

**UNIVERSIDAD DE PANAMÁ**  
**CAMPUS CENTRAL OCTAVIO MÉNDEZ PEREIRA**  
**FACULTAD DE CIENCIAS NATURALES EXACTA Y TECNOLOGÍA**  
**ESCUELA DE BIOLOGÍA**

**TESIS DE GRADO DE OPCIÓN AL TRABAJO DE GRADUACIÓN**

**ELABORADO POR:**  
**YUREIBY DEL CARMEN VALDES JIMENEZ**  
**3 - 738 - 1336**

**TRABAJO DE GRADUACIÓN PARA OPTAR POR EL TÍTULO DE**  
**LICENCIATURA EN BIOLOGIA CON ORIENTACION EN GENETICA Y**  
**BIOLOGIA MOLECULAR**

**PANAMÁ**  
**2022**



**UNIVERSIDAD DE PANAMÁ**  
**CAMPUS CENTRAL OCTAVIO MÉNDEZ PEREIRA**  
**FACULTAD DE CIENCIAS NATURALES EXACTA Y TECNOLOGÍA**  
**ESCUELA DE BIOLOGÍA**

**TESIS DE GRADO**

**METAGENÓMICA DE SUELOS ACRIsoles HÁPLICO DE PANAMÁ**

**POR:**

**YUREIBY DEL CARMEN VALDÉS JIMENEZ**

Trabajo de graduación sometido a consideración por la Escuela de Biología, Facultad de Ciencias Naturales Exactas y Tecnología, como requisito parcial para optar por el título de Licenciada en Biología con orientación en Genética y Biología Molecular

**PANAMÁ, REPÚBLICA DE PANAMÁ.**

**2022**

**UNIVERSIDAD DE PANAMÁ  
FACULTAD DE CIENCIAS NATURALES EXACTAS Y TECNOLOGÍA  
ESCUELA DE BIOLOGÍA**

---

**TRIBUNAL EXAMINADOR**

---

**Título: METAGENÓMICA DE SUELOS ACRIsoles HÁPLICO DE PANAMÁ**

**Por: Yureiby Del C. Valdes Jimenez** \_\_\_\_\_  
**3-738-1336**

**Asesora: Dra. Magaly de Chial** \_\_\_\_\_

**Jurado: Dr. Carlos Ramos** \_\_\_\_\_

**Mgtr. Brenda Mayorga** \_\_\_\_\_

**Fecha:** \_\_\_\_\_

## **DEDICATORIA**

*Este trabajo se lo dedico a Dios por darme sabiduría, prudencia y dedicación. A mi madre **Caridad Jimenez Ávila** quien es ejemplo de vida y me ha enseñado que los logros se construyen a base de sacrificios, esfuerzos y que las cosas materiales no nos dan felicidad; porque a pesar de sus escasos recursos económicos logró darme lo necesario para cubrir mis estudios, también a mi familia, hermanos, sobrinos y mi hermana **Isaura Salazar Jimenez** a quien Dios tenga en su santa gloria por darme aún más motivos para salir adelante y lograr ayudar a sus hijas. A todas las personas de mi querida provincia de Colón, les exhorto que salgamos hacia delante y con orgullo dejemos nuestra provincia y país en alto. A ti que lees esto por tomar de tu tiempo e informarte a través de esta investigación.*

***Yureiby del Carmen Valdes Jimenez***

## **AGRADECIMIENTO**

*Agradezco a Dios por haber estado conmigo cada día de mi vida y darme las virtudes necesarias para culminar de manera favorable este estudio, permitiéndome cumplir esta etapa y logro en mi vida. A la Dra. Magaly de Chial por brindarme sus conocimientos, destrezas y habilidades, también por haber aceptado ser la asesora principal de esta investigación y por haberme enseñado técnicas de investigaciones científicas a un nivel más elevado de lo que se hace a nivel de licenciatura, por enseñarme a utilizar programas y comandos para análisis bioinformáticos a través de Qiime2. Al Dr. Carlos Ramos por abrirme las puertas a esta investigación siendo el director de este estudio metagenómico en suelo por primera vez realizado en la República de Panamá, también por brindarme su tiempo en técnicas moleculares y elaboración de metodología en diversas ocasiones. Al Ing. Belgis Chial Sánchez por sus conocimientos de informática y haberme ayudado en los resultados de este estudio. A mi asesora la profesora Mgtr. Brenda Mayorga quienes me brindaron su ayuda y paciencia, así como sus conocimientos en la revisión de la escritura de este trabajo. Al laboratorio de Biología molecular de la vicerrectoría de Investigación y postgrado por haberme permitido utilizar las instalaciones y equipos de laboratorio. A mi madre quien con sus ánimos y humildad me inspiro, ayudo y apoyo a lo largo de este camino para poder culminar esta licenciatura. Y para finalizar a la Licda. Criseida Aquilar asistente del laboratorio de biología molecular y a todos los que desinteresadamente y sin reproches, aportaron a que este trabajo de graduación fuera posible.*

***Mis más sinceros agradecimientos***

# Índice

# Índice de Contenido

DEDICATORIA	4
AGRADECIMIENTO	5
Índice	6
Resumen	12
Resumen	13
Abstract	14
Abstract	15
Introducción	16
Capítulo I	19
Revisión Bibliográfica	19
Objetivos	40
I. Objetivos:	41
B. Objetivos Específicos:	41
Capítulo II	42
Metodología	42
Metodología Experimental	43-51
Capitulo III	51
Resultados y Discusión	51
Resultados	52
Discusión	97
Capitulo IV	102
Conclusiones	102
Recomendación	104
Bibliográficas	105
Anexo	123-134

# Índice de Tablas

Tabla 1. Sitos de colecta de suelos Acrisoles Háplico Cutánico y Acrisoles Háplico Hyperdístico.....	43
Tabla 2. Resultados iniciales de la secuenciación de la biblioteca.....	53
Tabla 3. Datos Físicoquímicos .....	54
Tabla 4. Sample metadata de las muestras de suelos analizadas. ....	55
Tabla 5. Total, de las secuencias procesadas de cada muestra de este estudio .....	55
Tabla 6. Recuento de secuencias (forward) por muestras .....	55
Tabla 7. Resumen de la longitud de las secuencias demultiplexadas .....	56
Tabla 8. Resumen de recuentos de las secuencias por muestra (Forward).....	56
Tabla 9. Estadística obtenida con DADA2 de las secuencias filtradas y no quiméricas por muestra de suelo.....	58
Tabla 10. Estadística del largo de las secuencias filtradas y el número de características (features) de las cuatro muestras .....	59
Tabla 11. Se asigna un id “Feature id” por lo que cada una de estas secuencias representativas ya está “mapeada”.....	59
Tabla 12. Artefacto FeatureData [Secuencia] QIIME 2 (table.qzv), que asigna identificadores de características en FeatureTable a las secuencias que representan en cada muestra de suelo estudiada. 60	
Tabla 13. Obtenida de la visualización tabla.qzv indicando el número de ASV (amplicon sequence variants) para las cuatro muestras de suelo y su frecuencia en estas.....	60
Tabla 14. variables de la metadata con la medida de alfa diversidad (p-value < 0.05).....	62
Tabla 15. P- value para esta métrica para la filogenética de Faith en cada uno de los parámetros físicoquímicos.....	64
Tabla 17. Permanova Tipo de Muestra.....	85
Tabla 18. Resultado Permanova Condición.....	85
Tabla 20. Tabla de valores para dominio.....	90
Tabla 21. Tabla de valores para filo .....	91
Tabla 22. Tabla de Valores para Clase.....	92



# Índice de Ilustraciones

Figura 1.	Horizontes Del Suelo.....	21
Figura 2.	12 grupos de suelos y sus 23 Clasificadores de Panama.....	25
Figura 3.	Análisis metagenómico. por Raimondi, 2013..	33
Figura 4.	ARNribosomal 16S .....	34
Figura 5.	Esquema del proceso seguido para la creación de bibliotecas mediante amplificación del gen del ARNr 16S.....	35
Figura 6.	Vista cercana - Chiriqui, Petro-Terminal Panamá-Puerto Armuelles. ....	44
Figura 7.	Vía hacia Boca Chica-Chiriquí. Coordenadas 8.266635, -82.190089.....	44
Figura 8.	Secuencias de ADN unidos al adapter y primer .....	47
Figura 9.	Procesamiento de metagenómica de amplicones para muestra de ADN y obtención final de data.....	48
Figura 10.	Gel de agarosa con las extracciones de DNA de las muestras de suelo. ....	52
Figura 11.	Visualización Qiime2 del resume los resultados de demultiplexación .....	57
Figura 12.	Grafico PCoAde la betadiversidad medida cualitativo con el aluminio.....	66
Figura 13.	Grafico PCoAde la betadiversidad medida cuantitativo con el aluminio.....	67
Figura 14.	Grafico PCoAde la betadiversidad medida cualitativo con la condición.....	68
Figura 15.	Grafico PCoAde la betadiversidad medida cuantitativo con la condición.....	69
Figura 16.	Grafico PCoAde la betadiversidad medida cualitativo con la localización.....	70
Figura 17.	Grafico PCoAde la betadiversidad medida cuantitativo con la localización.....	71
Figura 18.	Grafico PCoAde la betadiversidad medida cualitativo con el tipo de muestra....	72
Figura 19.	Grafico PCoAde la betadiversidad medida cuantitativo con tipo de muestra. ....	73
Figura 20.	Grafico PCoAde la betadiversidad medida cualitativo con el SampleID. ....	74
Figura 21.	Grafico PCoAde la betadiversidad medida cuantitativo con con SampleID. l.....	75
Figura 22.	Grafico PCoAde la betadiversidad medida cualitativo con el Magnesio. ....	76
Figura 23.	Grafico PCoAde la betadiversidad medida cuantitativo con el Magnesio. ....	77
Figura 24.	Grafico PCoAde la betadiversidad medida cualitativo con contenido orgánico.78	
Figura 25.	Grafico PCoAde la betadiversidad medida cuantitativo con el contenido orgánico. 79	
Figura 26.	Grafico PCoAde la betadiversidad medida cualitativo con el pH. ....	80
Figura 27.	Grafico PCoAde la betadiversidad medida cuantitativo con el pH.....	81
Figura 28.	Grafico PCoAde la betadiversidad medida cualitativo con el fosforo .....	82
Figura 29.	Grafico PCoAde la betadiversidad medida cuantitativo con el fosforo. ....	83
Figura 30.	Composición taxonómica de las muestras. ....	89

# Índice de Grafico

Grafico 1. Caja de bigotes para la diversidad filogenetica de faith en suelos intervenidos y no intervenidos .....	62
Grafico 2. Caja de bigotes de uniformidad de Pielou representando la uniformidad de las especies en las muestras y su relación con una las variables de la metadata. (aluminio) .....	63
Grafico 3. Permanova Tipo de Muestra .....	85
Grafico 4. Permanova Condición.....	86
Grafico 5. Columna SampleMetadata Barcode sequence – Metric Shannon.....	87
Grafico 6. Rarefacción Alfa. Esta visualización de resultado nos muestra la profundidad.....	88
Grafico 7. Taxonomía Nivel 1.....	90
Grafico 8. Taxonomía Nivel 2.....	91
Grafico 9. Taxonomía Nivel 3.....	92
Grafico 10. Taxonomía Nivel 4.....	93
Grafico 11. Taxonomía Nivel 5.....	94
Grafico 12. Taxonomía Nivel 6.....	95
Grafico 13. Taxonomía Nivel 7.....	96

## **Índice de Anexos**

1. Resumen de los procedimientos analíticos para la caracterización del suelo	124
2. Script de análisis bioinformático con QIIME2	130
3. Clasificación de Suelo de Panamá 2010	134

# **Resumen**

## Resumen

Durante la clasificación de suelo Acrisoles Háplico (Cutanico y Hiperdistrico) los factores fisicoquímicos de acidez, condición y contenido orgánico entre otros contribuyen en al alto o bajo índice de diversidad microbiana (procariota) de microorganismo no conocidos. Los estudios de diversidad microbiana basados en la secuenciación metagenómica del gen ARNr 16S han revelado que el suelo es un entorno particularmente rico en diversidad procariota con comunidades de microorganismo no cultivable que contribuyen a la riqueza de nutrientes y caracterización del suelo permitiendo la identificación de nuevos microorganismos procariotas de este modo la clasificación de cada suelo y su diversidad dependerá de factores fisicoquímico y practica culturales. En Panamá no existen aún estudios sobre la diversidad microbiológica asociada a la clasificación de suelos. Para iniciar este análisis se realizó la extracción de ADN siguiendo la secuenciación masiva del gen ARNr 16S de muestras de suelo de 4 sitios de la provincia de Chiriquí (Boca chica y puerto armuelle), obteniéndose 829,385 secuencias y 637,269 AVS.

Se observaron índices de diversidad elevados para las muestras todas las muestras con diferencias significativas en muestras de suelo respecto al tipo de suelo y pH. Los análisis taxonómicos para todas las muestras mostro bacterias mayor al 50%, como filo mayoritario en suelo a WS6, Actinobacteria, Proteobacteria con predominancia de SC72, Actinobacteria, Alphaproteobacteria. Y solo una muestra marco presencia del dominio Archaea (ACHD28. - Acrisoles Cutanico) con un porcentaje menos a 1%. el Crenarchaeota (especialmente MCG (Misceláneo Crenarchaeota con Predominancia de orden pGrfC26 termofilico. Estos hallazgos, constituyen un primer aporte al conocimiento de las diversidades bacterianas en suelos Acrisoles Háplico (Cutanico y hiperdistrico) de región en desarrollo intervenida y no intervenida y contribuirán al descubrimiento de nuevos microorganismos y prácticas agrícolas sustentables.

# **Abstract**

## **Abstract**

During the classification of Haplic Acrisols soil (Cutaneous and Hyperdistric) the physicochemical factors of acidity, condition and organic content, among others, contribute to the high or low index of microbial (prokaryotic) diversity of unknown microorganisms. Microbial diversity studies based on metagenomic sequencing of the 16S rRNA gene have revealed that the soil is a particularly rich environment in prokaryotic diversity with communities of non-cultivable microorganisms that contribute to the richness of nutrients and characterization of the soil allowing the identification of new microorganisms. prokaryotes in this way the classification of each soil and its diversity will depend on physicochemical factors and cultural practices. In Panama there are still no studies on the microbiological diversity associated with the classification of soils. To start this analysis, DNA extraction was carried out following the massive sequencing of the 16S rRNA gene of soil samples from 4 sites in the province of Chiriquí (Boca Chica and Puerto Armuelle), obtaining 829,385 sequences and 637,269 AVS.

