EXPLORING THE SOIL-ASSOCIATED BACTERIAL MICROBIOME OF COFFEE PLANTATIONS IN DIFFERENT REGIONS OF COLOMBIA: A METABARCODING APPROACH †

[EXPLORACIÓN DEL MICROBIOMA BACTERIANO ASOCIADO AL SUELO DE PLANTACIONES DE CAFÉ EN DIFERENTES REGIONES DE COLOMBIA: UN ENFOQUE DE METABARCODING]

Lorena Jacqueline Gómez-Godínez1, Víctor Ochoa2, Valeria Faggioli3 and Marco Cristancho2†


2Plant Pathology Department, National Center for Coffee Research- CENICAFE, Pedro Uribe Mejía, km 4 via Chinchiná-Manizales, C.P. 170009. Chinchiná-Manizales, Colombia. Email: marco.cristancho@cafedecolombia.com

3Instituto Nacional de Tecnología Agropecuaria INTA, Av. Rivadavia 1439. Ciudad Autónoma de Buenos Aires. C.P. 11033AAE. Buenos Aires, Argentina

*Corresponding author

SUMMARY

Introduction: Coffee is one of the leading tropical crops produced worldwide. Colombia ranks third in coffee production in the world. Microorganisms associated with coffee plants can have many biotechnological applications, such as plant growth promotion and biological control. **Objective:** To describe the bacteria present in the soil associated with coffee cultivation. This was done using a bulk sequencing or 16S rRNA metabarcoding approach. **Methodology:** Soil samples were collected to analyze their microbiome from three different departments of Colombia (Cauca, Risaralda and Magdalena). The plants were of different ages, and the crop management was different (conventional and organic). Subsequently, an analysis was carried out using Qiime2 to describe the communities associated with coffee cultivation and soil chemical properties. **Results:** Some important genera were identified, such as *Janthinobacterium*, *Bacillus*, *Actinomadura* and *Actinoallomurus*. These genera can be used as plant growth promoters, organic matter transformers, producers of antibiotics and metabolites with potential biotechnological applications. **Implications:** The study presented in this manuscript describes the communities associated with the soil in coffee cultivation. However, a much more complex approach could be through metagenomics, where the bacterial communities and the functions of; however, this vision is up to ten times higher in cost. **Conclusion:** Through the results obtained, it is concluded that the studies carried out through metabarcoding help to understand the composition of the microorganisms associated with coffee and, with this, try to elucidate the functions of the associated microorganisms.

**Key words:** Metabarcoding; coffee; uncultured bacteria species.

RESUMEN

Introducción: El café es uno de los principales cultivos tropicales producidos a nivel mundial. Colombia ocupa el tercer lugar en producción de café en el mundo. Los microorganismos asociados a las plantas de café pueden tener una gran cantidad de aplicaciones biotecnológicas, como la promoción de crecimiento vegetal y el control biológico. **Objetivo:** Describir las bacterias presentes en el suelo asociadas al cultivo del café. Esto se hizo mediante un enfoque de secuenciación masiva o metabarcoding del gen ARNr 16s. **Metodología:** Se recolectaron muestras de microbioma del suelo de tres departamentos diferentes de Colombia (Cauca, Risaralda y Magdalena). Las plantas tuvieron diferentes edades y el manejo del cultivo fue diferente (convencional y orgánico). Posteriormente se realizó un análisis mediante Qiime2 para describir las comunidades asociadas al cultivo de café. **Resultados:** Se identificaron algunos Filo importantes como, por ejemplo, Acidobacterias, Actinobacterias, Firmicutes y Verrucomicrobia, Dentro de
los géneros importantes, se lograron identificar a *Janthinobacterium*, *Bacillus*, *Actinomadura* y *Actinoallomur*us. Estos géneros pueden ser utilizados como promotores del crecimiento vegetal, transformadores de materia orgánica, productores de antibióticos y metabolitos con potenciales aplicaciones biotecnológicas. **Implicaciones**: El estudio presentado en este manuscrito es descriptivo, de las comunidades asociadas al suelo en el cultivo de café, sin embargo, un enfoque mucho más complejo podría ser a través de metagenómica donde se puede describir a las comunidades de bacterias y las funciones de estas, sin embargo, esta visión es hasta diez veces más elevada en cuestiones de costo. **Conclusión**: A través de los resultados se concluye que los estudios realizados a través de metabarcoding ayudan a entender la composición de los microorganismos asociados al café y con esto tratar de dilucidar las funciones de los microorganismos asociados.  
**Palabras clave**: metabarcode; café; especies de bacterias no cultivadas.

