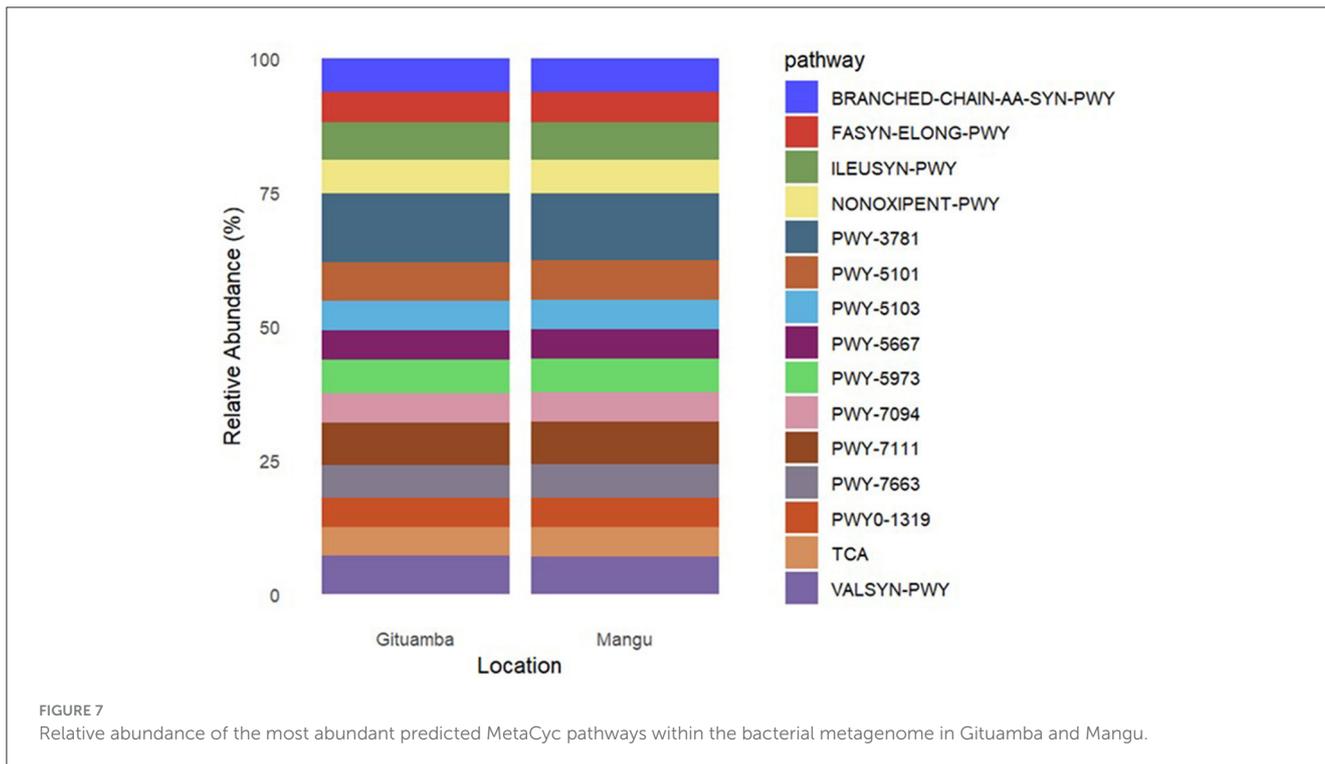


FIGURE 7 Relative abundance of the most abundant predicted MetaCyc pathways within the bacterial metagenome in Gituamba and Mangu.