High diversity indices were observed for all samples with significant differences in soil samples regarding soil type and pH. The taxonomic analyzes for all the samples showed bacteria greater than 50%, as the majority phylum in soil at WS6, Actinobacteria, Proteobacteria with a predominance of SC72, Actinobacteria, Alphaproteobacteria. And only one sample marked the presence of the Archaea domain (ACHD28. - Acrisols Cutanico) with a percentage less than 1%. the Crenarchaeota (especially MCG (Miscellaneous Crenarchaeota with a Predominance of the thermophilic pGrfC26 order). These findings constitute a first contribution to the knowledge of bacterial diversities in Haplico Acrisol soils (Cutanico and hyperdistrict) of an intervened and non-intervened developing region and will contribute to the discovery of new microorganisms and sustainable agricultural practices.

## Introducción

El suelo es considerado un espacio heterogéneo definido por sus propiedades físicas, químicas y biológicas (Reyes & Valery, 2007). La superficie suelo en condiciones naturales logra desarrollar un equilibrio dinámico entre sus diferentes propiedades, la cual genera las condiciones adecuadas para crear un microbioma y microbiota ampliamente diverso siendo transformadores y descomponedores de sustratos, entre otros. Todos estos organismos establecen relaciones en formas muy variadas y complejas (Nogales, 2005). El estudio de clasificación más recientes en Panamá fue el 2010 donde la estadística muestra 12 grupos de suelos y sus 23 Clasificadores distintos (Villarreal, et al., 2010). Panamá pertenece al hotspot (punto caliente) mesoamericano de biodiversidad (Suzart *et al.*, 2015), la cual lo hace atractivo para estudios de biodiversidad). Dentro de estos y objeto de este estudio, el suelo Acrisol tiene un clasificador principal llamado Haplico el cual se divide en dos clasificadores suplementarios: Hyperdystric y cutanic, los cuales están distribuidos en la provincia de Chiriquí (Panama). Este tipo de suelos se caracterizan en general por presentar elevada acidez y alta saturación de aluminio, lo que incrementa la actividad de los microorganismos encargados de la mineralización de la materia orgánica del suelo y, debido a ello, aumenta la disponibilidad de nitrógeno, fósforo, azufre y boro (Sadeghian, 2016), entre otros.

Es de gran importancia estudiar la abundancia y diversidad de microorganismos procariotas en suelo ya que esto nos lleva al descubrimiento de nuevos microorganismos, las comunidades presentes y las interacciones entre estos microorganismos. Además, el suelo constituye una de las fuentes más importantes de diversidad microbiana y por lo tanto un reservorio de genes novedales con potencial biotecnológico. En un gramo de suelo forestal se estima  $4 \times 10^7$  células procariotas (Richter, 1995). Anteriormente la caracterización de la diversidad microbiana estaba restringida a aquellos microorganismos cultivables, sin embargo, con el advenimiento de la metagenómica ha sido posible determinar las comunidades microbianas presentes en el suelo aún de aquellos microorganismos que no son cultivables.

La secuenciación metagenómica ofrece muchas ventajas al estudiar la diversidad microbiana, incluido el potencial de proporcionar información sobre la distribución ecológica de múltiples familias de genes simultáneamente. Los datos metagenómicos también permiten muestrear una amplia diversidad de genes a la vez, en lugar de centrarse en uno. Al respecto conviene decir que los datos metagenómicos pues puede proporcionar información sobre la ecología y la



evolución microbianas. Estudios previos de diversidad microbiana utilizando datos metagenómicos generalmente han cuantificado la estructura de ensamblajes microbianos agrupando secuencias metagenómicas en grupos taxonómica o funcionalmente similares basados en la similitud general de secuencia, o en genes marcadores únicos, tal como el gen 16S SSU-rRNA (el gen rRNA de la subunidad ribosomal pequeña 16S) (Woese & Fox, 1977, Daniel, 2005). En los últimos años ha habido avances significativos en el desarrollo estadístico de la diversidad filogenética para cuantificar la importancia relativa de procesos como la dispersión, la competencia y el filtrado ambiental en la configuración estructural de la comunidad microbiana (Woese & Fox, 1977).

En esta investigación se determinó la diversidad Procariota y la correlación de estas mismas con los factores fisicoquímico que permiten predecir la abundancia del índice de diversidad alfa, beta obtenida y la diferencia significativa cualitativa y cuantitativa en los mismos de dos tipos de suelos Acrisoles Háplico cutanico y Acrisoles Háplico hiperdistrico,

Se inicio el análisis de esa diversidad aplicando, métodos de metagenómicos (que incluye colecta y extracción de ADN, análisis fisicoquímico a escala de laboratorio a escala) y bioinformáticos con Qiime2 la cual proporciona las condiciones óptimas de empleo para la obtención resultados de calidad diferenciada para estudiar el microbiota de suelo, mediante la secuenciación de biblioteca de la región:

16S Amplicon PCR Forward Primer =

5'TCGTCCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG 16S

Amplicon PCR Reverse Primer =

5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATC

C del gen ARNr. Qiime2 permite determinar así posibles relaciones entre los índices de diversidad alfa y beta obtenida y los parámetros fisicoquímicos presentes en este tipo de suelo generando información no existente sobre las bacterias presentes en este tipo de suelos en Panamá. Esto nos permite generar información básica de gran utilidad para la caracterización de la diversidad microbiana en otros tipos de suelos. Además, con las muestras obtenidas se pueden confeccionar bibliotecas genómicas para la búsqueda de genes de interés entre los que se incluyen genes que codifican para antibióticos, enzimas de interés industrial como lipasas, celulasas entre otros. La selección de bibliotecas de suelos metagenómico, son especialmente con enfoques basados en la actividad metabólica y esto ha llevado a la identificación de varias

biomoléculas novedosas, que incluyen enzimas y antibióticos de importancia médica e industrial.

# Capítulo I

## Revisión Bibliográfica

## **Definición de suelo**

El suelo es considerado un espacio heterogéneo definido por sus propiedades físicas, químicas y sus diferentes atributos, lo que genera las condiciones adecuadas para una diversidad de organismos transformadores y descomponedores de sustratos. El suelo es una mezcla compleja de organismos vivos, materia orgánica, minerales, agua y aire. En general, se considera que el microbiota del suelo, conformada principalmente por bacterias y hongos, juega un papel importante en la fertilidad, reciclaje de nutrientes, evolución, estructura y conservación de este (Reyes & Alexis Valery, 2007).

El suelo se compone de partículas orgánicas, de materias vegetales y animales, descompuestas que provienen de plantas y animales vivos; partículas minerales tales como arena, arcilla, piedras o grava que, alguna vez, fueron parte de rocas mayores.

Según su textura, estructura y consistencia, los diversos tipos de suelos retienen mayor o menor cantidad de agua y aire. Todos los suelos están comprendidos en dos categorías principales: los suelos minerales y los suelos orgánicos. Los suelos minerales provienen de una roca madre denominada material de partida. Se forman con el tiempo al quebrarse el material de partida por diversos procesos físicos, químicos y biológicos ocasionados por el clima, el drenaje, la lixiviación, la erosión, la vegetación y los organismos vivos. Esto se denomina meteorización. Por ejemplo, las altas temperaturas del suelo rompen las piedras en fragmentos menores mediante el calentamiento y enfriamiento. El material de partida se reduce gradualmente a partículas; las superficies mayores entran en contacto con el agua y cambia la composición química de los minerales presentes. Las sustancias químicas solubles son arrastradas o se lixivian, hacia las capas más profundas del suelo, mientras los elementos menos solubles permanecen en las capas superiores del suelo. Continúa el proceso de meteorización (desintegración o ruptura en la roca, sin afectar a su composición química o mineralógica.) y con el tiempo se forman suelos minerales como los que conocemos hoy.

La formación de los suelos comienza con el proceso de meteorización originando horizontes que son capas que caracterizan a cada tipo de suelo. Al igual que existen muchos tipos y variaciones de suelo, también hay variaciones en las capas horizontales que son típicas de todos

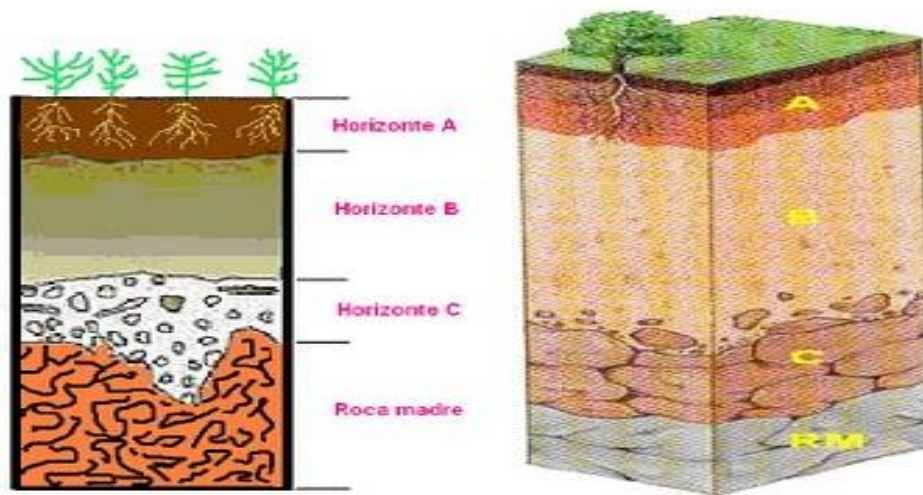
los suelos. Las capas de suelo tienden a variar de un lugar a otro en el número, grosor individual, color, y características físicas y químicas.

Las capas principales del suelo y el subsuelo de los suelos minerales) se subdividen en capas más delgadas llamadas horizontes dominantes que a su vez se pueden tener sub-horizontes.

Las propiedades físicas de los estratos del suelo, desde la superficie del terreno hasta una profundidad de unos 1.5 - 2 metros o, en ocasiones, a mayor profundidad, se ven afectadas por los cambios en el contenido de agua y la temperatura según la estación, y por diversos agentes biológicos como raíces, gusanos, insectos y bacterias. Lo que permite diferenciar los horizontes en el suelo

La parte superior del suelo mineral, el horizonte dominante A, está sometida a los efectos mecánicos de la meteorización y a la pérdida de algunos de sus elementos debido a la lixiviación. En la parte inferior del estrato del suelo, el horizonte dominante B, se precipitan y acumulan algunas de las sustancias lixiviadas de la parte superior. (Figura, 1)

Debajo del horizonte dominante B, la naturaleza del suelo se determina por el tipo de roca madre del que se formó, la manera en que se depositó y por hechos geológicos ulteriores.



**Figura 1.** Horizontes Del Suelo son las distintas capas que componen al suelo sobre el que habitan los seres vivos de la biosfera, es decir, a la capa superficial de la tierra. Algunos lugares tienen pocas de estas capas, mientras que otros tienen muchas, todas con colores, texturas y propiedades físicas y químicas diferentes. Horizonte A: zona de lixiviación, horizonte B: Zona de Acumulación, horizonte C: material parental. (Anon., Jorge L. Tamyó)

Este puede tener cualquier color, edad, ser muy poco profundo o profundo, y consiste principalmente en una mezcla estructurada de arena, limo y arcilla (inorgánicos), rocas y material orgánico (vivo y muerto) (Hartemink, 2016). Cada suelo está formado por procesos pedogénicos (consisten en una serie de reacciones y redistribuciones de materia, de tal forma que un material originario sufre una progresiva diferenciación de horizontes y modifica su grado de organización "obtiene estructura") que son bióticos y abióticos (Targulian & Krasilnikov, 2007).

### **Biodiversidad de Suelo**

El convenio sobre la diversidad biológica define la diversidad biológica como "la variabilidad entre los organismos vivos de todas las fuentes, incluidos, entre otros, los ecosistemas terrestres, marinos, otros ecosistemas acuáticos y los complejos ecológicos que forman parte; esto incluye la diversidad dentro de las especies, entre especies y de ecosistemas "(CDB, 1992, Artículo 2)

El concepto se usa convencionalmente en un sentido taxonómico y denota el número de especies distintas, pero puede extenderse para abarcar diversidad genética, fenotípica (expresada), funcional, estructural o trófica. Se acepta que la biomasa total bajo tierra iguala o potencialmente excede la biomasa aérea. Y aunque la biodiversidad en el suelo supera a la de otros sistemas terrestres en órdenes de magnitud, particularmente a escala microbiana, sigue estando notablemente infravalorada (FAO, et al., 2020)

Los suelos se consideran uno de los hábitats con mayor diversidad biológica de la tierra. Se ha estimado que 1 gramo de suelo contiene hasta mil millones de células bacterianas, que comprenden decenas de miles de taxones.

Los suelos son uno de los principales reservorios globales de biodiversidad, más del 40% de los organismos vivos en los ecosistemas terrestres están asociados durante su ciclo de vida directamente con los suelos (Decaëns et al., 2006). De hecho, los suelos son uno de los principales reservorios mundiales de biodiversidad (Bardgett y van der Putten, 2014; Carey, 2016). Este reservorio incluye bacterias y arqueas, hongos, protistas y muchos más eucariotas, como nematodos, ácaros oribátidos, ciempiés y milpiés, enquitreidos, tardígrados, colémbolos, hormigas, escarabajos terrestres y lombrices de tierra (Zhang, 2013; Stork, 2018; Coleman y Whittman), 2005). El suelo es un sistema complejo y heterogéneo, que comprende agregados organominerales de diferentes tamaños y componentes orgánicos, que crea hábitats para la biodiversidad en múltiples escalas espaciales; la diversidad en la composición del hábitat con

poros de diferentes tamaños llenos de aire y / o agua permite que un número increíble de taxones de diferentes tamaños y ecología lo habiten (Andre, Ducarme y Lebrun, 2002). Las comunidades de suelo son sistemas jerárquicos donde varios tipos de organismos habitan volúmenes de suelo críticamente diferentes. Esto incluye la película de agua de un micrómetro de espesor alrededor de las partículas del suelo que albergan organismos acuáticos como bacterias, protistas, nemátodos y tardígrados (la microfauna), el espacio poroso lleno de aire para los animales del suelo entre 100  $\mu\text{m}$  y 2 mm de ancho (la mesofauna), los puntos calientes de nutrientes y otros recursos alrededor de las raíces de las plantas para los microorganismos, y la macrofauna y megafauna que perciben el suelo como un todo en el que hacen pasajes que pueden penetrar todos los horizontes del suelo a través de volúmenes significativos de suelo (Pokarzhevskii et al. , 2003). La principal fuerza impulsora de la gran diversidad de animales del suelo es este fraccionamiento del tamaño corporal, pero también su diferenciación funcional. Una variedad de nichos ecológicos en el suelo, tanto en términos de tamaño como en la gama de recursos proporcionados, conduce a una diferenciación funcional significativa de los organismos del suelo.