**INTRODUCTION**

Coffee is the second most consumed beverage after water and the most traded tropical agricultural product in the world (Mussatto *et al.*, 2011); it is a perennial plant belonging to the Rubiaceae family. The *Coffea* genus consists of about 103 species, where the *Coffea arabica* species is the most used for the production of beverages worldwide (Davis *et al.*, 2006). Colombia is one of the world’s leading coffee producers, ranking third after Brazil and Vietnam. Within Colombia, Cauca is the fourth-largest coffee-producing department, Risaralda is sixth, and Magdalena is fourteenth. Soils are one of the most dynamic ecosystems on the planet. They are home to a diverse community of microorganisms, including bacteria, fungi, viruses, archaea, and protozoa. These microorganisms interact with each other and with the soil environment to perform essential ecosystem functions (Lladó *et al.*, 2017).

The study of plant-associated microorganisms is important for a number of biotechnological applications, such as the biological control of plant pathogens, the promotion of plant growth, and the isolation of active compounds. Biodiversity characterization has also been used to monitor and assess the state of the environment (Bonatelli *et al.*, 2021; Lahlali *et al.*, 2022; Lopes *et al.*, 2021). Changes in the structure of the microbial community can influence the stability of populations in the face of disturbances. Therefore, the identification and conservation of soil microbial components is essential for resilience under changing conditions, as well as to support current and future food security (Bertola *et al.*, 2021).

The importance of identifying the microorganisms associated with coffee is a crucial step to allow the prospecting of beneficial microorganisms that can be used as biofertilizers in soils and starter cultures to improve the quality of the coffee drink since it is known that the profile microbial promotes changes in the sensory and chemical profile in coffee fermentation (Junior *et al.*, 2021; Duong *et al.*, 2020). Furthermore, coffee cultivars are more adapted to some regions than others, and their associated microorganisms may also be shaped by climatic and edaphic characteristics, which may be related to the sensory experience obtained from this product (terroir) (Véloso *et al.*, 2020; Williams *et al.*, 2022).

One of the factors that can modify the communities of microorganisms in the soil is tillage. It is known that conventional tillage causes physical alterations of the upper horizon of the soil and can decrease organic matter (Brangari *et al.*, 2022). On the other hand, zero tillage and organic management are practices that can favor the enrichment of soil with organic matter and nutrients, reduce soil erosion, increase water availability, and improve the stability of the soil system in the face of changes in humidity and temperature (Sun *et al.*, 2018; Schmidt *et al.*, 2018; Xin *et al.*, 2018). It has been shown that various agronomic management can significantly alter soil biomass, as well as the composition of the microbial communities present. For example, greater microbial diversity and more microbial biomass have been reported in soil treated under no-till management compared to soils with conventional management (Wang *et al.*, 2017; Yin *et al.*, 2017; Mathew *et al.*, 2012; Guo *et al.*, 2016).

Most studies on crop-associated bacteria and their community structure have been conducted using culture-dependent approaches. However, since a high percentage of natural bacteria remain in a non-culturable state (Bodor *et al.*, 2020), the use of culture-independent methods provides additional information on the diversity of bacterial communities (Rincon-Florez *et al.*, 2013).