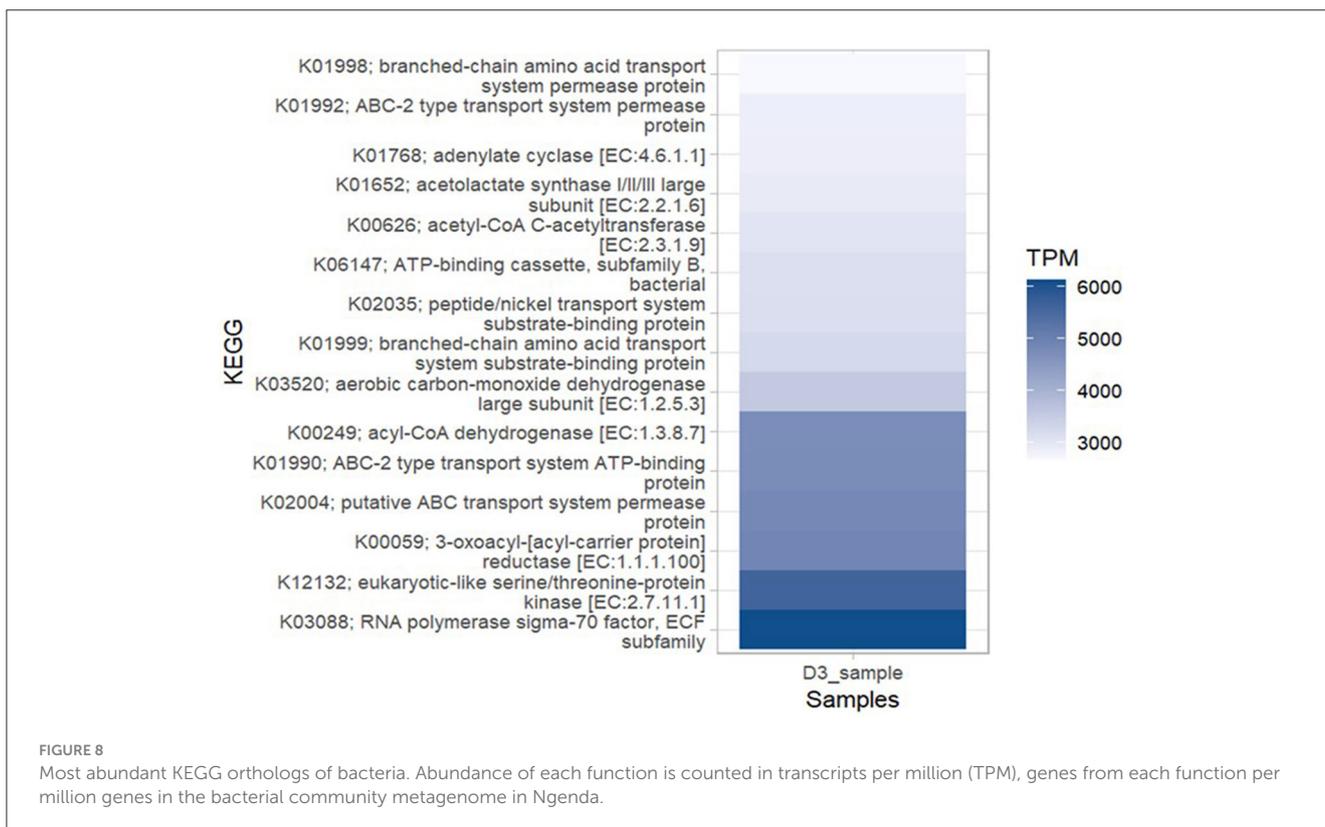
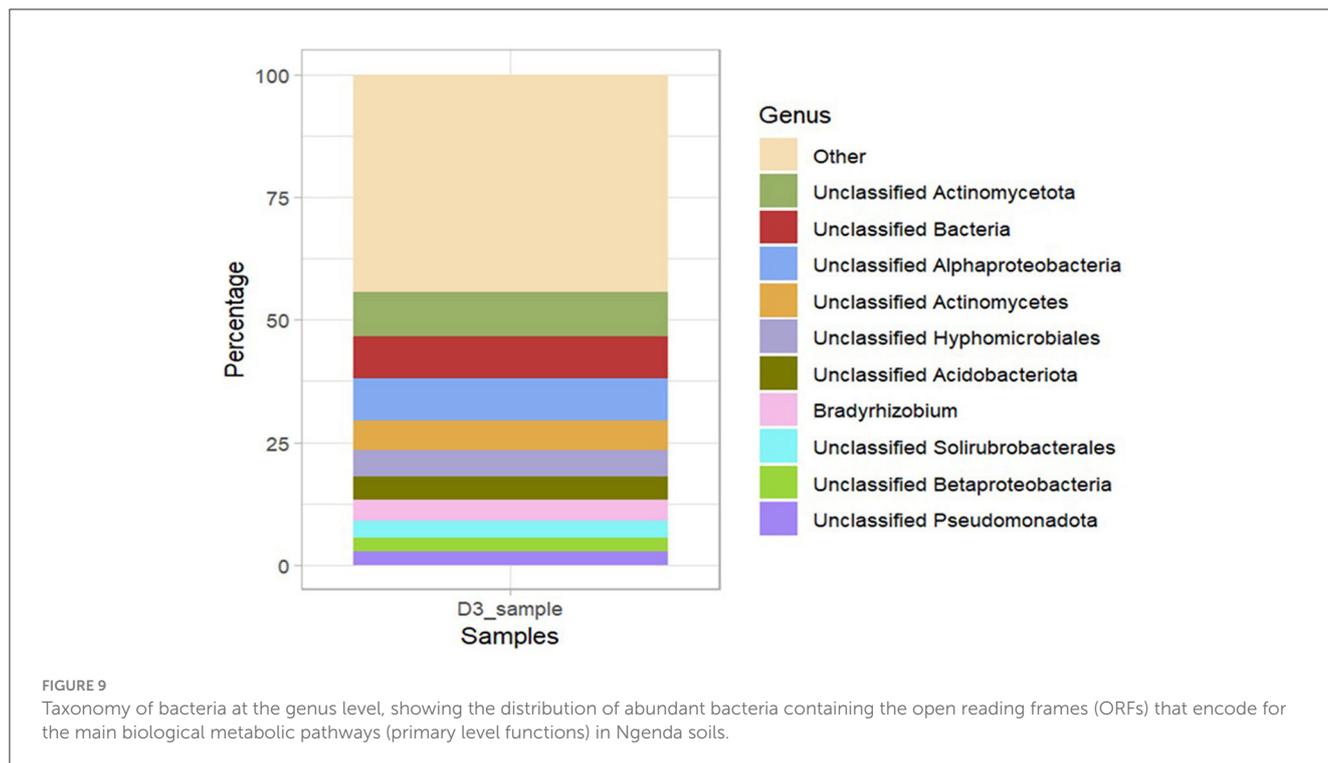


FIGURE 8 Most abundant KEGG orthologs of bacteria. Abundance of each function is counted in transcripts per million (TPM), genes from each function per million genes in the bacterial community metagenome in Ngenda.

characterizing the soil parameters. The variations in soil nutrient properties (TOC, N, and OM) showed a similar trend across our study sites. Previous reports suggest that the higher levels of soil nutrients in Gituamba and Ngenda can be attributed to high biomass production by the banana plants and other

crops, mulching practices, effective crop residue management and favorable soil microbial activities (Gerke, 2022; Githongo et al., 2022; Mushtaq et al., 2023). Spearman’s correlation revealed that soil pH had weak correlations with most of the parameters (EC, TOC, OM, N, and P) in our study sites suggesting that soil pH



did not strongly affect most soil parameters. Contrary to our findings, a related study by Zhang et al. (2019) revealed that soil pH showed positive correlations with other soil parameters in different soil horizons, making it a key factor in controlling soil nutrient availability. The findings of our study suggest that soil pH management practices should be prioritized across our study sites to enhance sustainable banana production. Overall, soils in Gituamba and Ngenda depicted better soil fertility characteristics with higher TOC, OM, and N levels compared to Mangu. The variations in soil parameters across the study sites could be due to variations in soil management practices, farming techniques, soil types, climate, and soil microbial communities (Singh et al., 2020).