Las comunidades bacterianas del suelo afectan directamente la estructura y funcionalidad del suelo a través de varios mecanismos. Como parte de su ciclo de vida, muchas bacterias del suelo secretan sustancias poliméricas extracelulares (EPS) que unen las partículas del suelo y mejoran la agregación del suelo. Esto mejora la porosidad del suelo, la capacidad de retención de agua, la aireación y la penetración de las raíces de las plantas (Costa et al., 2018). Las bacterias generalmente tienen una alta tasa de crecimiento con un tiempo de generación muy corto y compiten por los mismos recursos en el suelo (Madigan et al., 2015). En un suelo sano y sin perturbaciones, las bacterias no patógenas suelen superar a las patógenas (Lowenfels y Lewis, 2010). Además, las bacterias actinomicetos (por ejemplo, estreptomicetos) secretan antibióticos que matan o inhiben el crecimiento de patógenos vegetales, proporcionando un ambiente saludable para el crecimiento de las plantas (Bhatti et al., 2017).

Los suelos son un reservorio clave de la biodiversidad global, que abarca desde microorganismos hasta flora y fauna. Esta biodiversidad tiene un papel fundamental en el apoyo a las funciones del suelo y, por lo tanto, a los bienes y servicios del ecosistema asociados con los suelos. Por tanto, es necesario mantener la biodiversidad del suelo para salvaguardar estas funciones. La biodiversidad del suelo consiste en la variedad de vida subterránea, desde los

genes y las especies hasta las comunidades que forman, así como los complejos ecológicos a los que contribuyen y a los que pertenecen, desde los microhábitats del suelo hasta los paisajes. Estos constituyen diversas funciones que incluyen la producción de biomasa, almacenamiento, filtrado y transformación de nutrientes, sustancias y agua, reserva de biodiversidad, física y entorno cultural, fuente de materias primas, actuando como reserva de carbono, archivo del patrimonio geológico y cultural (McBratney, et al., 2014).

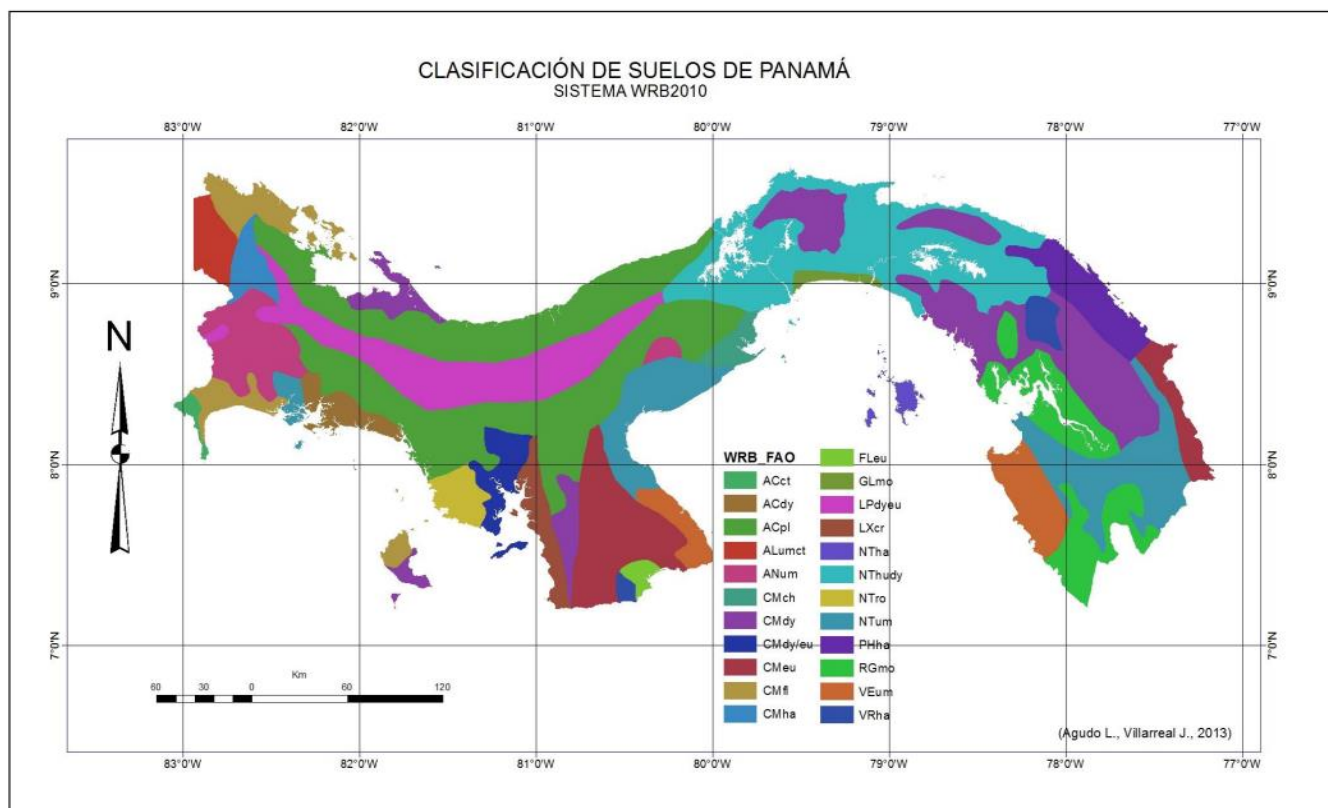
La Guía para la descripción de suelos (FAO, 2006) debe seguirse cuando se describe el suelo y sus características. Es posible hacer una clasificación preliminar en campo utilizando todas las propiedades y características observables o fácilmente medibles en el suelo y en el terreno asociado. Sin embargo, la clasificación final puede hacerse solo cuando se dispone de datos analíticos físicoquímicos (anexo1) (Van Reeuwijk, 2002, IUSS Working group WRB, 2014). La combinación descrita de horizontes, propiedades y materiales de diagnóstico se compara con la Guía WRB con el fin de asignar al suelo el Grupo de Suelo de Referencia (GSR) apropiado y que permite en adición designar los calificadores primarios y secundarios de los suelos según la Base referencial mundial del recurso suelo 2014 (IUSS Working group WRB, 2014).

### **Suelos de Panamá**

Panamá es una franja de tierra angosta orientada de oeste a este y es bañada en sus costas por el mar Caribe y el océano Pacífico, que se ubica en extremo del istmo centroamericano, entre los 7° 12' 07" y 9° 38' 46" de Latitud norte y entre los 77° 09' 24" y 80° 03' 07" de Longitud oeste. Limita al oeste con la República de Costa Rica y al este con la República de Colombia, al norte con el océano Atlántico (mar Caribe) y al sur con el Pacífico. La superficie total del país es de 75,517 kilómetros cuadrados. En algunas regiones de Panamá encontramos suelos fértiles como los de Volcán, Cerro Punta y Boquete. Existen algunos suelos no fértiles pero que con nutrientes y abonos los hacen aptos para la agricultura; estos los encontramos en Colón, Chorrera, Santiago. Suelo de origen volcánico: se localiza en las tierras altas de Chiriquí, son propios para la agricultura (Josefina Gaitán, 2021).



El mapa de clasificación de suelos de Panamá, la base de datos recopilada y el atlas de suelos de Latinoamérica y el Caribe, constituyen un valioso aporte para la planificación, uso y manejo del suelo, de acuerdo con sus propiedades fisicoquímicas en una determinada región del país que agrupa a los suelos en 12 grupos diferentes y 23 unidades ) y que permiten establecer la equivalencia de los diferentes órdenes, según el Soil Taxonomy, con el World Reference Base of Soil Resources (WRB (Villarreal, et al., 2010). (Figuras, 2) Tipo de suelo según su abreviatura. Códigos recomendados por (FAO, 2006 - IUSS Working group WRB, 2014). Anexo 4



**Figura 2.** 12 grupos de suelos y sus 23 Clasificadores distintos según (Villarreal, et al., 2010) en Panamá: Mapa de clasificación de suelos de Panamá utilizando el sistema WRB 2010. Los códigos de abreviaturas y colores representativos recomendados nos permiten establecer la equivalencia de los diferentes órdenes, según el Soil Taxonomy, con el World Reference Base of Soil Resources (WRB).

## Características de los suelos Acrisoles Haplico

Los acrisoles presentan un marcado horizonte de acumulación de arcilla y una baja saturación de bases (menor al 50%); los más abundantes son acrisoles órticos, férricos y gleicos. Haplic (ha) (del griego haplous, simple): que tiene una expresión típica de ciertos rasgos (típica en el sentido de que no hay una caracterización adicional o significativa); se usa sólo si no aplica ninguno de los calificadores previos.

La mayoría de los Acrisoles tienen una capa fina, marrón, horizonte superficial argico, particularmente en regiones con pronunciado horizonte seco; Los colores más oscuros se encuentran donde el anegamiento (periódico) retarda la mineralización de la materia orgánica del suelo. El subyacente El horizonte subsuperficial álbico tiene elementos de estructura débilmente desarrollados y puede incluso ser masivo; Normalmente es de blanquecino a amarillo y se superpone a un color más fuerte de amarillo a rojo argico. Estos suelos contienen pocos minerales resistentes en la interperie. Los contenidos de óxidos de Fe, Al y Ti son comparables a otros suelos como los Ferralsoles; el SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub> en proporciones menos de 2. Los suelos acrisoles bajo una cubierta forestal protectora son superficiales porosos. Si se tala el bosque, el valioso horizonte A se degrada y se pierde la cual provoca una corteza de superficie dura. La corteza permite una penetración insuficiente de agua durante las lluvias con una erosión devastadora de la superficie (baja estabilidad de la estructura) como una inevitable consecuencia (Calderón Sáenz, 2016). Muchos Acrisoles en posiciones bajas del paisaje muestran signos de saturación de agua periódica; suelos horizontes superficiales son casi negros mientras que los colores de la matriz son cercanos al blanco en el horizonte álbico eluvial.

La mayoría de los Acrisoles tienen una microestructura débil y una macroestructura masiva, especialmente en la superficie y suelo subterráneo poco profundo que se ha empobrecido de sesquióxidos. Enlace entre sesquióxidos y las arcillas de baja actividad cargadas negativamente son menos fuertes que en los Ferralsoles. En consecuencia, la proporción de agua la 'arcilla natural' dispersable sobre la 'arcilla total' es más alta que en Ferralsoles.

Los acrisoles tienen malas propiedades químicas. Los niveles de nutrientes de las plantas son bajos y la toxicidad del aluminio y la absorción tiene fuertes limitaciones. Como la actividad biológica es baja en Acrisoles, la regeneración natural, por ejemplo, del suelo superficial que fue degradado por operaciones mecánicas, es muy lento. (Madigan et al., 2015; Wang et al., 2017).

La preservación de la superficie del suelo con su materia orgánica es de suma importancia para la agricultura en Acrisoles. La limpieza mecánica de los bosques naturales con la extracción de raíces y relleno posterior de agujeros con material superficial puede producir tierra inadecuada para la agricultura, por altos niveles tóxicos de aluminio del antiguo subsuelo. Sistemas de cultivo adaptados tras una fertilización completa y generosa, así como un manejo cuidadoso son requeridos si se desea practicar una agricultura sedentaria (es decir anual o con décadas de descanso) sobre los Acrisoles.

Se recomienda la agroforestación como alternativa al cultivo itinerante para proteger el suelo y lograr rendimientos más altos sin requerir insumos caros. La agricultura de bajos insumos sobre Acrisoles no es muy rentable. Cultivos comerciales poco exigentes y ácido-tolerantes como la piña, el anacardo, el té y el caucho se pueden cultivar con cierto éxito. Áreas crecientes de Acrisoles están plantadas con palma de aceite (por ejemplo, en Malasia y Sumatra). Grandes áreas de Acrisoles están bajo bosque, desde la selva alta y densa hasta bosques abiertos. La mayor parte de las raíces de los árboles se concentran en el horizonte superficial de humus y sólo unas pocas raíces se extienden hacia abajo en el suelo subsuperficial. Los Acrisoles son adecuados para la producción de cultivos de riego y de secano sólo después de encalado y fertilización completa. La rotación de cultivos anuales con pasturas mejoradas mantiene el contenido de materia orgánica.

En América del Sur, Acrisoles también se encuentran debajo de la savana. Los Acrisoles son adecuados para la producción de cultivos de secano y regadío únicamente después del encalado y la fertilización completa. -La rotación de cultivos anuales con pastos mejorados mantiene el contenido de materia orgánica (Medina et al., 2008)

### **Actividad bacteriana en el suelo**

Las bacterias se encuentran en todos los entornos de la tierra, desde fuentes termales y respiraderos oceánicos de aguas profundas hasta la atmósfera y la nieve ártica. También colonizan en los cuerpos de otros organismos, como plantas, animales y humanos, como patógenos, simbioses o simplemente formando una comunidad comensal. Se estima que la tierra alberga aproximadamente  $2,5 \times 10^{30}$  células. El contenido colectivo de carbono de todas estas células bacterianas es comparable al de todas las plantas de la tierra, y su contenido total de nitrógeno y fósforo es mucho mayor que el de toda la vegetación (Madigan et al., 2015; Wang et al., 2017), lo que convierte a estos microorganismos en la principal fuente de nutrientes

indispensables para la vida. Desde una perspectiva funcional, las bacterias se clasifican en tres grupos principales: fotoautótrofas como las cianobacterias que utilizan el CO<sub>2</sub> atmosférico como fuente de carbono y fijan el CO<sub>2</sub> utilizando energía luminosa, produciendo compuestos orgánicos que pueden ser utilizados por otros organismos (similar a la fotosíntesis de las plantas); quimioautótrofos o quimiolitótrofos que utilizan el CO<sub>2</sub> atmosférico como fuente de carbono (similar a los fotoautótrofos), pero obtienen energía de la oxidación de compuestos inorgánicos como amoníaco, hierro y azufre, y utilizan esta energía para la fijación de CO<sub>2</sub> y la producción de compuestos orgánicos; y heterótrofos o quimioorganótrofos que utilizan materiales orgánicos como fuente de energía y de carbono (similar a los animales) (Madigan et al., 2015; Graham et al., 2016).