In the last decade, the use of high-throughput sequencing has substantially improved our understanding of the global biogeography and ecology of soil bacterial and fungal communities (Howe *et al.*, 2021). This strategy, along with metabarcoding, represents a fast, accurate, and cost-effective technology for testing thousands of environmental samples simultaneously.

The objective of this study was to analyze the communities associated with coffee, of different ages and grown in different regions of Colombia, as well as under different agricultural systems, conventional and organic, using a crop-independent approach.
MATERIAL AND METHODS

Description of the study site

The study was conducted in three Colombian coffee-growing regions: El Tambo (Cauca), Quinchía (Risaralda), and Minca (Magdalena). A total of 20 coffee farms were sampled, 8 in Cauca, 8 in Risaralda, and 4 in Magdalena. The soil samples were collected between July and August 2021, under the climatic and environmental conditions of each region. Around the coffee-plant, the surface of each area was cleaned to remove any organic material present. A sterile metal spatula was used to collect approximately 50-150 g of soil (a mixture of soil, sand, and small soil grains) from a depth of 0.2 meters, the distance between coffee plant for sampling was 1.5 meters. Five different coffee plants samples were collected from each farm, for a total of 30 soil samples. The samples were stored in sterile Ziploc-type plastic bags and transported to the Laboratory of Molecular Interactions of Microorganisms in Agriculture - LIMMA, located at the Universidad de Los Andes, for further processing (Figure 1).

Physical-chemical soil analysis

The physicochemical analyses were performed on 100 g of each sample at the Soil and Foliar Laboratory of the Technological University of Pereira, Colombia. The pH was measured in water (1:1 w/v aqueous extract) using a potentiometer. The Organic Matter (OM) was determined with the Walkley-Black photometric method. Electrical Conductivity (EC) was measured in paste saturated with water using a conductivity meter (Keren 2016; Kalra and Yash 1995). The following ions were determined using the atomic absorption technique (Acetate+EDTA): Iron (Fe), Manganese (Mn), Zinc (Zn), and Copper (Cu). Boron (B) was determined by extraction with monocalcium phosphate, azomethine, and photometric detection (Mauya et al., 2018). Sulfur (S) was determined by the turbidimetric method under extraction with monocalcium phosphate (Wall et al., 1980). Phosphorus (P) was analyzed with the Bray II photometric method. Aluminum was also measured (Wuensch et al., 2015; Lin and Coleman et al., 1960).

The texture was determined by Bouyoucos and the exchangeable acidity by the volumetric method under extraction with KCl (Beretta et al., 2014). The bases Potassium (K), Calcium (Ca), Magnesium (Mg), and Sodium (Na) were analyzed by the ammonium acetate method, using the atomic absorption technique. Solubles were measured in water-saturated paste, using the same technique. Finally, the Cation Exchange Capacity (CEC) was analyzed volumetrically (Sumner et al., 1996; Soylak et al., 2003; Yunan et al., 2028).

Figure 1. Graphical overview of the methods of this study.
Sample processing and metagenomic DNA extraction

Each Ziploc bag was tempered before processing. Using sterile metallic mini-spatulas, 250 mg of soil was taken and added to a sterile 2 ml Eppendorf tube. DNA was extracted using the DNeasy® PowerSoil® Pro-Kit (Qiagen) and EZ-10 Spin Column Soil DNA Miniprep Kit (BioBasic) following the manufacturer's instructions. Possible contaminants were removed with the kit's own resins, and then the DNA was separated with silica filters. The DNA extraction was performed individually for each replicate. DNA was quantified using spectrophotometry with the NanoDrop™ 2000 (Thermo Fisher Scientific™) and Qubit 2.0 (Thermo Fisher Scientific™) kits following the manufacturer's protocols. The integrity of the DNA was observed by electrophoresis in a 1% agarose gel. There were 6 conditions with 3 or 5 replicates per condition, having a total of 25 samples for metagenomic sequencing. These were stored at -80°C in sterile Eppendorf tubes until further processing (Figure 1).