Since rhizospheric soil is closest to the roots, it provides an avenue to explore microbial communities (Sui et al., 2019). From our metagenomics analyses of rhizospheric bacterial communities and their potential functions, we hypothesized that diverse microbial populations in banana soils play beneficial roles in supporting plant health. Metagenomic profiling, as demonstrated in this study, facilitates the identification of previously unknown bacterial taxa and their functional roles. These bacteria are essential for boosting plant growth, strengthening disease resistance, and enhancing soil health. Additionally, we anticipate that these microbial communities will display distinct biogeographical patterns across different banana-growing regions, influenced by variations in soil characteristics, climate, and agricultural practices. The makeup of microbial communities in rhizospheric soil is strongly linked to soil quality and the overall health of crops (Shi et al., 2021). Soil-associated microbiota significantly influences how plants adapt and respond to environmental changes, including stressors like salinity and drought (Zhang et al., 2017).

In Gituamba and Mangu agroecosystems, our primary goal was to characterize the bacterial composition and diversity. This guided the choice of amplicon sequencing, particularly 16S rRNA gene sequencing, which provided a cost-effective and high-throughput method to profile the microbial communities at the taxonomic level. On the other hand, the Ngenda agroecosystems soil was subjected to shotgun metagenomic sequencing. Due to relatively better soil fertility characteristics in Ngenda compared to other sites, we aimed to go beyond taxonomic classification to analyze the functional traits and metabolic potential of the microbial communities. Shotgun metagenomics thus allowed whole genome reconstruction and functional annotation, providing detailed insights into the potential roles of soil bacteria in ecosystem attributes. By using both amplicon and shotgun sequencing approaches, we ensured a balanced approach in characterizing the taxonomic composition in all sites while obtaining more functional insights in Ngenda than Gituamba and Mangu.

Phylum *Proteobacteria* was the most prevalent across all our sites and has been consistently observed in studies evaluating bacterial communities in the banana rhizosphere (Fu et al., 2017; Kaushal et al., 2022a,b; Mia et al., 2010). Members of the phylum *Proteobacteria*, a Gram-negative bacterium, are also enriched in the rhizosphere soil of many plants (Fu et al., 2022; Gao et al., 2019; Zhang et al., 2022). They play an active role in organic matter decomposition, ammonia oxidation, nitrogen fixation and degradation of inorganic compounds (Ling et al., 2022). The phylum *Actinobacteria*, a group of Gram-positive bacteria, plays a crucial role in decomposing organic matter and produces antibiotics that inhibit plant pathogens in soil (Dhakal et al., 2017). In addition, *Cyanobacteria*, *Planctomycetes*, and *Chloroflexi* are known to play a role in nitrogen cycling within agroecosystems

(Li et al., 2021). A previous study by Beltran-Garcia et al. (2021) also identified key processes mediated by *Proteobacteria*, *Actinobacteria*, and *Cyanobacteria* within banana agroecosystems.

The predominant genus, *Pantoea*, a Gram-negative bacterium belongs to the *Enterobacteriaceae* family, and its members are versatile biocontrol agents, protecting against pathogenic microbes in several plant varieties (Duchateau et al., 2024). Other abundant genera including *Brasilonema*, a genus of *Cyanobacteria*, play a significant role in banana soils, particularly in mitigating Panama disease, a soil-borne infection, caused by pathogenic fungi, *Fusarium*. Other genera including *Serratia*, *Klebsiella*, *Burkholderia*, and *Streptomyces* contribute to plant growth promotion by production of phytohormones, mitigation of plant abiotic stresses and inhibition of plant pathogens (Dos Santos et al., 2022; Kulkova et al., 2024; Olanrewaju and Babalola, 2019). The bacterial communities in Ngenda soils exhibited a moderate level of alpha diversity, suggesting that the bacterial community diversity was moderate with a balanced representation of different species. This suggests a relatively healthy and stable ecosystem. Phyla *Pseudomonadota* and *Actinobacteria* were the most prominent in Ngenda soils, consistent with findings by Liu et al. (2022) who investigated the vertical distribution patterns of soil microbial communities in peatlands under varying environmental conditions. Members of *Actinobacteria* bacteria are an important phylum that act as plant growth promoters due to their vital role in the cycling of organic matter and inhibition of several plant pathogens (Ibrahimi et al., 2023). At the genus level, we found *Bradyrhizobium* as the most abundant in Ngenda soils and its members actively promote plant growth and yield by fixing nitrogen in soil (Shahrajabian et al., 2021). A study by Chalasani et al. (2021) also reported the positive roles of *Bradyrhizobium* enhancing legume nodulation, which helps to maintain plant community structure and restore degraded ecosystems. A previous study by Wahome et al. (2023) explored the abundance of soil bacteria as well the effects of soil parameters on their distribution in small-holder banana farms in Kisii, Nyamira, and Embu Counties of Kenya. Similar to our findings, they reported *Proteobacteria* phylum as the most prevalent. They also noted a variation in soil pH which affects the bacterial community composition in all their study sites. Overall, their study revealed the importance of soil health and management practices in shaping microbial diversity, which in turn affects the productivity of banana agroecosystems in these regions. In our study, the adaptability of soil bacteria likely accounts for their significant presence in the study sites and underscores their potential role in maintaining soil health and enhancing crop yield. These sensitive bacterial species may be crucial for evaluating the impact of human activities on banana microbiomes and their associated ecological functions.