En los ecosistemas terrestres, la heterotrofia se considera el modo trófico dominante en el que las bacterias tienden a prevalecer en las capas basales de las redes tróficas (Steffan y Dharampal, 2018). Estos distintos grupos tróficos están estrechamente vinculados con plantas y animales a través de asociaciones simbióticas o parasitarias, o como descomponedores (Steffan et al., 2015). Por ejemplo, a escala local, algunos de los controles dominantes sobre la composición de la comunidad bacteriana y sus actividades son el tipo de planta y la calidad de la hojarasca (Myers et al., 2012; Sun et al., 2014; Jankowski, Schindler y Horner-Devine, 2014). Las bacterias, como los principales heterótrofos del suelo, desempeñan un papel clave en las transformaciones del carbono y el ciclo de los nutrientes, mejorando la fertilidad del suelo. También pueden regular la estructura del suelo y crear entornos de suelo saludables para proteger a las plantas de agentes patógenos y aumentar el rendimiento de los cultivos (Sylvia et al., 2005). Transformaciones de carbono: una gran parte de la biomasa vegetal (caída al suelo y no consumida por herbívoros) y otra biomasa viva del suelo es diseminada por detritívoros invertebrados y microorganismos a través de las redes tróficas (Moore y de Ruiter, 2012; Paul, 2016). Además de descomponer los detritos vegetales en el suelo, liberando elementos esenciales para el crecimiento de las plantas, muchas bacterias pueden transformar diferentes tipos de hidrocarburos saturados y aromáticos como el aceite y los químicos / pesticidas sintéticos (Brzeszcz y Kaszycki, 2018). Las bacterias del suelo juntas como consorcios sintróficos juegan un papel importante en la descomposición de los contaminantes ambientales y la circulación de compuestos recalcitrantes complejos a través de la red alimentaria (Madigan et al., 2015). Por lo tanto, tienen el potencial de ser utilizados para la biorremediación de los

ecosistemas contaminados (Dong y Lu, 2012). Ciclo de nutrientes: el flujo de energía y nutrientes dentro de un ecosistema se produce en una amplia gama de escalas y niveles tróficos, desde la superficie hasta el subsuelo, que definen la jerarquía de las redes tróficas (Steffan et al., 2015). La actividad en un nivel trófico podría influir en el proceso que ocurre en otros niveles tróficos (Wollrab et al., 2012). El correcto funcionamiento de las redes tróficas depende en gran medida del ciclo de los nutrientes, y las bacterias son los principales motores de dicha renovación de sustancias y ciclos biogeoquímicos en la Tierra (Graham et al., 2016). Mientras que los quimioorganótrofos como descomponedores se caracterizan por su capacidad para mineralizar los compuestos orgánicos y el ciclo del carbono, los quimio-trofos obtienen su energía de compuestos inorgánicos y son impulsores clave del ciclo de azufre, nitrógeno, hierro y otros elementos transformando los compuestos inorgánicos en formas utilizables por las plantas. (Ingham, 2009; Madigan et al., 2015). Las bacterias influyen enormemente en la fertilidad del suelo en diferentes tipos de ecosistemas, y su diversidad funcional es la clave para el mantenimiento de la biodiversidad del suelo, la estabilidad de las redes tróficas y el funcionamiento del ecosistema (Li et al., 2019; Graham et al., 2016).

El nitrógeno se considera a menudo como uno de los factores de crecimiento limitantes en el suelo para el establecimiento y desarrollo de plantas (Ågren et al., 2012). Las bacterias fijadoras de nitrógeno simbióticas y de vida libre (no simbióticas), que se encuentran en la mayoría de los tipos de suelos (no perturbados), pueden convertir nitrógeno a amoníaco, que los nitrificantes convierten en nitrato que es fácilmente asimilado por las plantas (Orr et al., 2011; Mus et al., 2016). Los entornos del suelo dominados por bacterias suelen tener un pH y un contenido de nitrógeno más altos que promueven el crecimiento de las plantas y estabilizan aún más la cohesión del suelo sobre el suelo a través de la vegetación establecida (Madigan et al., 2015).

En las redes complejas de las cadenas alimentarias naturales, las bacterias del suelo actúan como descomponedores, quimiolitotrofos, patógenos y simbioses se pueden encontrar en diferentes niveles tróficos (Steffan y Dharampal, 2018). Es probable que la diversidad trófica de las comunidades microbianas del suelo, como las bacterias, regule la identidad trófica de los organismos de nivel superior (Steffan y Dharampal, 2018). Dentro de las cadenas alimentarias, las huellas microbianas se encuentran en los componentes proteicos y lipídicos de la estructura celular de los consumidores de nivel superior (Larsen et al., 2009; Arthur et al., 2014). Como tales, las bacterias son componentes primarios ubicuos de todas las redes tróficas y tienen un

fuerte impacto en el equilibrio de las redes tróficas del suelo y la estabilidad del funcionamiento del ecosistema (Moore y de Ruiter, 2012; Crotty, 2011).

Las bacterias del suelo son los actores fundamentales en las redes tróficas (Steffan y Dharampal, 2018). Se están convirtiendo rápidamente en componentes integrales de los sistemas modelo que investigan los vínculos mecánicos entre la biodiversidad y el funcionamiento de los ecosistemas (Birtel et al., 2015). Sin embargo, los estudios de diversidad microbiana global utilizando técnicas moleculares independientes del cultivo sugieren la existencia de un grupo potencialmente extenso de especies bacterianas no clasificadas (Mora et al., 2011; Yarza et al., 2014; Louca et al., 2019). Estos microorganismos no descritos pueden tener una amplia gama de capacidades metabólicas que influyen en los servicios de los ecosistemas, e incluso pueden ser de particular interés para fines específicos de gestión industrial y ambiental (Birtel et al., 2015; Louca et al., 2019). Una mejor comprensión de cómo la pérdida de biodiversidad por actividades antropogénicas podría afectar la estabilidad de las redes tróficas y el funcionamiento de los ecosistemas requiere un conocimiento de las brechas actuales en la diversidad microbiana del suelo (Louca et al., 2019). También es un requisito previo para medir la posición trófica de un microbio dentro de su comunidad respectiva mientras está integrado dentro de la red alimentaria. El conocimiento de la diversidad trófica del microbioma y macrobioma del suelo nos permite evaluar de manera más integral la diversidad funcional y la productividad de los ecosistemas (Steffan et al., 2015).

### **Taxonomía Procariota**

La taxonomía incluye áreas principales como la identificación, la clasificación y la nomenclatura. En la década de los años 70, Carl Woese y colaboradores propusieron la clasificación de los organismos vivos en tres dominios: Archaea (incluyendo a las denominadas entonces como “arqueobacterias”), Bacteria (que agrupaba a las conocidas entonces como “eubacterias”) y Eukarya (que englobaba a todos los seres eucariotas), basándose en la utilización de la subunidad menor del ARN ribosómico 16S como marcador filogenético (Woese, 1987; Woese et al., 1990). Estos tres dominios se desarrollaron a partir de un hipotético ancestro común (Penny y Poole, 1999), que es la raíz del árbol universal de la vida.

Los organismos procariotas se encuadran dentro de los dominios Archaea y Bacteria, si bien se ha demostrado que los organismos que constituyen el dominio Archaea están más estrechamente relacionados con el dominio Eukarya, que con el dominio Bacteria (Madigan et al., 2015).

En sus inicios y durante mucho tiempo, la taxonomía bacteriana se basaba en el análisis fenotípico para la identificación y la clasificación de las bacterias en grupos o taxones. La clasificación es la organización de organismos en grupos en base a su parecido fenotípico o su parentesco evolutivo. La identificación, permite situar un organismo desconocido dentro de un taxón previamente descrito y la nomenclatura es el proceso de asignar nombre a los grupos taxonómicos. La nomenclatura es la única parte de la taxonomía que está reglamentada (Lapage et al., 1992; Brenner et al., 2005a; Madigan et al., 2009), mientras que la clasificación y la caracterización de los procariotas son áreas que no están formalmente reguladas y en las que han tenido lugar numerosos cambios en los últimos 50 años.

Las primeras definiciones de especie procariota estaban basadas en grupos monotéticos definidos por un único carácter fenotípico. Posteriormente, esta definición adquirió un carácter politético de modo que las especies fueron definidas por una combinación de caracteres fenotípicos independientes, co-variables y de igual peso (Sneath y Sokal, 1973; Goodfellow y col., 1997). Finalmente, gracias al desarrollo y aplicación de las técnicas moleculares y de secuenciación que permiten un análisis genético y genómico, se dio un nuevo enfoque al concepto de especie procariota (Konstantinidis y Tiedje, 2004; Oren, 2004; Stackebrandt, 2004; Staley, 2006). Una especie procariota se define conceptualmente como “un grupo monofilético y genómicamente coherente de organismos individuales que muestran un elevado grado de similitud global con respecto a muchas características independientes, y que es diagnosticable por unas propiedades fenotípicas discriminativas” (Rosselló-Móra y Amann, 2001;). Más recientemente, el avance en la taxonomía microbiana basada en 58 estudios genómicos está abriendo camino para crear un concepto de especie más universal y transparente; no obstante, todavía se encuentra en una etapa de transición hasta llegar a convertirse en un criterio sólido para la descripción de nuevas especies microbianas, que se fundamentará en una combinación entre un estudio polifásico de la especie a estudiar y un estudio genómico detallado (Vandamme y Peeters, 2014; Thompson et al., 2015; Al-saari et al., 2015). Llegados a este nivel, se pueden estimar las semejanzas genómicas completas entre cepas (inter- e intraespecífica), estimar el efecto que tiene la transferencia horizontal de genes (THG), establecer de manera inequívoca

las verdaderas relaciones evolutivas entre las cepas (filogenómica), dilucidar qué caracteres fenotípicos son análogos, etc (Oren et al., 2014; Thompson et al., 2015).

### **Diversidad de bacterias**

La diversidad (patrones de parentesco filogenético entre organismos en comunidades ecológicas) proporciona información importante sobre los mecanismos subyacentes al ensamblaje de la comunidad. Los estudios que miden la diversidad filogenética en comunidades microbianas han utilizado la subunidad pequeña del gen de rRNA (SSU-rRNA) para cuantificar las relaciones filogenéticas entre taxones microbianos (Kembel, et al., 2011) (Rusch, et al., 2007) (Venter J, 2004. La composición de una comunidad microbiana puede ser inferida a partir de la secuenciación de regiones del gen SSU-rRNA y de allí establecer posibles relaciones filogenéticas entre los microorganismos presentes en estas comunidades (Rush, et al. 2007; Venter, 2004).

### **Metagenómica**

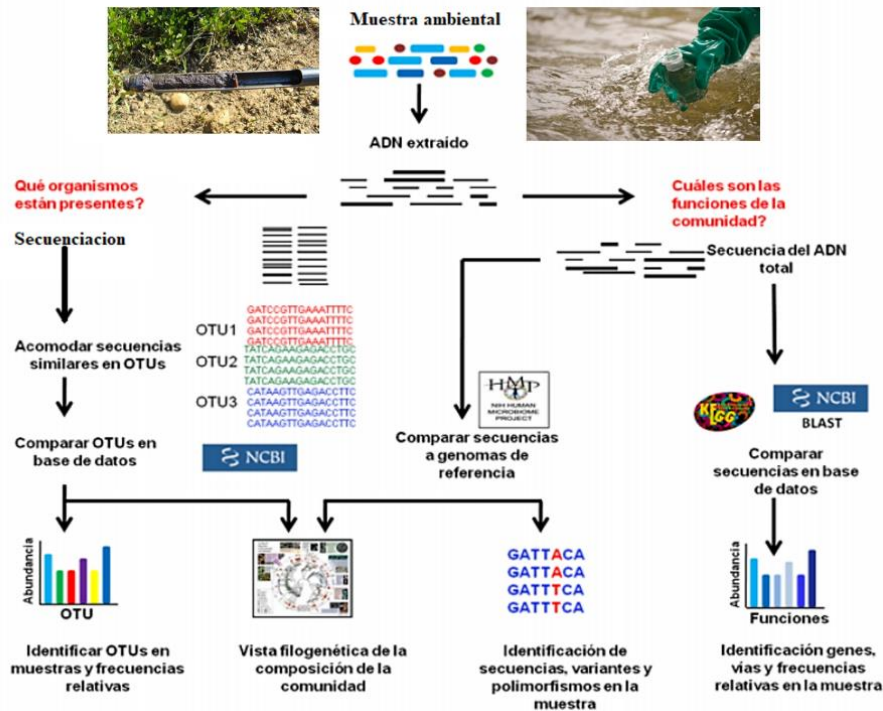
El término Metagenómica se descompone en " dos palabras "meta" y "genómica". La genómica se concibe como la colección de secuencias de ADN y meta la colección de las secuencias de todos los organismos presentes en la muestra (Segre &, 2001)

La metagenómica es una ciencia que surge como una rama de las ciencias genómicas, la cual se refiere al estudio del metagenoma de un nicho en particular (Handelsman, 2004; Riesenfeld et al., 2004). El metagenoma se puede definir como la totalidad de ADN obtenido de una muestra ambiental. Hasta el momento, se han investigado metagenomas de diversos ambientes, incluyendo suelos agrícolas y forestales, ecosistemas acuáticos y minas entre otros (Rondon et al., 2000; Wang et al., 2000; Lee et al., 2004; Tyson et al., 2004; Craig et al., 2009). En algunos casos, se han descubierto novedosos elementos genéticos que podrían tener aplicación en la industria (Rondon et al., 2000; Wang et al., 2009), mientras que, otros, han aportado aspectos novedosos de la ecología microbiana en un ecosistema en particular (Tyson et al., 2004).

En particular, se ha propuesto que entre 80-90% de los microorganismos que habitan el suelo son desconocidos (Alexander, 1977). Esto representa una limitante para descubrir el verdadero potencial genético de estos sistemas. Al estudiar el metagenoma de un ambiente en particular, es muy probable que este consista en gran parte de bacterias u otros organismos no cultivables; sólo conocemos una pequeña proporción de éstos, los cuales podemos reproducir en condiciones



de laboratorio (Handelsman 2004; Riesenfeld et al., 2004). De esta manera, el ADN metagenómico se puede clonar en vectores y expresar en diversos huéspedes procariotes, dependiendo de los objetivos del estudio (Rondon et al., 2000; Wang et al., 2000). Todo este material genético podría codificar nuevas o mejores actividades metabólicas. De la misma manera, se pueden emplear métodos de secuenciación masiva que generarán genomas completos de organismos no cultivables (Handelsman, 2004; Tyson et al., 2004).



**Figura 3.** Análisis metagenómico. por Raimondi, 2013. Las metagenómicas inician con la colecta de la muestra ya sea líquida o sólida siguiendo la extracción de ADN, el ADN extraído tiene dos funciones puede ser secuenciado para determinar que organismo esta presente en la muestra a través de OTUs o AVS y el ADN total secuenciado ser comparado en NCBI con otras secuencias en la base de datos para determinar la función de los genes identificados.

En los últimos años se ha registrado interés por la metagenómica lo cual ha contribuido juntamente con los avances tecnológicos a incrementar la cantidad de investigaciones en esta disciplina. Los avances en la tecnología de secuenciación, por ejemplo, han ampliado los enfoques y las preguntas que se pueden considerar con la metagenómica. La metagenómica ha sido utilizada para estudiar genes individuales (celulasas), vías biosintéticas (síntesis de

antibióticos), organismos (Archaea) y comunidades (microorganismos en la película de drenaje mina de ácido).