Sequencing of the metagenome

Libraries were constructed using the Illumina Metagenomics Sequencing technique and the MiSeq PE250 (Matched Ends 2x250) Sequencing Kit. The 16S V3/V4 region of bacteria was amplified using primers 341F (5′-CCTAYGGGRBGCASCAG-3′) and 806R (5′-GGACTACNNGGTTATCTAAT-3′) (Yu et al., 2005). All sequencing was performed by external services of the National Institute of Agricultural Technology - Marcos Juárez Agricultural Experimental Station, in Argentina (Figure 1).

Bioinformatics Analysis

Sequence quality was assessed using FastQC (v0.11.9) (Andrews, 2010). Quality control and processing were performed using QIIME2 (v2021.4.0) (Bolten et al., 2019). Sequences were demultiplexed using a Q-score of 30, Illumina adapters, and poor sequencing quality were removed with cutadapt and DADA2 (Callahan et al., 2016). Clean sequences were compared to the greengenes database (gg-13-8-99-515-806-nb-classifier.qza), and taxonomic assignment (ASVs) was performed. Alpha diversity indices (observed species, Chao1, Shannon, and Simpson) were calculated from the rarefaction curves. Beta diversity indices were calculated using Euclidean distance and UniFrac (weighted and unweighted) and visualized through principal coordinate analysis (PCoA) (Figure 1).

RESULTS

Soil analysis

Soils from three colombian regions (Cauca, Risaralda, and Magdalena) were characterized. The pH of the soils varied, with Cauca having the lowest pH of 4.5, followed by Risaralda 4.9 and 5.0 to Magdalena (Table 1). The percentage of organic matter also varied, with Risaralda_9 having the lowest percentage (7.9%), followed by Risaralda_15 (8.7%), Risaralda_12 (14.6%), and Cauca (23%). The nitrogen content varied, with Risaralda_9 having the lowest percentage (0.33%), followed by Risaralda_15 (0.42%), Risaralda_12 (0.58%), and Cauca (0.72%). The percentage of phosphorus was highest in Risaralda_9 (32%), followed by the other samples (4%). The percentage of potassium was similar across all samples, ranging from 1% to 1.5%. The site with the highest amount of iron was Risaralda_12, with 218 parts per million (ppm), followed by Cauca with 169 ppm; Manganese oscillated between 31 ppm (Risaralda_12), and the lowest concentration was presented in Cauca and Risaralda_15 with 18 ppm. Zinc was found in a concentration of 14 ppm in Cauca and 4 ppm in Risaralda_15 (Table 1). The samples from this site are divided into Magdalena_17 (30-year-old plants, conventional management) and Magdalena_19 (new plants, conventional management). Samples from Magdalena_19 were not analyzed for physical-chemical properties; this is denoted as NA in Table 1. The main reason the analysis was not carried out was due to the conditions of the sample (soil texture characteristics, made DNA extraction difficult) and the fact that there was no opportunity to repeat the sampling due to difficult access to the place.

Profile of bacterial communities associated with coffee cultivation

Amplicons from 25 samples corresponding to the different treatments were obtained and analyzed as described above. A total of 2,649,699 reads were obtained, with an average of 115,204 reads per sample. After filtering, 1,901,635 reads were obtained and 132 ASVs were identified. These were assigned to 14 phyla, 24 classes, 14 orders, 23 families, and 42 bacterial genera. The diversity indices, calculated from the rarefaction curves, remained constant across the different treatments.

The most abundant phylum in the soil samples from the Cauca department was Firmicutes. Chloroflexi and Acidobacteria were the next most abundant phyla. In the soil samples from the Risaralda department, Acidobacteria was the most abundant phylum, followed by Nitrospirae and Proteobacteria. In the soil samples from the Magdalena department, the most
abundant phyla were Planctomycetes, Acidobacteria, and Verrucomicrobia. The results are summarized in Figure 2.