The functional diversity of soil microbes is a vital indicator of community composition and ecological roles, essential for understanding the contributions of microorganisms in different environments (Chen et al., 2020). We focused on bacterial functional predictions of pathways and enzymes encoded within bacterial metagenomes. Metabolic gene functions in the 16S rRNA gene sequences revealed a significantly higher abundance in Gitumba than in Mangu. Metabolic pathways such as aerobic respiration I (PWY-3781), pyruvate fermentation to isobutanol

(PWY-711), L-isoleucine biosynthesis II (PWY-5101), L-isoleucine biosynthesis I from threonine (ILEUSYN-PWY), and L-valine biosynthesis (VALSYN-PWY) were prominent. These findings resonate with studies by Tang et al. (2024) and Zhou et al. (2023) who reported that most bacterial metabolic pathways in soils were associated with carbon IV oxide assimilation, acetate, carbohydrates, amino acids biosynthesis, and inorganic compounds degradation. In our study, most of the predicted pathways are linked to crucial processes such as nitrate reduction and the breakdown of various anaerobic aromatic compounds, including catechol, creatinine, toluene, salicylate, L-arabinose, L-valine, and the biosynthesis of adenosine nucleotides, 1,5-anhydrofructose, and pyrimidine ribonucleotides (Zhou et al., 2023). Therefore, the predicted pathways, based on PICRUSt2 analysis, are associated with functions that may contribute to plant growth, development, and the synthesis of bioactive substances. However, it is important to note that these predictions are based on 16S rRNA gene profiles and do not represent direct functional measurements. Previous studies have linked similar microbial functions, such as energy metabolism, fermentation/alternative energy metabolism and amino acid biosynthesis pathways, to enhanced plant performance (Adedayo et al., 2022). Recent reports indicate that the most common bacterial phyla that contain metabolic pathways in rhizospheric soils that contribute to soil health and plant growth include *Proteobacteria*, *Actinobacteria*, *Cyanobacteria*, *Acidobacteria*, *Bacteroidetes*, *Firmicutes*, *Chloroflexi*, and *Planctomycetes* (Li et al., 2014, 2020; Lopes et al., 2016; Nicolas et al., 2021).

The most abundant KEGG Orthologs/Enzymes within the bacterial metagenome in Ngenda were RNA polymerase sigma-70 factor, Extracytoplasmic function (ECF) subfamily, eukaryotic-like serine/threonine-protein-kinase, 3-oxoacyl- [acyl-carrier protein] reductase, putative ABC transport system permease protein, ATP-binding cassette (ABC)-2 type transport system ATP-binding protein, acyl-CoA dehydrogenase and aerobic carbon-monoxide dehydrogenase large subunit. These results align with a similar investigation by Wang et al. (2023) that explored the microbial communities of rhizosphere soil and root endophytes in *Panax notoginseng* from various geographical regions. The KEGG pathways in this study site are associated with key processes that contribute to improved soil health and crop production in agroecosystems as demonstrated by Abulfaraj et al. (2024). By connecting taxonomy and function, we discovered that genes possessed by *Bradyrhizobium*, *Actinomycetota*, *Alphaproteobacteria*, *Betaproteobacteria*, *Pseudomonadota*, *Solirubacterales*, *Acidobacteria*, and *Hyphomicrobiales* were associated with metabolic pathways for metabolism, genetic information processing, cellular processes, environmental information processing, organismal systems and human disease processes. Analyzing the functional microbiome revealed potential bacteria that could benefit plants and promote their growth at our study locations.

While functional prediction provided initial insights into the contribution of microbial communities to gene functions in banana rhizospheric soils, it is recommended to apply several omics approaches such as metagenomics, metatranscriptomics, metabolomics and proteomics in future studies. These approaches

can help address the known biases associated with functional prediction tools. Also, the selection between shotgun and amplicon sequencing methods for microbiome analysis typically depends on the specific research objectives. For example, amplicon sequencing is ideal for analyzing larger samples, such as in longitudinal studies, but it provides limited taxonomic and functional resolution (Jovel et al., 2016). In contrast, shotgun metagenomics generally offers more comprehensive information on numerous genes, resulting in enhanced taxonomic and functional resolution of sequences (Peterson et al., 2021; Zaheer et al., 2018). In the current study, both approaches revealed almost similar taxonomic groups whereas functional profiling yielded more information in the shotgun approach compared to 16S rRNA gene sequencing. The findings from our study describe the need for the application of omics techniques to provide more research insights in banana-soil microbiome interactions, for sustainable agriculture in different growing regions.

## Conclusion

This study provides baseline data on the bacterial composition and predicted gene functions associated with the rhizospheres of banana plants in Kenya. It also highlights differences in soil physicochemical properties across agroecosystems, emphasizing the need for future research to explore how these factors influence microbial community structure. Through amplicon and shotgun metagenomic sequencing, we demonstrated that the abundance of beneficial rhizosphere bacteria across diverse environments is linked to their ecological functions, contributing to soil health and crop productivity. A comprehensive understanding of key microbial metabolic pathways that facilitate plant development could empower farmers to adopt sustainable soil management practices, enhancing both environmental and economic outcomes.

However, we acknowledge certain limitations in this study, including the use of only two composite amplicon samples per site without biological replication and the application of shotgun metagenomics to a single sample. These constraints limit the generalizability and robustness of our diversity and functional predictions. Future research should incorporate multiple biological replicates and broader shotgun metagenomic analyses to strengthen statistical validity and provide a more comprehensive understanding of microbial roles in the banana rhizosphere. Expanding such efforts could reveal novel microbial taxa and functions with potential applications in ecosystem management, agriculture, and biotechnology.

## Data availability statement

Sequences from this study were deposited into the SRA under the BioProject ID PRJNA1146306 (<https://www.ncbi.nlm.nih.gov/search/all/?term=PRJNA1146306>).

## Author contributions

EM: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original

draft. GK: Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. OM: Investigation, Methodology, Visualization, Writing – original draft. AK: Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing. VG: Supervision, Validation, Writing – review & editing. WM: Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research received support from the World Bank Group through the National Agricultural and Rural Inclusive Growth Project (NARIGP) in Kiambu County and the National Research Fund, Kenya (NRF/2/MMC/158).