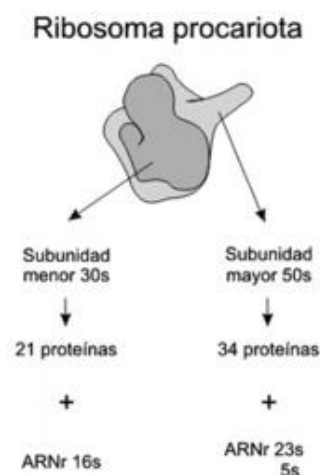
La metagenómica funcional ha contribuido a la identificación de nuevos antibióticos, enzimas hidrolíticas y enzimas involucradas en rutas de biosíntesis como también proteínas de membrana. Alternativamente, secuencias que portan regiones conservadas de genes asociados con la biosíntesis de antibióticos, como las policétido sintasas y péptido sintetasas, pueden identificarse por secuenciación pantallas basadas en que no requieren expresión génica heteróloga. (Daniel, 2005)

### Gen de ARNr 16S

El gen que codifica para el ARN ribosomal 16S genera un ARN que forma parte de la subunidad 30S del ribosoma procariota. La "S" en 16S es el coeficiente de sedimentación, es decir, un índice que refleja el tamaño de una partícula con base a la tasa de sedimentación en un campo centrífugo. Cuanto mayor sea el valor, mayor será la molécula. El gen 16S ARNr altamente conservado, pero incluye regiones altamente variables muy útiles para la identificación bacteriana-(Anon., 2015).

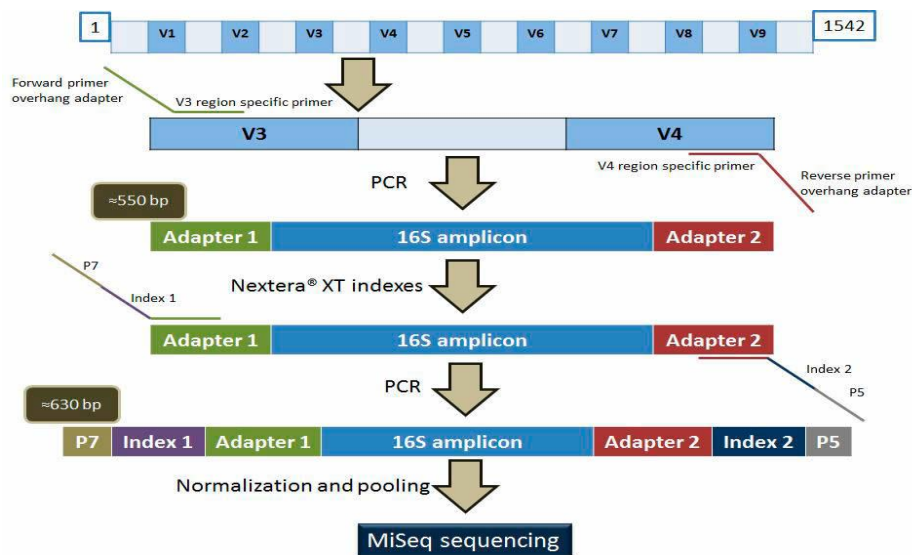
Las regiones conservadas son compartidas por todas las bacterias mientras que las regiones variables difieren entre bacterias con especificidad a nivel género y especie. Por lo tanto, es posible diseñar cebadores universales a partir de las regiones conservadas y específicos a partir de las regiones variables.

El gen que codifica para el ARN ribosomal 16S es de 1500 bp y contiene aproximadamente 50 dominios funcionales.



**Figura 4.** ARNribosomal 16S. Por María del Rosario Rodicio,2004

La secuenciación de segunda generación de amplicones del gen 16S r RNA ofrece varias ventajas incluyendo la factibilidad de obtener de manera masiva la secuencia de una gran cantidad de microorganismos; en comparación con la secuenciación capilar o los enfoques basados en PCR. La secuenciación de segunda generación (NGS) del gen 16S rRNA e ITS permite la identificación cepas que no se pueden encontrar utilizando métodos tradicionales. A diferencia de la secuenciación capilar o los enfoques basados en PCR, la secuenciación de próxima generación es un método libre de cultivo que permite el análisis de toda la comunidad microbiana dentro de una muestra. El análisis de ITS con NGS permite una rápida identificación de hongos para ayudar a avanzar en nuestra comprensión del microbioma. Además, NGS ofrece la capacidad de combinar múltiples muestras en una secuenciación. El diagrama de flujo presentado en la figura 5 resume las etapas para la construcción de las bibliotecas de rRNA 16S para la secuenciación de segunda generación. La región amplificada comprende a las regiones variables V3 y V4 y la región conservada que se encuentra entre ellas. Una primera PCR se realiza con los primers específicos V3-V4 que contienen los adaptadores para los índices Nextera. En una segunda PCR, se añaden los índices para introducir una combinación única para cada muestra y las secuencias P5 y P7 (que hibridan con los adaptadores de la celda de flujo del sistema MiSeq). Tras normalizar la cantidad de cada muestra que se añade al pool final, las librerías están listas para ser secuenciadas mediante MiSeq.



**Figura 5.** Esquema del proceso seguido para la creación de bibliotecas mediante amplificación del gen del ARNr 16S.

## **Análisis Estadístico Qiime 2**

En este estudio se utilizó QIIME2 (Quantitative Insights Into Microbial Ecology, <https://qiime2.org>) para el análisis bioinformático, ya que se trata de una plataforma de código abierto escrito en Python y ejecutado en Linux, con herramientas interactivas ampliamente utilizada para el análisis de secuencias del gen 16S y 18S obtenidas en plataformas como Illumina y otras.

En comparación con otras herramientas, QIIME2 es fácil de instalar, se puede ejecutar en máquinas virtuales como VirtualBox, es reproducible, sus resultados pueden ser visualizados de forma virtual, admite características personalizadas y se une al proceso de análisis. QIIME 2 permite a los investigadores comenzar un análisis con datos de secuencias de ADN sin procesar y terminar con cifras de calidad de publicación y resultados estadísticos. Qiime2 proporciona una herramienta, Qiime view(<https://view.qiime2.org/>) que permite visualizar los artefactos tipo qzv y que a su vez permite ver la procedencia de este. Los estudios estadísticos que permite realizar qiime2 incluyen la asignación taxonómica, reconstrucción filogenética, la alfa diversidad, beta diversidad que permiten relacionar con los metadatos de la muestra. Los metadatos pueden incluir detalles técnicos, como los códigos de barras de ADN que se usaron para cada muestra en una ejecución de secuenciación multiplexada, descripciones de las muestras, parámetros fisicoquímicos entre otros.

### **Diversidad Alfa**

a-diversity se refiere a la diversidad dentro de una muestra tal como muestra de sólida o líquida (Gilbert & Lynch, 2019). Hay tres índices de diversidad a que se usan a menudo en medicina. El índice Chao 1, una métrica de riqueza estima el número total de especies en una muestra. Tiene en cuenta los tres factores siguientes: el número de especies, el número de taxones únicos y el número de taxones dobles. Esto significa que no puede reflejar la abundancia del microbiota. El índice de Shannon-Wiener combina riqueza y uniformidad. Da más peso a las especies raras, lo que significa que es mayor cuando aumenta el número de especies raras. Su valor generalmente no supera 5,0; cuanto más alto es su valor, más abundante es la diversidad a. El índice de Simpson también combina riqueza con uniformidad. Sin embargo, a diferencia del índice de Shannon-Wiener, pone más énfasis en las especies comunes. Su valor oscila entre 0 y

casi 1; cuanto más alto es su valor, más abundante es la diversidad  $\alpha$ . En los índices anteriores, la riqueza se refiere al número total de especies en una muestra, mientras que la abundancia se refiere a los recuentos de lectura sin procesar de una especie. Tenga en cuenta que la abundancia relativa se utiliza cuando los recuentos de lectura sin procesar se normalizan o se convierten en porcentajes.

En este estudio usamos el índice de Shannon para el análisis de la alfa diversidad. Este índice se representa como  $H$  y se expresa con un número positivo, que en la mayoría de los ecosistemas naturales varía entre 0,5 y 5, aunque su valor normal está entre 2 y 3. Es uno de los índices más empleados para cuantificar la biodiversidad específica. El índice tiene en cuenta la cantidad de especies que existen en una muestra (riqueza) y la cantidad relativa de individuos que hay para cada una de las especies (abundancia).

Para poder ver si nuestros resultados eran significativos, a nivel de alfa diversidad se puede realizar una prueba de comparaciones por parejas. Se utiliza la prueba de Kruskal-Wallis como enfoque no paramétrico, para poder aplicar un análisis de las varianzas (ANOVA).

Finalmente, representamos los valores medios de abundancia según el índice de Shannon ( $H$ ) con la ayuda de un diagrama.

## **Diversidad Beta**

La diversidad beta mide las diferencias en la composición o estructura de las comunidades microbianas entre muestras. Estas diferencias se pueden medir de forma cuantitativa o de forma cualitativa.

Las métricas de diversidad beta más utilizadas son la disimilitud de Bray-Curtis y el análisis UniFrac no ponderado. La métrica UniFrac es una métrica filogenética empleada con frecuencia para comparar la historia evolutiva de las diferentes comunidades microbianas. El hecho de ser no ponderada significa que la métrica usa solo la riqueza de las características y no la abundancia. La distancia de BrayCurtis a partir de una matriz de distancias calcula cómo de diferentes son las muestras entre sí. Esta métrica opera de manera independiente a los datos de los árboles filogenéticos, mientras que UniFrac es una métrica filogenética.

La disimilitud de Bray-Curtis es una medida estadística utilizada para cuantificar la disimilitud de composición entre dos muestras o grupos. Su valor varía de 0 a 1, donde 0 significa que las dos muestras o grupos comparten todas las especies y 1 significa que no comparten ninguna.

Además, da más peso a las especies comunes. Tenga en cuenta que Bray-Curtis no es una medida de distancia real, por lo que el término "disimilitud de Bray-Curtis" es más apropiado que "distancia de Bray-Curtis".

La distancia UniFrac, que puede ponderarse o no ponderarse, estima las diferencias entre muestras o grupos en función de la distancia filogenética. La distancia UniFrac no ponderada tiene en cuenta la presencia y ausencia de taxones. Es sensible para detectar cambios de riqueza en especies raras, pero ignora la abundancia información en el cómputo. La distancia UniFrac ponderada incorpora la información de abundancia y reduce la contribución de especies raras.

El análisis de coordenadas principales se utiliza para observar la similitud o disimilitud entre muestras. Cuando más similares sean las muestras respecto a sus características, más cerca entre sí estarán en la gráfica de ordenación de PCoA.

Creamos el objeto `trans_beta` para invocar la matriz de distancia calculada a partir del índice de disimilitud de Bray-Curtis. Ordenamos la matriz de distancia de Bray-Curtis por el método PCoA y finalmente graficamos los resultados (véase Figura 18 en el apartado de Resultados). También representamos en un diagrama de caja y bigotes los valores medios de la matriz de distancia según la disimilitud de Bray-Curtis en función del grupo experimental (véase Figura 19 en el apartado de Resultados).

Para ver la significancia de la beta diversidad se realizó un PERMANOVA. El PERMANOVA nos permitió medir si los grupos de muestras eran significativamente diferentes entre sí mediante una prueba estadística centrada en las permutaciones<sup>45</sup>, basada en el método Anderson<sup>46</sup> y la función `adonis` del paquete `vegan` de R<sup>47</sup>. La significancia probabilística se estimó usando 999 permutaciones. Este enfoque no requirió ningún supuesto de distribución. Se efectuó un análisis basado en permutaciones del grupo, para ver la significancia estadística del factor grupo, otro análisis entre grupos para ver si alguna de las comparaciones era significativa y, por último, un análisis donde se añadió al modelo el factor tanque para ver la importancia de esta variable en la variabilidad de las muestras.

**La PERMANOVA** La diversidad  $\beta$  de diferentes comunidades se puede comparar utilizando varios métodos o modelos estadísticos, como PERMANOVA, la prueba de Mantel, el análisis de similitud y los procedimientos de permutación de respuestas múltiples. El PERMANOVA es el más popular y considerado como el más poderoso. Se implementa a través de la función

adonis en el paquete R vegan. Se pueden procesar cuatro métricas de disimilitud o distancia en el paquete vegano: la disimilitud de Bray Curtis, la distancia Jaccard y las distancias UniFrac ponderadas y no ponderadas. Si el valor P de la prueba de permutación es menor que 0,05, indica que la diversidad  $\beta$  entre diferentes comunidades es estadísticamente significativa. Otro resultado de la prueba es  $R^2$ , que indica qué parte de la varianza total puede explicarse mediante el factor de agrupación

# **Objetivos**



## **I. Objetivos:**

### A. Objetivo General:

- Determinar posibles relaciones entre factores físicos–químicos del suelo Acrisoles Háplico y la diversidad bacteriana mediante la secuenciación de segunda generación de regiones variables (V3, V4) del gen ARNr 16S.

### B. Objetivos Específicos:

- Determinar los ASV (amplicon-sequence variants) de muestras de suelo Acrisoles Háplico para establecer la diversidad bacteriana de estos suelos.
- Determinar la diversidad alfa y beta de las muestras de suelo.
- Establecer la relación de la diversidad bacteriana con los parámetros fisicoquímicos a través de los índices de diversidad de Faith, uniformidad de Pielou para la diversidad alfa y la métrica de Unifrac para el cálculo de la diversidad beta.
- Utilizar el método de Kruskal- Wallis para realizar las comparaciones por muestras y ver si existen diferencias significativas ( $p$ -value  $<0.05$ ) de alfa diversidad y Beta diversidad.

## **Hipótesis**

Utilizando la secuenciación metagenómica para secuenciar las secuencias diana del ARN ribosómico 16S como método de investigación nos permitirá determinar si en la diversidad microbiana en suelo Acrisoles Háplico de Panamá influyen los factores fisicoquímicos como elevada acidez y la presencia de aluminio.