In most samples, the most abundant family was Hyphomicrobiaceae, however, in the samples from Risaralda_19 and Risaralda_9, the most abundant family was Koribacteraceae, and Pseudomonadaceae, respectively. The Rhodospirillaceae family was the second most abundant in the samples from Cauca, Magdalena_17 and Magdalena_19. In the case of the samples from Risaralda_9 and Risaralda_12, the second most abundant family was Micrococcaceae, and in the case of the sample from Risaralda_15 Koribacteraceae. Koribacteraceae was the third family in the samples from Cauca and Magdalena_17, in the samples from Risaralda_15 and Magdalena_19 the family Syntrophobacteraceae and Sinobacteraceae, ranked third respectively (Figure 3).

Finally, some of the genera identified in Cauca were Conexibacter, Bacillus, Actinomadura, Rhodococcus, Syntrophobacter, Pedobacter, and Burkholderia. In the Risaralda_9 samples, Chryseobacterium, Pedobacter, Janthinobacterium, Pseudomonas and Rhodococcus were found, in Risaralda_12, Arthrobacter, Labrys, Pilimelia, Clostridium, and Rhodococcus, in Risaralda_15, Anaeromyxobacter, Steroidobacter, Nitrospira, Mesorhizobium and Pedobacter. In Magdalena_17, Nostocoida, Conexibacter, Actinomadura, Burkholderia and Edaphobacter, and finally in Magdalena_19 Novosphingobium, Kaistobacter, Methylibium, Flavobacterium and Phenyllobacterium (Figure 4).

The analysis of beta diversity, based on PCoA, did not allow us to corroborate differences in the structure of the microbiota associated with the coffee soil in the different regions of Colombia (Figure 5).

**DISCUSSION**

Soils intended for agricultural crops are often low in organic matter, with contents of less than 4% to 6% (Hatten & Liles 2019). However, this is not the case for coffee crops, which are typically grown under shade and benefit from the mulching effect of leaf litter, finding percentages of organic matter from 4% to 11% (López-Báez et al., 2016; Darfis and Putri 2021). This helps to protect the soil from erosion and enhances the accumulation of organic matter. In our study, we found that the organic matter content of coffee soils ranged from 5.5% to 23%. Organic matter is important for plant growth because it provides nutrients, such as nitrogen (Zheng et al., 2023). In Colombia, mathematical models have been developed to estimate the nitrogen content of coffee soils based on the organic matter content (Salazar et al., 2023). Our results confirm this relationship, as we found that the soil samples from Cauca, which had the highest organic matter content (23%), also had the highest nitrogen content (0.72%).

![Table 1. Description of study sites and soil chemical properties. Where: pH: potential of hydrogen, N: Nitrogen, OM: Organic Matter, CEC: Cation exchange capacity, K: Potassium, Ca: Calcium, Mg: Magnesium, Al: Aluminum, Fe: Iron, Mn: Manganese, Zn: Zinc, Cu: Copper, B: Boron, S: Sulfur, P: Phosphorus and NA: Does not apply.](image-url)
Figure 2. Phylum bacterial communities associated with soil samples obtained from different regions of Colombia. The bars represent the relative abundance. Sample 1 a,c,d: samples obtained from the department of Cauca in the municipality of Tambo, 4-month-old plants with organic management; sample 9 a,b,d: samples obtained from the department of Risaralda in the municipality of Quinchia, 10-year-old plants with organic management; sample 12 a,b,c,d and e: samples obtained from the department of Risaralda in the municipality of Quinchia, 5-year-old plants with organic management; sample 15 a,c,d and e: samples obtained from the department of Risaralda in the municipality of Quinchia, 5-year-old plants with organic management; sample 17 a,b,c,d and e: samples obtained from the department of Magdalena in the municipality of Minca, 30-year-old plants with conventional management; sample 19 a,b,c,d and e: samples obtained from the department of Magdalena in the municipality of Minca, new plants with organic management.