## Acknowledgments

We thank Pwani University for providing laboratory space at the Pwani University Bioscience Research Centre (PUBReC) to conduct this research. Additionally, the authors acknowledge the National Research Fund for their research grant support.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2025.1582271/full#supplementary-material>

## References

- Abulfaraj, A. A., Shami, A. Y., Alotaibi, N. M., Alomran, M. M., Aloufi, A. S., Al-Andal, A., et al. (2024). Exploration of genes encoding KEGG pathway enzymes in rhizospheric microbiome of the wild plant *Abutilon fruticosum*. *AMB Express* 14:27. doi: 10.1186/s13568-024-01678-4
- Adedayo, A., Fadiji, A., and Babalola, O. (2022). The effects of plant health status on the community structure and metabolic pathways of rhizosphere microbial communities associated with *Solanum lycopersicum*. *Horticulturae* 8:404. doi: 10.3390/horticulturae8050404
- Alawiye, T., and Babalola, O. (2019). Bacterial diversity and community structure in typical plant rhizosphere. *Diversity* 11:179. doi: 10.3390/d11100179
- Alawiye, T., and Babalola, O. (2021). Metagenomic insight into the community structure and functional genes in the sunflower rhizosphere microbiome. *Agriculture* 11:167. doi: 10.3390/agriculture11020167
- Alemu, M. M. (2017). Banana as a cash crop and its food security and socioeconomic contribution: the case of Southern Ethiopia, Arba Minch. *J. Environ. Prot.* 08, 319–329. doi: 10.4236/jep.2017.83024
- Barbera, P., Kozlov, A. M., Czech, L., Morel, B., Darriba, D., Flouri, T., et al. (2018). Data from: EPA-ng: massively parallel evolutionary placement of genetic sequences (Version 1, p. 16601287 bytes) [Dataset]. Dryad. doi: 10.1101/291658
- Bargali, S. S. (2024). Soil microbial biomass: a crucial indicator of soil health. *Curr. Agric. Res. J.* 12, 01–06. doi: 10.12944/CARJ.12.1.01
- Beltran-Garcia, M. J., Martinez-Rodriguez, A., Olmos-Arriaga, I., Valdez-Salas, B., Chavez-Castrillon, Y. Y., Di Mascio, P., et al. (2021). Probiotic endophytes for more sustainable banana production. *Microorganisms* 9:1805. doi: 10.3390/microorganisms9091805
- Birt, H. W. G., Pattison, A. B., Skarszewski, A., Daniells, J., Raghavendra, A., and Dennis, P. G. (2022). The core bacterial microbiome of banana (*Musa* spp.). *Environ. Microbiome* 17:46. doi: 10.1186/s40793-022-00442-0
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9
- Buchfink, B., Reuter, K., and Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* 18, 366–368. doi: 10.1038/s41592-021-01101-x
- Chalasan, D., Basu, A., Pullabhotla, S. V., Jorin, B., Neal, A. L., Poole, P. S., et al. (2021). Poor competitiveness of *Bradyrhizobium* in pigeon pea root colonization in Indian soils. *MBio* 12, 10–1128. doi: 10.1128/mBio.00423-21
- Chen, S., Huang, T., Zhou, Y., Han, Y., Xu, M., and Gu, J. (2017). AfterQC: automatic filtering, trimming, error removing and quality control for fastq data. *BMC Bioinformatics* 18(S3):80. doi: 10.1186/s12859-017-1469-3
- Chen, Z.-J., Shao, Y., Li, Y.-J., Lin, L.-A., Chen, Y., Tian, W., et al. (2020). Rhizosphere bacterial community structure and predicted functional analysis in the water-level fluctuation zone of the danjiangkou reservoir in china during the dry period. *Int. J. Environ. Res. Public Health* 17:1266. doi: 10.3390/ijerph17041266
- Cordero, J., De Freitas, J. R., and Germida, J. J. (2020). Bacterial microbiome associated with the rhizosphere and root interior of crops in Saskatchewan, Canada. *Can. J. Microbiol.* 66, 71–85. doi: 10.1139/cjm-2019-0330
- Czech, L., Barbera, P., and Stamatidis, A. (2020). Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics* 36, 3263–3265. doi: 10.1093/bioinformatics/btaa070
- Dhakal, D., Pokhrel, A. R., Shrestha, B., and Sohng, J. K. (2017). Marine rare actinobacteria: isolation, characterization, and strategies for harnessing bioactive compounds. *Front. Microbiol.* 8:1106. doi: 10.3389/fmicb.2017.01106
- Dixit, M., Ghoshal, D., Lal Meena, A., Ghasal, P. C., Rai, A. K., Choudhary, J., et al. (2024). Changes in soil microbial diversity under present land degradation scenario. *Total Environ. Adv.* 10:200104. doi: 10.1016/j.teadv.2024.200104
- Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* 14, 927–930. doi: 10.1111/j.1654-1103.2003.tb02228.x
- Dos Santos, I. B., Pereira, A. P. D. A., De Souza, A. J., Cardoso, E. J. B. N., Da Silva, F. G., Oliveira, J. T. C., et al. (2022). Selection and characterization of burkholderia spp. for their plant-growth promoting effects and influence on maize seed germination. *Front. Soil Sci.* 1:805094. doi: 10.3389/fsoil.2021.805094
- Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2020). PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 38, 685–688. doi: 10.1038/s41587-020-0548-6
- Duchateau, S., Crouzet, J., Dorey, S., and Aziz, A. (2024). The plant-associated *Pantoea* spp. as biocontrol agents: mechanisms and diversity of bacteria-produced metabolites as a prospective tool for plant protection. *Biol. Control* 188:105441. doi: 10.1016/j.biocontrol.2024.105441