# **Capítulo II**

## Metodología

## Metodología Experimental

### Sitios de Colecta

Se colectaron muestras de suelos Acrisoles Háplicos en cuatro sitios de la Provincia de Chiriquí, dos intervenidos y dos no intervenidos. Los sitios de colecta y las coordenadas se presentan en la tabla 2. El primer sitio ubicado al Sur de Puerto Armuelles representa un suelo no intervenido y un área compuesta por más de 20 km de área boscosa. El segundo sitio Petro-Terminal Panamá-Puerto Armuelles, suelo intervenido, área con varios tanques de reserva de petróleo (Figura 6). Tercer sitio Horconcitos-Chiriquí, suelo intervenido. El cuarto sitio Vía hacia Boca Chica-Chiriquí suelo no intervenido (Figura 7)

Tabla 1. Sitos de colecta de suelos Acrisoles Háplico Cutánico y Acrisoles Háplico Hyperdístico

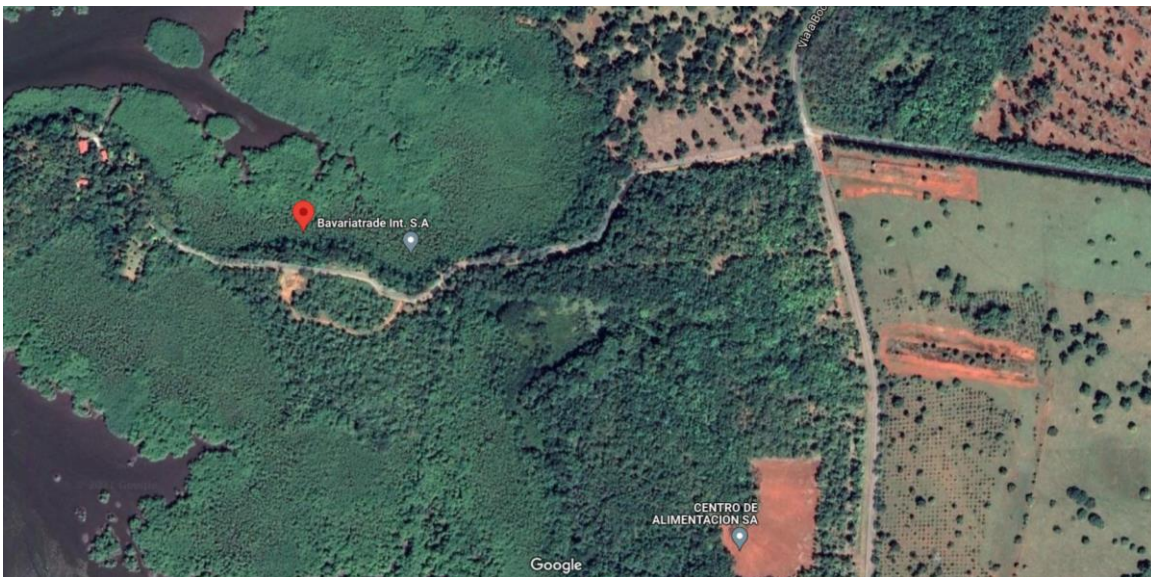
MUESTRA	COORDENADAS GEOGRÁFICAS	TIPO DE SUELO (WRB)	LUGAR	CARACTERÍSTICA	FECHA DE COLECTA
ACHD28	8.227170, -82.889004	Achd:HAPLIC ACRISOLS CUTANIC	Sur de Puerto Armuelles	no intervenido	09/03/2019
ACHD29	8.208398, -82.877976	Achd:HAPLIC ACRISOLS CUTANIC	Petro-Terminal Panamá-Puerto Armuelles	Intervenido	09/03/2019
ACHD30	8.316704, -82.151697	Achd:HAPLIC ACRISOLS hyperdistic	Horconcitos-Chiriquí	Intervenido	09/04/2019
ACHD31	8.266635, -82.190089	Achd:HAPLIC ACRISOLS hyperdistic	Vía hacia Boca Chica-Chiriquí	no intervenido	09/04/2019



**Figura 6.** Vista cercana - Chiriquí, Petro-Terminal Panamá-Puerto Armuelles.

Coordenadas 8.208398, -82.877976

Por: Google Maps



**Figura 7.** Vía hacia Boca Chica-Chiriquí. Coordenadas 8.266635, -82.190089

Por: Google Maps

## Colecta De Muestras De Suelo

Los sitios de muestreo fueron inicialmente seleccionados con la ayuda del mapa de clasificación de suelos para Panamá (WRB 2010) y Google Maps. Una vez localizado el sitio de muestreo procedimos a determinar la localización. Las coordenadas del sitio de colecta fueron anotadas en la bolsa ziploc previamente rotulada con el número de muestra, tipo de suelo y localidad.

Las muestras de suelo colocadas en bolsas ziploc fueron utilizadas para los análisis fisicoquímicos y se colectaron aproximadamente dos libras de suelo de cada muestra.

Se rotuló también un tubo eppendorf de 5 ml estéril con el número de muestra. Las muestras colectadas y transferidas a los tubos eppendorf fueron utilizadas para la extracción del ADN y se colectó suficiente cantidad de suelo para llenar el tubo.

Con la finalidad de minimizar la contaminación, el equipo utilizado en la colecta fue lavado con agua y desinfectado con alcohol al 70% antes y después de la colecta de cada muestra. Durante la colecta y la manipulación de las muestras se utilizaron guantes desechables estériles de látex. Las muestras para la extracción de ADN fueron colectadas con un muestreador de suelo (Figura 16 colocar foto de nuestro muestreador) mientras que las muestras para el análisis fisicoquímico fueron colectadas con una pala (figura X). La muestra para la extracción de ADN se colecta primero. Se introdujo el muestreador de suelo en sitio de colecta previa remoción de la vegetación u hojarasca en caso de que estuvieran presentes. El muestreador se introdujo a una profundidad de aproximadamente 10 cm. Luego de sacar el muestreador se procedió a colectar de la columna de suelo contenida en el interior del muestreador y con la ayuda de una espátula el suelo de la región comprendida entre 7.5 a 10 cm de la columna de suelo. Esta región corresponde a la región comprendida entre la superficie y una profundidad de 2.5 cm; profundidad a la que se reporta una mayor cantidad de microorganismos del suelo.

Una vez colectadas las muestras para la extracción de ADN se procedió a colectar las muestras para el análisis físicoquímico en el cuadrante que tiene como centro el sitio en el cual se introdujo el muestreador a una profundidad de 2.5 cm. (<https://www.youtube.com/watch?v=7KhRALD50Tc>). Las muestras para la extracción de ADN fueron colocadas en una hielera, transportadas al laboratorio y colocadas a 4 ° C. Las muestras para el análisis fisicoquímico, una vez en el laboratorio se vaciaron en bandejas de plásticos planas y se dejaron secar por 2 o 3 días. Luego entregadas al laboratorio de análisis de suelo de la Facultad de Ciencias Agropecuarias de la Universidad de Panamá.



Figura 1. Extracción de muestra de suelo utilizando el tubo de extracción de suelo.

### **Extracción de ADN**

La extracción de ADN se realizó con el kit DNeasy PowerSoil ProKit (<https://www.qiagen.com/us/resources/resource/detail?id=86f0cbb2-4c25-4533-893b-4811d2e6e6d0&lang=en>), siguiendo las instrucciones del fabricante (Qiagen, Inc). La cuantificación y calidad del ADN se determinó utilizando el espectrofotómetro Eppendorf, Modelo BioSpectrometer Basic, 230V, 50-60 Hz (6135000009) a longitudes de onda de 260 y 280 nm. La relación de las absorbancias 260/280 fue utilizada como criterio de la pureza del ADN. Una relación 260/280 entre 1.8-2.0 es considerada como ADN de buena calidad (Sambrook y Russell, 2001).

### **Electroforesis de ADN**

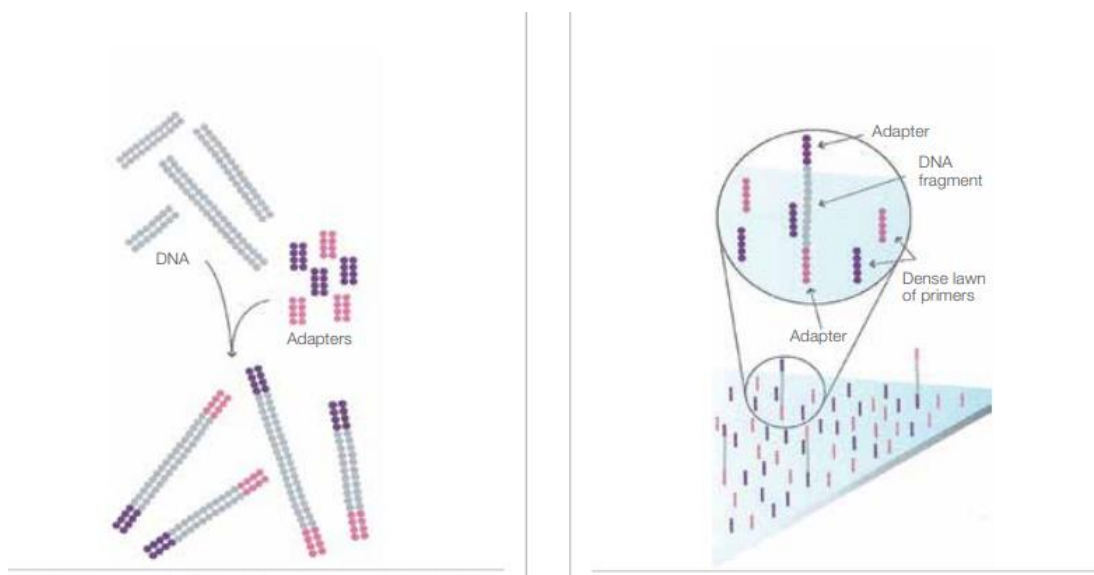
Se preparó un gel de agarosa a 0.8%, en TAE1X con 0.5µg/ml de Gel Red. Se cargaron 3.0 µl de marcador de peso molecular Lambda Hind III y 3.0 µl del ADN de cada muestra. La electroforesis se corrió a 90 V durante 45 minutos y el ADN fue visualizado en un documentador de geles (marca). La presencia de ADN de alto peso molecular es indicativa de buena calidad



mientras que la presencia de una mancha (smear) se asocia a ADN degradado. La obtención de ADN de buena calidad es requerida para la construcción de la biblioteca.

## SECUENCIACIÓN

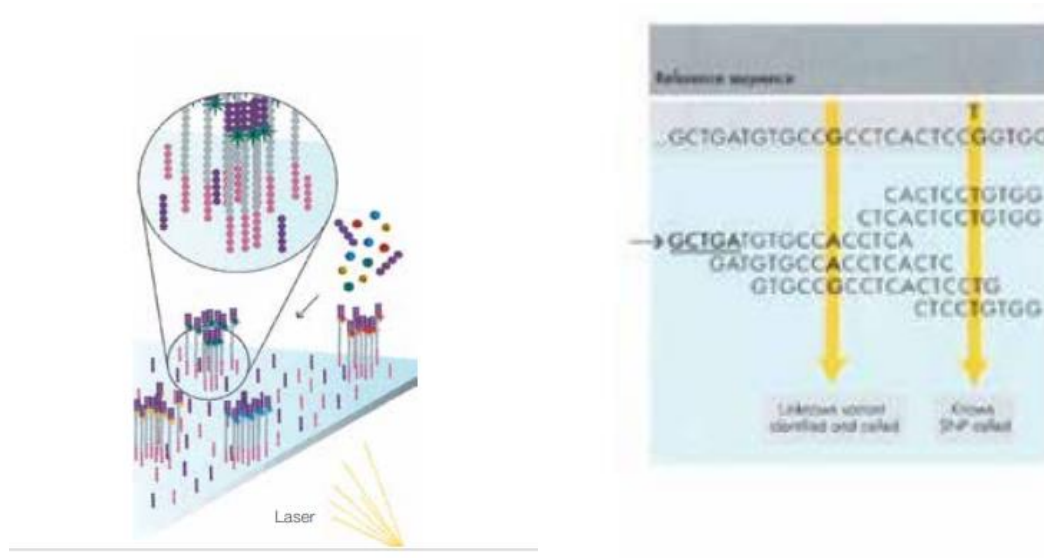
El ADN total extraído de las muestras de suelo (ACHD28, ACHD29, ACHD30 y ACHD31) fué enviado a la empresa NOVOGENE UC Davis Sequencing Center (<https://en.novogene.com>, California, Estados Unidos) para la construcción de la biblioteca metagenómica y la secuenciación de la biblioteca de cada muestra de suelo.



**Figura 8.** Secuencias de ADN unidos al adapter y primer

La biblioteca metagenómica se construyó amplificando la región variable V3-V4 del gen ribosomal 16S con los cebadores degenerados 341F (5'CCTAYGGGRBGCASCAG3') y 806R (5'GGACTACNNGGGTATCTAAT3) que amplifican una región de 470 bp de este gen. En adición a la secuencia específica para la amplificación de esta región cada primer posee una secuencia que actúa como adaptador (adapter) para unir los amplicones a la plataforma Illumina, una secuencia que actúa como código de barra (Golay barcode) que se utiliza para identificar cada muestra, una secuencia que previene la formación de estructuras secundarias (hairpin) y una secuencia usualmente de 2 nucleótidos denominada "linker". La biblioteca fue secuenciada utilizando el método de secuenciación por síntesis (SBS) con la plataforma Illumina NovaSeq 6000 y la estrategia PE250 (Paired-End) que consiste en secuenciar en direcciones opuestas dos regiones contiguas de 250 bp que se sobreponen por 26 nt en la región 3' de la secuencia

generada (Figura 9). Las etapas que conllevan a la construcción y secuenciación de la biblioteca se muestran en la figura 8.



**Figura 9.** Procesamiento de metagenómica de amplicones para muestra de ADN y obtención final de data.

### Procesamiento de las secuencias

Los datos sin procesar o crudos (“raw data”) que se obtienen inicialmente a partir de la biblioteca incluyen secuencias que no son útiles ni confiables y que constituyen lo que se denomina secuencias o datos sucios (“dirty data”). Las secuencias fueron combinados y luego filtradas para obtener secuencias limpias que permitan generar resultados más precisos y confiables.

Las secuencias obtenidas fueron procesadas con FASTQ Processor NOVOGENE UC Davis Sequencing Center (California, Estados Unidos). Posteriormente, las secuencias fueron analizadas con QIIME2. Eventualmente, luego de clasificar las secuencias de cada muestra; las secuencias de los códigos de barras, de los cebadores y de los adaptadores se eliminaron usando esta herramienta.



## **Análisis bioinformático de datos**

La Compañía Novogene envió las secuencias en formato FASTQ que fueron procesadas con líneas de comando del programa QIIME 2 (versión 2021.2) y siguiendo el protocolo de Caporaso et al. (2011).

- El procedimiento señalado en Qiime2 según la versión 2021.2 son las siguientes:
- Instalación de la máquina virtual y de QIIME2.
- Descarga de los datos.
- Importación de los datos y demultiplexación de las muestras.
- Filtraje de las muestras.
- Control de calidad de las secuencias con DADA2.
- Construcción de la tabla de características y generación de los gráficos de rarefacción.
- Clasificación taxonómica y generación de un árbol filogenético.
- Análisis de alfa y beta diversidad.
- Análisis de la composición taxonómica y abundancia diferencial.
- Interpretación y revisión de los resultados.