Microorganisms are the living component of soil, and their abundance and activity depend on a variety of factors, such as soil type and management, plant type and growth, and root exudates. In this study, we report the diversity of bacteria associated with coffee plants of different ages, under conventional and organic management. We found that the Acidobacteria phylum was abundant in samples from the Cauca department, Risaralda_9, Magdalena_17, and Magdalena_19. The Acidobacteria phylum is positively correlated with organic matter and is involved in carbon sequestration. Additionally, Acidobacteria dominate soils with high organic matter content and are involved in the degradation of lignocellulose (Rawat et al., 2012).
Figure 3. Family bacterial communities associated with soil samples obtained from different regions of Colombia. The bars represent the relative abundance.

On the other hand, the Firmicutes phylum was most abundant in the soils of Cauca in comparison with Risaralda and Magdalena sites. This phylum is typically found in soil, but in lower abundance than other phyla (Youssef and Elshahed, 2008). Firmicutes include gram-positive bacteria that can form endospores, which allow them to survive in adverse environmental conditions. Bacillus is a genus of Firmicutes that is commonly found in soil.

In the Risaralda_9 sample, we found a high percentage of the Proteobacteria phylum, which is common in agricultural ecosystems and plays a variety of roles in the soil (Valinsky et al., 2002), including carbon, nitrogen, and sulfur cycling (Spain et al., 2009). The Nitrospirae phylum was also identified in the Risaralda_9 and Magdalena_19 samples. This phylum is characterized by a single class, Nitrospira, and its members are aerobic chemolithotrophs that can perform a variety of metabolic functions, including nitrification, sulfate reduction, and magnetotaxis. Some species of Nitrospira are important participants in the nitrogen cycle (Hu., et al., 2021).

Another interesting and consistent phylum found in the samples was Verrucomicrobia. It was found in high percentages in the samples from Risaralda_12, Magdalena_17, and Magdalena_19. Verrucomicrobia are difficult to cultivate, and most are facultative mesophiles or strict anaerobes (Bergmann et al., 2011). They are saccharolytic and oligotrophic. Little is known about their function, but some representatives of Verrucomicrobia associated with rice plants have been found to incorporate plant-derived carbon for growth (Hernández et al., 2015).

This may explain the high abundance of Verrucomicrobia in the Risaralda_12 samples, which had the highest percentage of organic matter.
Figure 4. Family bacterial communities associated with soil samples obtained from different regions of Colombia. The bars represent the relative abundance.

The phylum Chlamydiae is a small group of obligate intracellular bacteria that were once thought to only infect humans and other animals (Elwell et al., 2016). However, recent research has shown that Chlamydiae are actually much more widespread and can be found in a wide variety of host organisms, including protists, arthropods, and vertebrates. Chlamydiae have also been found in a variety of environmental samples, suggesting that they may play a role in the cycling of nutrients in the environment.