- Enagbonma, B. J., Ajillogba, C. F., and Babalola, O. O. (2020). Metagenomic profiling of bacterial diversity and community structure in termite mounds and surrounding soils. *Arch. Microbiol.* 202, 2697–2709. doi: 10.1007/s00203-020-01994-w
- Fu, L., Penton, C. R., Ruan, Y., Shen, Z., Xue, C., Li, R., et al. (2017). Inducing the rhizosphere microbiome by biofertilizer application to suppress banana *Fusarium wilt* disease. *Soil Biol. Biochem.* 104, 39–48. doi: 10.1016/j.soilbio.2016.10.008
- Fu, L., Yan, Y., Li, X., Liu, Y., and Lu, X. (2022). Rhizosphere soil microbial community and its response to different utilization patterns in the semi-arid alpine grassland of northern Tibet. *Front. Microbiol.* 13:931795. doi: 10.3389/fmicb.2022.931795
- Gao, X., Wu, Z., Liu, R., Wu, J., Zeng, Q., and Qi, Y. (2019). Rhizosphere bacterial community characteristics over different years of sugarcane ratooning in consecutive monoculture. *Biomed Res. Int.* 2019, 1–10. doi: 10.1155/2019/6916189
- Gerke, J. (2022). The central role of soil organic matter in soil fertility and carbon storage. *Soil Syst.* 6:33. doi: 10.3390/soilsystems6020033
- Giri, B., and Varma, A. (Eds.). (2020). *Soil Health, Vol. 59*. New York, NY: Springer International Publishing. doi: 10.1007/978-3-030-44364-1
- Githongo, M., Kiboi, M., Muriuki, A., Fliessbach, A., Musafiri, C., and Ngetich, F. K. (2022). Organic carbon content in fractions of soils managed for soil fertility improvement in sub-humid agroecosystems of Kenya. *Sustainability* 15:683. doi: 10.3390/su15010683
- Han, C., Zhang, Z., Gao, Y., Wang, W., Chen, J., and Wang, Y. (2023). Microbiome reveals the effects of biogas fertilizer on soil microbial community structure and diversity in perennial apple orchards. *Horticulturae* 9:1023. doi: 10.3390/horticulturae9091023
- He, Z., Yuan, C., Chen, P., Rong, Z., Peng, T., Farooq, T. H., et al. (2023). Soil microbial community composition and diversity analysis under different land use patterns in Taojia River Basin. *Forests* 14:1004. doi: 10.3390/f14051004
- Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119
- Ibrahimi, M., Loqman, S., Jemo, M., Hafidi, M., Lemee, L., and Ouhdouch, Y. (2023). The potential of facultative predatory *Actinomycetota* spp. and prospects in agricultural sustainability. *Front. Microbiol.* 13:1081815. doi: 10.3389/fmicb.2022.1081815
- Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., et al. (2016). Characterization of the gut microbiome using 16S or shotgun metagenomics. *Front. Microbiol.* 7:459. doi: 10.3389/fmicb.2016.00459
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361. doi: 10.1093/nar/gkw1092
- Kang, S., Van Nostrand, J. D., Gough, H. L., He, Z., Hazen, T. C., Stahl, D. A., et al. (2013). Functional gene array-based analysis of microbial communities in heavy metals-contaminated lake sediments. *FEMS Microbiol. Ecol.* 86, 200–214. doi: 10.1111/1574-6941.12152
- Kaushal, M., Kolombia, Y., Alakonya, A. E., Kuate, A. F., Ortega-Beltran, A., Amah, D., et al. (2022a). Subterranean microbiome affiliations of plantain (*Musa* spp.) under diverse agroecologies of Western and Central Africa. *Microb. Ecol.* 84, 580–593. doi: 10.1007/s00248-021-01873-x
- Kaushal, M., Mahuku, G., and Swennen, R. (2020a). Metagenomic insights of the root colonizing microbiome associated with symptomatic and non-symptomatic Bananas in fusarium wilt infected fields. *Plants* 9:263. doi: 10.3390/plants9020263
- Kaushal, M., Swennen, R., and Mahuku, G. (2020b). Unlocking the microbiome communities of banana (*Musa* spp.) under disease stressed (*Fusarium wilt*) and non-stressed conditions. *Microorganisms* 8:443. doi: 10.3390/microorganisms8030443
- Kaushal, M., Tumuhairwe, J. B., Kaingo, J., Richard, M., Nakamanya, F., Taulya, G., et al. (2022b). Compositional shifts in microbial diversity under traditional banana cropping systems of Sub-Saharan Africa. *Biology* 11:756. doi: 10.3390/biology11050756
- Kelley, K., Lai, K., and Wu, P.-J. (2008). "Using R for data analysis: a best practice for research," in *Best Practices in Quantitative Methods*, ed. J. Osborne (London: SAGE Publications, Inc), 535–572. doi: 10.4135/9781412995627.d40
- Kong, Z., and Liu, H. (2022). Modification of rhizosphere microbial communities: a possible mechanism of plant growth promoting rhizobacteria enhancing plant growth and fitness. *Front. Plant Sci.* 13:920813. doi: 10.3389/fpls.2022.920813
- Kulkova, I., Wróbel, B., and Dobrzyński, J. (2024). *Serratia* spp. as plant growth-promoting bacteria alleviating salinity, drought, and nutrient imbalance stresses. *Front. Microbiol.* 15:1342331. doi: 10.3389/fmicb.2024.1342331
- Kutwa, A. A., Muhingi, W. N., and Kokonya, D. (2016). Smallholder Rural Youth Farming in Kiambu County, Kenya. *J. Cult. Soc. Devel.* 25, 60–71

- Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. *Bioinformatics* 31, 1674–1676. doi: 10.1093/bioinformatics/btv033
- Li, J., Wen, Y., and Yang, X. (2021). Understanding the responses of soil bacterial communities to long-term fertilization regimes using DNA and RNA sequencing. *Agronomy* 11:2425. doi: 10.3390/agronomy11122425
- Li, M., Zhang, K., Yan, Z., Liu, L., Kang, E., and Kang, X. (2022). Soil water content shapes microbial community along gradients of wetland degradation on the Tibetan Plateau. *Front. Microbiol.* 13:824267. doi: 10.3389/fmicb.2022.824267
- Li, X., Rui, J., Xiong, J., Li, J., He, Z., Zhou, J., et al. (2014). Functional potential of soil microbial communities in the maize rhizosphere. *PLoS ONE* 9:e112609. doi: 10.1371/journal.pone.0112609
- Li, Y., Wang, C., Wang, T., Liu, Y., Jia, S., Gao, Y., et al. (2020). Effects of different fertilizer treatments on rhizosphere soil microbiome composition and functions. *Land* 9:329. doi: 10.3390/land9090329
- Ling, N., Wang, T., and Kuzuyakov, Y. (2022). Rhizosphere bacteriome structure and functions. *Nat. Commun.* 13:836. doi: 10.1038/s41467-022-28448-9
- Liu, L., Wang, Z., Ma, D., Zhang, M., and Fu, L. (2022). Diversity and distribution characteristics of soil microbes across forest-peatland ecotones in the permafrost regions. *Int. J. Environ. Res. Public Health* 19:14782. doi: 10.3390/ijerph192214782
- Lopes, L. D., Pereira E Silva, M. D. C., and Andreote, F. D. (2016). Bacterial abilities and adaptation to the rhizosphere colonization. *Front. Microbiol.* 7:1341. doi: 10.3389/fmicb.2016.01341
- Mia, M. B., Shamsuddin, Z. H., Wahab, Z., and Marziah, M. (2010). Rhizobacteria as bioenhancer and biofertilizer for growth and yield of banana (*Musa* spp. cv. 'Berangan'). *Sci. Hortic.* 126, 80–87. doi: 10.1016/j.scienta.2010.06.005
- Molano, L.-A. G., Vega-Abellaneda, S., and Manichanh, C. (2024). GSR-DB: A manually curated and optimized taxonomical database for 16S rRNA amplicon analysis. *mSystems* 9:e00950-23. doi: 10.1128/mSystems.00950-23
- Molefe, R. R., Amoo, A. E., and Babalola, O. O. (2021). Metagenomic insights into the bacterial community structure and functional potentials in the rhizosphere soil of maize plants. *J. Plant Interact.* 16, 258–269. doi: 10.1080/17429145.2021.1936228
- Mushtaq, H., Ganai, B. A., and Jehangir, A. (2023). Exploring soil bacterial diversity in different micro-vegetational habitats of Dachigam National Park in North-western Himalaya. *Sci. Rep.* 13:3090. doi: 10.1038/s41598-023-30187-w
- Muthee, A. I., Gichimu, B. M., and Nthakania, P. N. (2019). Analysis of banana production practices and constraints in Embu County, Kenya. *Asian J. Agric. Rural Dev.* 9, 123–132. doi: 10.18488/journal.1005/2019.9.1/1005.1.123.132
- Nicolas, A. M., Jaffe, A. L., Nuccio, E. E., Taga, M. E., Firestone, M. K., and Banfield, J. F. (2021). Soil candidate phyla radiation bacteria encode components of aerobic metabolism and co-occur with nanoarchaea in the rare biosphere of rhizosphere grassland communities. *mSystems* 6:e01205-20. doi: 10.1128/mSystems.01205-20
- Njiru, E. B. K. (2019). Urban expansion and loss of agricultural land: A GIS based study of Kiambu County. *Int. J. Sci. Res.* 8, 915–919.
- Nwachukwu, B. C., and Babalola, O. O. (2022). Metagenomics: a tool for exploring key microbiome with the potentials for improving sustainable agriculture. *Front. Sustain. Food Syst.* 6:886987. doi: 10.3389/fsufs.2022.886987
- Nyamamba, K. A., Ouna, T. O., Kamiri, H., and Pane, E. (2020). Effects of land use change on banana production: a case study of Imenti South Sub-County of Meru County in Kenya. *BloEx J.2*, 640–652. doi: 10.33258/bloex.v2i3.303
- Nzioka, S. M. (2009). Economic efficiency analysis of banana farmers in Kiambu East District of Kenya: Technical inefficiency and marketing efficiency. *J. Dev. Sustain. Agric.* 4, 118–127. doi: 10.11178/jdsa.4.118
- Olanrewaju, O. S., and Babalola, O. O. (2019). Streptomyces: implications and interactions in plant growth promotion. *Appl. Microbiol. Biotechnol.* 103, 1179–1188. doi: 10.1007/s00253-018-09577-y
- Omotayo, O., Igiehon, O., and Babalola, O. (2021). Metagenomic study of the community structure and functional potentials in maize rhizosphere microbiome: elucidation of mechanisms behind the improvement in plants under normal and stress conditions. *Sustainability* 13:8079. doi: 10.3390/su13148079
- Peterson, D., Bonham, K. S., Rowland, S., Pattanayak, C. W., Resonance Consortium, and Klepac-Ceraj, V. (2021). Comparative analysis of 16S rRNA gene and metagenome sequencing in pediatric gut microbiomes. *Front. Microbiol.* 12:670336. doi: 10.3389/fmicb.2021.670336
- Puente-Sánchez, F., García-García, N., and Tamames, J. (2020). SQMtools: automated processing and visual analysis of 'omics data with R and anvi'o. *BMC Bioinformatics* 21:358. doi: 10.1186/s12859-020-03703-2