Una breve descripción del procedimiento bioinformático consistió en importar las secuencias crudas (multiplexed sequences) con el comando `qiime tools import`. Luego las secuencias fueron separadas en archivos para cada muestra (demultiplexación) con la línea de comando `qiime cutadapt demux-single` \ utilizando el archivo `sample-metadata` que contiene la información de cada secuencia del código de barra (barcode) respectivo. Posteriormente se procedió a eliminar de las secuencias demultiplexadas el código de barra utilizando la línea de comando `qiime cutadapt trim-single` teniendo como resultado el producto (output) : `trimmed-seqs.qza` que puede ser visualizado con la línea de comando “`qiime demux summarize`” que tiene el producto `trimmed-seqs.qzv` para visualizarlo con una tabla interactiva con Qiime2 `view` y que permite determinar el rango de secuencias que deben ser removidas; es decir, aquellas con un quality score menor de 20. Una vez establecido el largo de las secuencias útiles se procede a eliminar este ruido mediante el método o algoritmo DADA2 (Callahan et al., 2016). El resultado de DADA2 es la tabla de característica (feature table) que da información sobre cuántas secuencias están asociadas con cada muestra y que resume los datos filtrados o eliminados y que sirve de base para los análisis de diversidad. Estos análisis requieren la generación de un árbol filogenético el cual se hizo con la línea de comando `qiime phylogeny align-to-tree-mafft-`

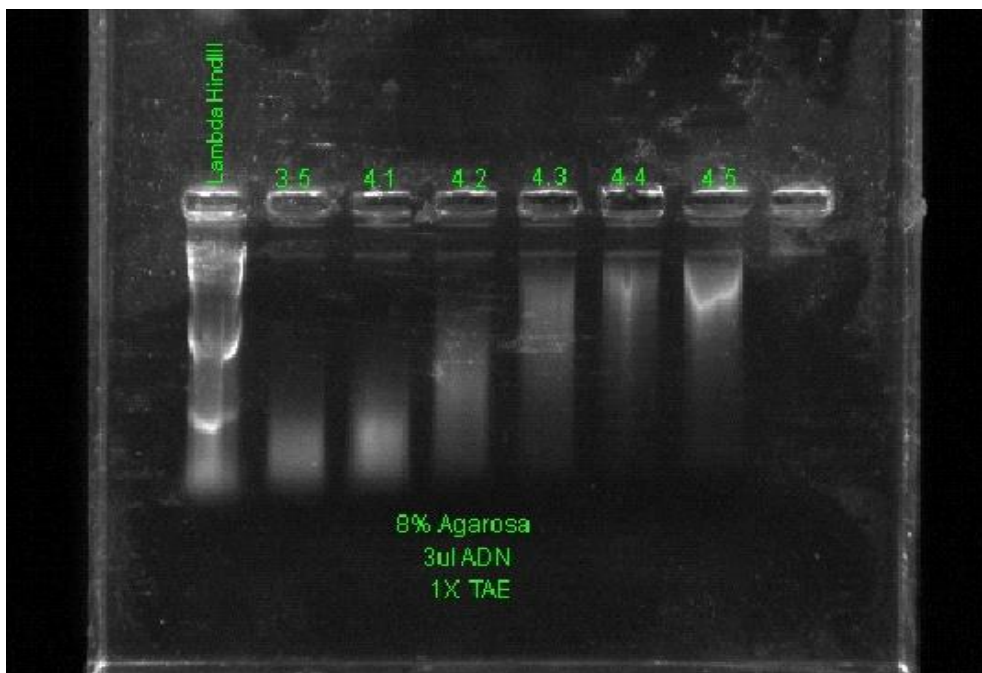
fasttree. Este árbol filogenético sirvió de base para el análisis de la diversidad alfa y beta. Para determinar la significancia de la beta diversidad se realizó un PERMANOVA. El PERMANOVA nos permitió medir si los grupos de muestras eran significativamente diferentes entre sí. El análisis taxonómico se usó una combinación de artefactos de QIIME2 y algunas funciones de R con el fin de obtener los porcentajes de abundancia relativa. A partir de la tabla de características de QIIME2, agrupamos las muestras en función de la columna categórica de interés grupo e indicamos que nos calculara los valores medios de abundancia relativa de las muestras para cada grupo. Luego se generaron las pruebas de abundancia diferencial para encontrar taxones significativos que evidencien diferencias entre comunidades bacterianas.

# **Capítulo III**

## Resultados y Discusión

## Resultados

Se logró la extracción eficiente de ADN de las cuatro muestras de suelo utilizadas en este estudio. En la electroforesis se observa ADN de alto peso molecular en los carriles 2, 3,4,5 correspondientes a las muestras ACHD28.1, ACHD29.1, ACHD30.1, ACHD31.1 (Figura 10). La relación de las absorbancias a 26/280 de cada una de las muestras generó valores entre 1.8 a 2.0 indicativo de un alto grado de pureza del ADN y las cantidades de ADN en cada muestra fueron mayores de 200 ng en un volumen de 20  $\mu$ l, mínimo requerido por la compañía para la construcción de la biblioteca.



**Figura 10.** Gel de agarosa con las extracciones de DNA de las muestras de suelo. En el primer carril se observa el marcador de peso molecular de LambdaHindIII (Thermo Fishe, Madrid, Thermo Scientific™ Marcador de ADN Lambda/HindIII: Productos bioquímicos y reactivos Ciencias de la vida | Fisher Scientific). El segundo carril corresponde a las muestras ACHD28.1, el tercero a ACHD29.1, el cuarto a ACHD30.1 y el quinto a ACHD31.1; esta electroforesis en gel determina el peso molecular adecuado de las muestras para una eficaz secuenciación.

## Datos crudos de secuenciación

La secuenciación de ADN se obtuvo los siguientes datos crudos sin procesamiento (Tabla 2) por Qiime2.

Tabla 2. Resultados iniciales de la secuenciación de la biblioteca

Sample Name	Raw PE(#)	Raw Tags(#)	Clean Tags(#)	Effective Tags(#)	Base(nt)	AvgLen(nt)	Q20	Q30	GC%	Effective%
ACHD28	208,707	156,726	154,697	100,794	41,921,557	416	98.09	93.98	57.50	48.29
ACHD29	204,074	158,618	156,508	98,801	41,026,284	415	98.16	94.12	57.07	48.41
ACHD30	200,386	149,155	146,916	97,658	40,418,596	414	98.10	94.07	57.21	48.73
ACHD31	216,341	170,233	167,884	111,051	45,639,295	411	98.03	93.84	57.75	51.33

Definición de tabla:

Raw PE: representa las lecturas de PE originales después de la secuenciación;

Raw Tags: representa etiquetas fusionadas de lecturas de PE;

Clean Tags: representa las etiquetas después del filtrado;

Las etiquetas efectivas: representan etiquetas después de filtrar la quimera y finalmente se pueden usar para análisis posteriores;

Base: es el número de bases de las Etiquetas Efectivas;

AvgLen: representa la longitud promedio de las etiquetas efectivas;

Q20 y Q30: son los porcentajes de bases cuyo valor de calidad en Etiquetas Efectivas es superior a 20 (tasa de error de secuenciación inferior al 1%) y 30 (tasa de error de secuenciación inferior al 0,1%);

GC (%): representa el contenido de GC en etiquetas efectivas;

+Efectivo (%): representa el porcentaje de Etiquetas Efectivas en Raw PE.

Directorio de resultados:

PE lee con código de barras y secuencia de cebador:  
resultado/00.RawData/Sample\_Name/\*.raw\_1(2).fq.gz;

PE lee sin código de barras ni secuencia de cebador:  
resultado/00.RawData/Sample\_Name/\*\_1(2).fq.gz;

Etiquetas sin procesar: resultado/00.RawData/Sample\_Name/ \*.extendedFragments.fastq;

Etiquetas efectivas:  
resultado/01.CleanData/Sample\_Name/\*.fastq;resultado/01.CleanData/Sample\_Name/\*.fna;

Información de códigos de barras y cebadores: resultado/00.RawData/SampleSeq\_info.xls

## Datos fisicoquímicos de muestras de suelo

Las muestras de suelo fueron analizadas mediante Espectrofotometría de Absorción atómica con un espectrofotómetro Shimadzu modelo AA-7000 en el laboratorio de suelos de la Facultad de Ciencias Agropecuarias de la Universidad de Panamá (Tabla, 3)

Tabla 3. Datos Fisicoquímicos

sample-id	barcode-sequence	linker-sequence	sample-type	Location	Condition	pH	Magnesi o	%organic content	Aluminium	phosphorus
#q2:types	categoric al	categorical	categorical	categorical	categorica l	categ orical	categoric al	categorical	categorical	numerical
ACHD28.1	GGCTA C	CCTAYGGGRBG CASCAG	Cutanic	Sur de Puerto Armuelles	no intervenid o	alto	medio	alto	alto	88.49
ACHD29.1	CTTGTA	CCTAYGGGRBG CASCAG	Cutanic	PetroTermina l-Puerto Armuelles	intervenid o	medi o	bajo	alto	medio	48.72
ACHD30.1	ATCAC G	CCTAYGGGRBG CASCAG	Hyperdistri c	Horconcitos Chiriquí	intervenid o	bajo	medio	medio	medio	13.99
ACHD31.1	CGATG T	CCTAYGGGRBG CASCAG	Hyperdistri c	Vía hacia Boca Chica- Chiriquí	no intervenid o	bajo	alto	medio	alto	19.12

## Clasificación de secuencias, filtrado y análisis de datos

El filtrado de las secuencias obtenidas en bruto supone un paso importante, a fin de obtener una correcta y exacta asignación taxonómica, en este estudio se utilizaron las secuencias forwards de las cuatro muestras de suelo. Empezando con la demultiplexación de las secuencias. Para esto se requería conocer el código de barra asociado a cada muestra contenido en la Metadata de las muestras (sample-metadata) (Tabla, 4). Esta columna era necesaria debido a que durante la secuenciación se juntan varias muestras amplificadas con primers específicos de cada muestra o Barcode. El demultiplexado consiste en separar las secuencias de cada muestra según el Barcode. Las secuencias demultiplexadas tenían adaptadores y cebadores de PCR que también fueron eliminados utilizando el “plugin cutadapt trim single en qiime2” y que genera el artefacto trimmed-seqs.qza (OUTPUT) dando un resumen de recuento de secuencias demultiplexada (Tabla, 6). Se generó un total de 829,389 secuencias procesadas y un 91% de las secuencias

tenía en promedio 227 nts (Tabla 7). También se obtuvo el mínimo, la mediana, la media y máximo del recuento de las secuencias (Tabla, 8).

Tabla 4. Sample metadata de las muestras de suelos analizadas.

Archivo de metadatos está formado por la columna SampleID que contiene el nombre de la muestra y la columna BarcodeSequence, que es la secuencia nucleotídica del código de barras que permite identificar las muestras. también se incluyen las columnas LinkerPrimerSequence, sample type, location y condition.

ID de Muestras	Barcode-sequence	Linker-sequence	Tipos de Muestras	Localizacion	Condicion
#q2:types	categorical	Categorical	categorical	categorical	categorical
AChD28.1	GGCTAC	CCTAYGGGRBGCASCAG	ACRISOLS-Cutanic	Sur de Puerto Armuelles	no intervenido
AChD29.1	CTTGTA	CCTAYGGGRBGCASCAG	ACRISOLS-Cutanic	PetroTerminal-Puerto Armuelles	intervenido
AChD30.1	ATCACG	CCTAYGGGRBGCASCAG	ACRISOLS-Hyperdistric	Horconcitos Chiriquí	intervenido
AChD31.1	CGATGT	CCTAYGGGRBGCASCAG	ACRISOLS-Hyperdistric	Vía hacia Boca Chica-Chiriquí	no intervenido

Tabla 5. Total, de las secuencias procesadas de cada muestra de este estudio

<b>Secuencias procesadas</b>	<b>829,508</b>
<b>Secuencias con adaptadores</b>	829,508 (100%)
<b>Secuencias que fueron muy cortas</b>	123 (0.0%)
<b>Secuencias filtradas</b>	829,389
<b>Pares de bases procesadas</b>	207,377,000 bp
<b>Pares de bases filtradas</b>	195,951,940 (94.5%)

Tabla 6. Recuento de secuencias (forward) por muestras

ID de la Muestra	Recuento secuencia foward
AChD29.1	225932
AChD28.1	212622
AChD31.1	198663
AChD30.1	192172
<b>Total</b>	<b>829,389</b>

Tabla 7. Resumen de la longitud de las secuencias demultiplexadas

<b>Forward reads</b>	
<b>Total, de secuencias muestreadas</b>	10000
<b>2%</b>	21 nts
<b>9%</b>	227 nts
<b>25%</b>	227nts
<b>50% (Mediana)</b>	227 nts
<b>75%</b>	227 nts
<b>91%</b>	227 nts
<b>98%</b>	244 nts

Tabla 8. Resumen de recuentos de las secuencias por muestra (Foward)

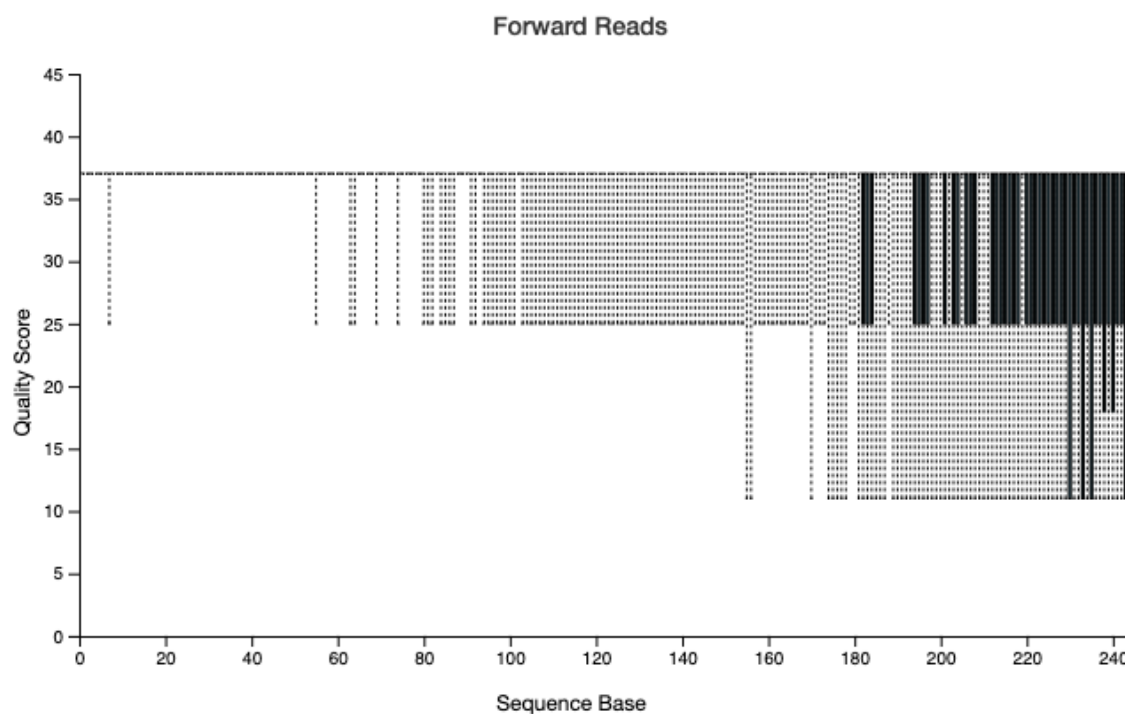
Se muestra el total de las secuencias por muestras. También se muestra el mínimo, la mediana, la media y máximo del recuento de las secuencias

<b>Resumen de recuentos de secuencias demultiplexadas</b>	
<b>Mínimo</b>	192172
<b>Mediana</b>	205642
<b>Media</b>	207347
<b>Máximo</b>	225932
<b>Total</b>	829,389

El artefacto demultiplexado nos permitió también crear un resumen interactivo de nuestras secuencias. Este resumen proporciona información útil para evaluar la calidad de la corrida de secuenciación de ADN, incluido el número de secuencias que se obtuvieron por muestra y la distribución de las puntuaciones de calidad (Phred quality score) de la secuencia en cada posición. Basados en la gráfica interactiva de calidad (interactive quality plot) de las secuencias demultiplexadas y visualizadas con Qiime2 view, las secuencias con largo mayor de 155 fueron



eliminadas ya que la calidad cae por debajo de 20 del puntaje de calidad (Phred quality score) (Figura 11).