The most abundant genera in the samples from Cauca were Conexibacter, Bacillus, and Actinoallomurus. Conexibacter is a genus of Actinobacteria that is typically found in acidic soils, such as those in the Cauca region (pH 4.5). These bacteria are known to contribute to the carbon cycle in soil environments and may also play a role in the nitrogen cycle (Pukall et al., 2010; Seki et al., 2012). Bacillus is another genus that was found in the samples. This genus is common in coffee plants and some species have been shown to inhibit the growth of pathogenic fungi, such as Colletotrichum gloeosporioides and Fusarium oxysporum (Kejela et al., 2016). Bacillus species also have a variety of other beneficial properties, including the ability to promote plant growth, solubilize and mineralize nutrients, produce phytohormones, and secrete antimicrobial compounds (Saxena et al., 2020). Actinoallomurus is the third most abundant genus in the samples from Cauca. These bacteria are of great interest due to their secondary metabolism, which could be used to find organisms with potential in the production of antibiotics against resistant pathogens (Pozzi et al., 2010).
In the samples from Risaralda_9 we found in greater abundance the genera *Chryseobacterium*, *Pedobacter*, *Janthinobacterium*. *Chryseobacterium* has been reported in different environments such as wastewater, clinical samples, and soil (Kämpfer et al., 2003, 2009; Shen et al., 2005). The strains that represent this genus are Gram-negative, immobile, and yellow in color. Some species have been reported as growth promoting bacteria (Chhetri et al., 2022). Bacteria within the *Pedobacter* genus are characterized by being resistant to antibiotics (Bjerketorp et al., 2021), and some species within this genus have been characterized by breaking down cellulose (Zhou et al., 2019). Within *Janthinobacterium* we find gram-negative bacteria, producers of some substances such as violacein and capable of inhibiting some pathogenic fungi such as *Pythium ultimum*, *Rhizoctonia solani* and *Rhizoctonia oryzae*, in crops such as wheat (Yin et al., 2021). It has been reported that when a soil is enriched with bacteria of this genus, it shows the ability to be less susceptible to dollar spot caused by the pathogenic fungus *Clarierea* spp. (Chou et al., 2021). Finally, *Janthinobacterium* has been shown to promote plant growth, thanks to the production of indole-3-acetic acid (Kuffner et al., 2008).

The most abundant genera in the Risaralda_12 samples are *Arthrobacter*, *Labrys*, and *Pilimelia*. *Arthrobacter* is a genus of obligate aerobic bacteria that grow in a rod-coccus cycle (Gobbetti and Rizzello, 2014). They are commonly found in soils, aerial surfaces of plants, and sewage sediments. They do not form endospores and are highly proteolytic. Some species of *Arthrobacter*, such as *Arthrobacter agilis*, have been reported as plant growth promoting bacteria. *A. agilis* was isolated from corn and has the ability to solubilize iron and increase the biomass of beans and *Medicago sativa* (Valencia-Cantero et al., 2007; Velázquez-Becerra et al., 2011).

*Labry* is a genus of aerobic, gram-negative, motile, and non-spore forming bacteria. Their shape varies from spherical to short rods. They were first described by Vasil’eva and Semenov (1984). Some species of *Labrys*, such as *Labrys metilaminiphilus*, can use methanol as a carbon source (Miller et al., 2005). *Pilimelia* is a genus of bacteria within the Actinobacteria. They produce branched, septate substratum hyphae, but the aerial mycelium does not develop. All *Pilimelia* species are aerobic, mesophilic, and gram-positive (Schaefer, 1973). Some Pilimelias have the ability to degrade plant polymers such as starch, cellulose, hemicellulose, pectin, or lignin. This may be due to the litter present in coffee crops.

The main genera in the Risaralda_15 samples were *Anaeromyxobacter*, *Steroidobacter*, and *Nitrospira*. *Anaeromyxobacter* are Fe (III) reducers, meaning they can reduce the oxidized form of iron (Fe(III)) to the reduced form (Fe(II)) (Sanford et al., 2002; Wang et al., 2020). This process is important in the cycling of iron in the environment. Steroidobacter are found in
soil, specifically organic soybean soils. They are believed to have a role in carbon decomposition, which is the process by which organic matter is broken down into smaller molecules (Lian et al., 2017). Nitrospira are found in acid soils and have an important role in nitrification. Nitrification is the process by which ammonia is oxidized to nitrate (Sêneca et al., 2020). The abundance of each genus was affected by the environmental conditions. For example, the abundance of *Anaeromyxobacter* was favored in rice cultures enriched with ferrihydrite and goethite. The abundance of *Nitrospira* decreased due to drought.