- Rossmann, B., Müller, H., Smalla, K., Mpiira, S., Tumuhairwe, J. B., Staver, C., et al. (2012). Banana-associated microbial communities in Uganda are highly diverse but dominated by Enterobacteriaceae. *Appl. Environ. Microbiol.* 78, 4933–4941. doi: 10.1128/AEM.00772-12
- Sarkar, S., Kumar, R., Kumar, A., Kumar, U., Singh, D. K., Mondal, S., et al. (2022). "Role of soil microbes to assess soil health," in *Structure and Functions of Pedosphere*, eds. B. Giri, R. Kapoor, Q.-S. Wu, and A. Varma (New York, NY: Springer Nature Singapore), 339–363. doi: 10.1007/978-981-16-8770-9\_14
- Sayers, E. W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K. D., and Karsch-Mizrachi, I. (2019). GenBank. *Nucleic Acids Res.* 47(D1), D94–D99. doi: 10.1093/nar/gkz956
- Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27, 863–864. doi: 10.1093/bioinformatics/btr026
- Shahrajabian, M. H., Sun, W., and Cheng, Q. (2021). The importance of Rhizobium, Agrobacterium, Bradyrhizobium, Herbaspirillum, Sinorhizobium in sustainable agricultural production. *Not. Bot. Horti Agrobot. Cluj-Napoca* 49:12183. doi: 10.15835/nbha49312183
- Shi, G., Sun, H., Calderón-Urrea, A., Li, M., Yang, H., Wang, W., et al. (2021). Bacterial communities as indicators of soil health under a continuous cropping system. *Land Degrad. Dev.* 32, 2393–2408. doi: 10.1002/ldr.3919
- Singh, S., Sharma, A., Khajuria, K., Singh, J., and Vig, A. P. (2020). Soil properties changes earthworm diversity indices in different agro-ecosystem. *BMC Ecol.* 20:27. doi: 10.1186/s12898-020-00296-5
- Sui, X., Zhang, R., Frey, B., Yang, L., Li, M.-H., and Ni, H. (2019). Land use change effects on diversity of soil bacterial, Acidobacterial and fungal communities in wetlands of the Sanjiang Plain, northeastern China. *Sci. Rep.* 9:18535. doi: 10.1038/s41598-019-55063-4
- Swafu, S. M., and Dlamini, P. E. (2022). Unlocking the land capability and soil suitability of makuleke farm for sustainable banana production. *Sustainability* 15:453. doi: 10.3390/su15010453
- Tamames, J., and Puente-Sánchez, F. (2019). SqueezeMeta, a highly portable, fully automatic metagenomic analysis pipeline. *Front. Microbiol.* 9:3349. doi: 10.3389/fmicb.2018.03349
- Tan, S., Jiang, L., Liu, J., Zeng, Z., Xiao, Y., Wu, X., et al. (2023). Rhizosphere microorganisms and soil physicochemical properties of restored wetland plant communities at cutting slash of *Populus deltoides* in Dongting Lake. *Wetlands* 43:48. doi: 10.1007/s13157-023-01696-1
- Tang, Y., Zhou, S., Xiao, Y., Zhang, T., Tao, X., Shi, K., et al. (2024). Exploring the microbial ecosystem of *Berchemia polyphylla* var. *leicladia*: a comprehensive analysis of endophytes and rhizospheric soil microorganisms. *Front. Microbiol.* 15:1338956. doi: 10.3389/fmicb.2024.1338956
- Wahome, C. N., Maingi, J. M., Ombori, O., Kimiti, J. M., and Njeru, E. M. (2021). Banana production trends, cultivar diversity, and tissue culture technologies uptake in Kenya. *Int. J. Agron* 2021, 1–11. doi: 10.1155/2021/6634046
- Wahome, C. N., Maingi, J. M., Ombori, O., Njeru, E. M., Muthini, M., and Kimiti, J. M. (2023). Diversity and abundance of bacterial and fungal communities in rhizospheric soil from smallholder banana producing agroecosystems in Kenya. *Front. Hortic.* 2:1061456. doi: 10.3389/fhort.2023.1061456
- Wang, X., Jarmusch, S. A., Frisvad, J. C., and Larsen, T. O. (2023). Current status of secondary metabolite pathways linked to their related biosynthetic gene clusters in *Aspergillus* section *Nigri*. *Nat. Prod. Rep.* 40, 237–274. doi: 10.1039/D1NP00074H
- Ye, Y., and Doak, T. G. (2009). A parsimony approach to biological pathway reconstruction/inference for genomes and metagenomes. *PLoS Comput. Biol.* 5:e1000465. doi: 10.1371/journal.pcbi.1000465
- Zaheer, R., Noyes, N., Ortega Polo, R., Cook, S. R., Marinier, E., Van Domselaar, G., et al. (2018). Impact of sequencing depth on the characterization of the microbiome and resistome. *Sci. Rep.* 8:5890. doi: 10.1038/s41598-018-24280-8
- Zhang, H., Ullah, F., Ahmad, R., Ali Shah, S. U., Khan, A., and Adnan, M. (2022). Response of soil proteobacteria to biochar amendment in sustainable agriculture - a mini review. *J. Soil Plant Environ.* 1, 16–30. doi: 10.56946/jspae.v1i2.56
- Zhang, H., Wang, R., Chen, S., Qi, G., He, Z., and Zhao, X. (2017). Microbial taxa and functional genes shift in degraded soil with bacterial wilt. *Sci. Rep.* 7:39911. doi: 10.1038/srep39911
- Zhang, Y.-Y., Wu, W., and Liu, H. (2019). Factors affecting variations of soil pH in different horizons in hilly regions. *PLoS ONE* 14:e0218563. doi: 10.1371/journal.pone.0218563
- Zhou, L., Wu, S., and Ma, M. (2023). First insights into diversity and potential metabolic pathways of bacterial and fungal communities in the rhizosphere of *Argemone mexicana* L. (Papaveraceae) from the water-level-fluctuation zone of Wudongde Reservoir of the upper Yangtze river, China. *Biodivers. Data J.* 11:e101950. doi: 10.3897/BDJ.11.e101950