The genera *Nostocoida*, *Conexibacter*, and *Actinomadura* were abundantly found in the Magdalena_17 samples. *Nostocoida* bacteria that were isolated from activated sludge (this bacterium needs nitrogen to grow). (Blackall et al., 2000). *Conexibacter* are bacteria that are common in soil. They may have a role in the decomposition of organic material. *Actinomadura* are bacteria that are common in soil. They can produce antibiotics, vitamins, enzymes, and other valuable biologically active substances (Trüper et al., 1992).

The Magdalena_19 samples contained genes from the genera *Novosphingobium*, *Kaistobacter*, and *Methylbium*. *Novosphingobium* are bacteria that have been found as endophytes of some crops, such as rice. Some species have growth promotion potential, and one has been identified as diazotrophic, meaning it can fix nitrogen (Addison et al., 2007; Gao et al., 2015; Zhang et al., 2016). *Kaistobacter* are bacteria that are favored when substances are added to the soil to protect plants against fungal diseases (Baćmaga et al., 2021). *Methylbium* were a genus described by Nakatsu et al. (2006). Many members of this genus hydrolyze urea, and some species have also shown their ability to biodegrade some undesirable compounds in biogas production (Boada et al., 2020; Nakatsu et al., 2006).

We found several genera of bacteria with great biotechnological potential, including *Janthinobacterium*, *Bacillus*, *Actinomadura*, and *Actinoallomurus*. These bacteria can produce metabolites or antibiotics that are used to treat multi-resistant bacteria, promote plant growth, and inhibit pathogenic fungi (Kuffner et al., 2008; Trüper et al., 1992; Kejela et al., 2016; Saxena et al., 2020).

Numerous studies have pointed out the importance of soil-associated microbial communities in plant growth promotion. These studies have documented the capacity of bacterial genera within communities, which are responsible for producing plant growth inducers, metabolites and antibiotics, among other substances (Poria et al., 2022; Katsenios et al., 2022; Gómez-Godínez et al., 2023). Within this work, we postulate the ability of bacteria to colonize the roots of coffee plants and how they could establish beneficial symbiotic interactions. This symbiosis could enhance the absorption of essential nutrients by the coffee plant, thus improving the crop’s ability to take advantage of nutrients efficiently, and the synthesis of compounds related to the promotion of plant growth could positively influence root development. Water absorption, resistance to adverse conditions, and the ability to inhibit pathogenic fungi could prevent soil diseases and directly affect coffee productivity. The selective suppression of pathogens the microbial communities could contribute to the overall health of the plants and, therefore, to an increase in the productivity of coffee beans (Souza et al., 2015; Glick, 2012; Abdelaal et al., 2021; Urgiles-Gómez et al., 2021; Asad et al., 2023).

Finally, it is expected that the study will not only deepen our understanding of the diversity associated with coffee plants but also lay the foundation for innovative and sustainable strategies in the controlled use of these bacterial communities, such as, using them in bioaugmentation. Moreover, the controlled introduction of beneficial bacteria can modulate the structure of the community, favoring interactions between community members, suppressing pathogens, and increasing plant growth and coffee bean production. Detailed exploration of these biotechnological potentials could have practical implications in agriculture, improving the sustainability and profitability of coffee globally.

**Funding.** This work was supported by CABANA (Capacity building for bioinformatics in Latin America) grant, on behalf of the Global Challenges Research 408 Fund [BB/P027849/1].

**Conflict of interests.** The authors declare no competing interests.

**Compliance with ethical standards.** This work does not require approval by a bioethical committee.

**Data availability.** The datasets generated during and/or analyzed during the current study are available in the european Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) with the project ID PRJEB50597 (https://www.ebi.ac.uk/). Or by email marco.cristancho@cafedecolombia.com

**Author contribution statement (CRediT).** L.J. Gómez-Godínez - Writing-original draft. Data curation and Writing-review & editing. V. Ochoa – Methodology and Writing-review & editing. V. Faggioli – Methodology and Writing-review & editing. M. Cristancho - Funding acquisition, Methodology, Project administration and Writing-review & editing.
REFERENCES