**Figura 11.** Visualización Qiime2 de la gráfica interactiva que resume los resultados de demultiplexación de las secuencias.

La eliminación de ruido de las secuencias se llevó a cabo mediante el método (plugin) de DADA2 de Qiime2. DADA2 es un paquete de R designado para determinar los ASV (amplicon sequence variants). En el pipeline de DADA2 las secuencias fueron sometidas a un proceso de filtrado, eliminación de ruido y quimeras mediante la aplicación del método DADA2 (Callahan *et al.*, 2016), usando el comando `qiime dada2 denoise-single` y resolviendo diferencias tan pequeñas como un nucleótido, a diferencia de la agrupación en unidades operacionales taxonómicas (OTUs) (Callahan *et al.*, 2017).

DADA2 (`dada2 denoise-single`) utiliza los parámetros `--p-trim-left-f` y `--p-trim-left-r` de DADA2, que requirieron utilizar la gráfica de calidad interactiva.

1. `P-trim-left` intercepta la secuencia de baja calidad en el extremo izquierdo, miramos el diagrama de caja en la figura anterior, la calidad del extremo izquierdo

es muy alta, no hay un área de baja calidad; pero se eliminó el linker y el barcode al asignarle el valor de 22.

2. -P-trunc-len La longitud de truncamiento de la secuencia también es para eliminar la secuencia de baja calidad en el extremo derecho. Vemos que después de que es mayor que 155, la calidad cae mucho e incluso la mediana cae por debajo de 20, por lo que debe eliminarse.

El resultado de DADA2 será un artefacto QIIME 2 FeatureTable (rep-seqs.qza: name of the output file where the sequences of each ASV are stored), que contiene recuentos (frecuencias) de cada secuencia única en cada muestra en el conjunto de datos. Es decir, son nuestras secuencias representativas, es de tipo, o lo que se podría considerar “OTU’s al 100%”, es decir aquellas secuencias únicas (feature: Una unidad de observación, como una unidad taxonómica operativa, una variante de secuencia, un gen, un metabolito, etc. Este término genérico se utiliza porque QIIME 2 puede admitir muchos tipos diferentes de características). Estas secuencias son los “Sequence Count” que en este estudio son 6095, también nos brinda información del tamaño que tienen, min, max, promedio (Tabla 9). Se observa directamente la secuencia resaltada en azul que esta enlazada con BLAST del NCBI y que al hacer clic en esta dará el resultado de BLAST de la secuencia. Además, se le asigna un id “Feature id”, por lo que cada una de estas secuencias representativas ya está “mapeada”. También podemos descargar estas secuencias en formato FASTA, sin embargo, no contienen datos de abundancia, por lo que solo son útiles para asignación taxonómica, no para índices de diversidad (Tabla, 10)

Tabla 9. Estadística obtenida con DADA2 de las secuencias filtradas y no quiméricas por muestra de suelo

Estadística del largo de las secuencias	
Conteo	6095
Mínimo	131
Máximo	131
Media	131
Rango	0
Std	0

Otro artefacto stats- dada2.qzv describe en números como fueron sufriendo cambios, en cada etapa de tratamiento las secuencias. Así como la proporción relativa a la entrada inicial de

secuencias y después del filtrado se obtuvo un total de 763,352 secuencias para las cuatro muestras (Tabla, 11)

Tabla 10. Estadística del largo de las secuencias filtradas y el número de características (features) de las cuatro muestras

### Sequence Table

To BLAST a sequence against the NCBI nt database, click the sequence and then click the View report button on the resulting page.

Download your sequences as a raw FASTA file

Click on a Column header to sort the table.

Feature ID	Sequence Length	Sequence
fde4eb9ae0f27146869198c45744d110	131	<a href="#">ATATTGCACAATGGGCGCAAGCCTGATGCAGCGACGCCGCTGAGGGATGACGGCCTTCGGGTTGTAACCTCTTTTCAGTAGGGGAAGAAGCGAAGTACGGTACCTGCAGAGAAGCGCCGGCTA</a>
bb80922e434328eb5ff93578813c2926	131	<a href="#">ATATTGCACAATGGGCGCAAGCCTGATGCAGCGACGCCGCTGAGGGATGACGGCCTTCGGGTTGTAACCTCTTTTCAGTAGGGGAAGAAGCGTAAAGTACGGTACCTGCAGAGAAGCGCCGGC</a>
16c2ba753453536da565b900a1a9c1f1	131	<a href="#">ATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCTGGGTGATGAAGGCCCTAGGGTTGTAAGCCCTTTTCGGCGGGGAAGATAATGACGGTACCCGAGAGAAGCCCGGGCTAACTC</a>

Tabla 11. Se asigna un id “Feature id” por lo que cada una de estas secuencias representativas ya está “mapeada

sample-id	input	filtered	percentage of input passed filter	denoised	non-chimeric	percentage of input non-chimeric
#q2:types	numeric	numeric	numeric	numeric	numeric	numeric
ACHD28.1	212622	197868	93.06	192490	167221	78.65
ACHD29.1	225932	192975	85.41	187669	163069	72.18
ACHD30.1	192172	180941	94.16	175946	151122	78.64
ACHD31.1	198663	191568	96.43	187414	155857	78.45

Un artefacto FeatureData [Secuencia] QIIME 2 (table.qzv), que asigna identificadores de características en FeatureTable a las secuencias que representan y por lo tanto indican su riqueza en cada muestra de suelo. Este artefacto tiene la relación de todos los Feature id (características) y su abundancia, esto está en la pestaña Feature Detail. Este archivo, es el que nos sirvió para los índices de diversidad. Otro aspecto importante es el Interactive Sample Detail ya que nos indica los *reads* (lecturas) por muestra. Así podemos saber cuál es la abundancia mínima de nuestro estudio (Tabla 12, 13)

Tabla 12. Artefacto FeatureData [Secuencia] QIIME 2 (table.qzv), que asigna identificadores de características en FeatureTable a las secuencias que representan en cada muestra de suelo estudiada.

<b>Muestra ID</b>	<b>Recuento de Características (feature count)</b>
<i>ACHD28.1</i>	167221
<i>ACHD29.1</i>	163069
<i>ACHD31.1</i>	155857
<i>ACHD30.1</i>	151122
<i>Total</i>	637269

Tabla 13. Obtenida de la visualización tabla.qzv indicando el número de ASV (amplicon sequence variants) para las cuatro muestras de suelo y su frecuencia en estas

<b>Resumen de la Tabla</b>	<b>Muestra</b>
<b>Número de Muestras</b>	4
<b>Número de características (Feature: ASV)</b>	6095
<b>Frecuencia total</b> (nos indica el número de veces que las características son observadas en las cuatro muestras)	637,269

### Árbol filogenético para análisis de biodiversidad

Los análisis de biodiversidad requieren la generación de un árbol filogenético el cual se hizo con el complemento (plugin) q2-phylogeny, usando el pipeline align-to-tree-mafft-fasttree. Primero, el pipeline usa el programa mafft para realizar una alineación de secuencia múltiple de las secuencias en nuestro FeatureData [Sequence] para crear un artefacto o tabla de secuencias alineadas (FeatureData [AlignedSequence] QIIME 2). A continuación, el pipeline enmascara (o filtra) la alineación para eliminar posiciones que son muy variables en el alineamiento. Generalmente se considera que estas posiciones agregan ruido al árbol filogenético resultante. Después de eso, el pipeline aplica el método FastTree para generar un árbol filogenético a partir de la alineación enmascarada. El programa FastTree crea un árbol sin raíz, por lo que en el paso

final de esta sección se aplica el método midpoint-root para generar un árbol con raíz en el punto medio de la distancia de punta a punta más larga en el árbol sin raíz.

### **Análisis de la diversidad alfa y beta**

Este trabajo previo ha permitido el estudio de la diversidad alfa y beta con el plugin q2- diversity con el método core-metrics-phylogenetic que rarifica la tabla de frecuencias de nuestras ASV a la profundidad (depth) de 151122 (determinada durante la visualización de la tabla de frecuencias). Este paso submuestreo los recuentos en cada muestra sin reemplazo de modo que cada muestra en la tabla resultante tenga un recuento total de 151122, porque las métricas de diversidad son sensibles a las diferentes profundidades de muestreo a lo largo de las distintas muestras.

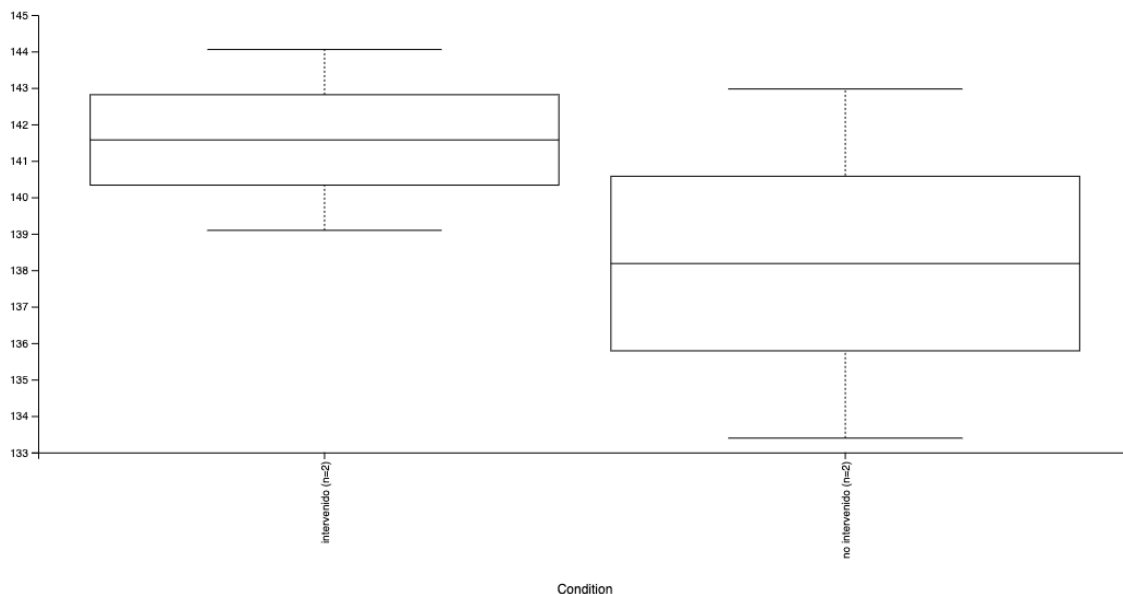
Este valor se eligió en función del número de secuencias en la muestra ACHD30.1 porque está cerca del número de secuencias en las siguientes muestras que tienen recuentos de secuencias más alto. Esto nos permitió conservar todas nuestras muestras.

### **Diversidad Alfa**

Después de calcular las métricas de diversidad, se analizó la composición microbiana de las muestras en el contexto de los metadatos de la muestra (sample-metadata) por medio de la filogenética de Faith (una medida de la riqueza de la comunidad) (Faith, 1992) y las métricas de uniformidad de Pielou. Esto se define como la suma de las longitudes de las ramas de un árbol filogenético y tiene en cuenta las diferencias filogenéticas entre especies. La diversidad filogenética (Faith's PD) usa la distancia filogenética para calcular la diversidad de una muestra dada. Se utilizó el método de Kruskal- Wallis para realizar las comparaciones por muestras y ver si existían diferencias significativas ( $p$ -value  $< 0.05$ ) de alfa diversidad. La medida basada en la longitud total de la rama de un árbol filogenético que comprende una muestra de Visualización de diagrama (Grafico, 1) de caja de índices de diversidad filogenética taxonómica (de Faith) para comunidades procariotas (gen 16S rRNA) asociadas con los factores fisicoquímicos respectivamente.

Cuando los resultados de la medida de la alpha diversidad fueron visualizados en QIIME2 VIEW, se produce una gráfica de bigotes, como se muestra en el gráfico 3 (tipo de muestra). Todas las variables de la metadata con la medida de alfa diversidad ( $p$ -value  $< 0.05$ ) correspondientes están listados en la tabla 14.

Los resultados de las pruebas de Kruskal-Wallis indican que no hay diferencias significativas de la diversidad microbiana (diversidad filogenética de Faith) entre la condición de suelo, tipo de suelo, pH, Magnesio, aluminio y contenido orgánico



**Grafico 1.** Caja de bigotes para la diversidad filogenetica de faith en suelos intervenidos y no intervenidos

Tabla 14. variables de la metadata con la medida de alfa diversidad (p-value < 0.05)

Variable de Metada	P-value
Tipo de suelo (cutánico e hiperdistrico)	0.12133525035848367
Condición (Intervenido, no intervenido)	0.43857802608099994
pH	0.2592402606458915
Magnesio	0.2592402606458915
% contenido orgánico	0.12133525035848367
Aluminio	0.43857802608099994

Se muestra las variables de la metadata y su respectivo valor de P como medida de la riqueza de las muestras estudiadas

La uniformidad de la comunidad es representada por el índice de uniformidad de Pielou. Este índice mide la regularidad o uniformidad con que los individuos están distribuidos dentro de las especies, y no tanto cuantas especies hay. Es posible calcular las medidas de uniformidad (también llamada en algunos libros equitatividad) de una comunidad mediante una ecuación sencilla usando el índice de Pielou:

Los resultados para esta métrica fueron similares que para la filogenética de Faith (Grafico 2), no existen diferencias significativas entre las diversas variables de la metadata y la uniformidad de la comunidad (Tabla, 15)

**Grafico 2.** Caja de bigotes de uniformidad de Pielou representando la uniformidad de las especies en las muestras y su relación con una las variables de la metadata. (aluminio)

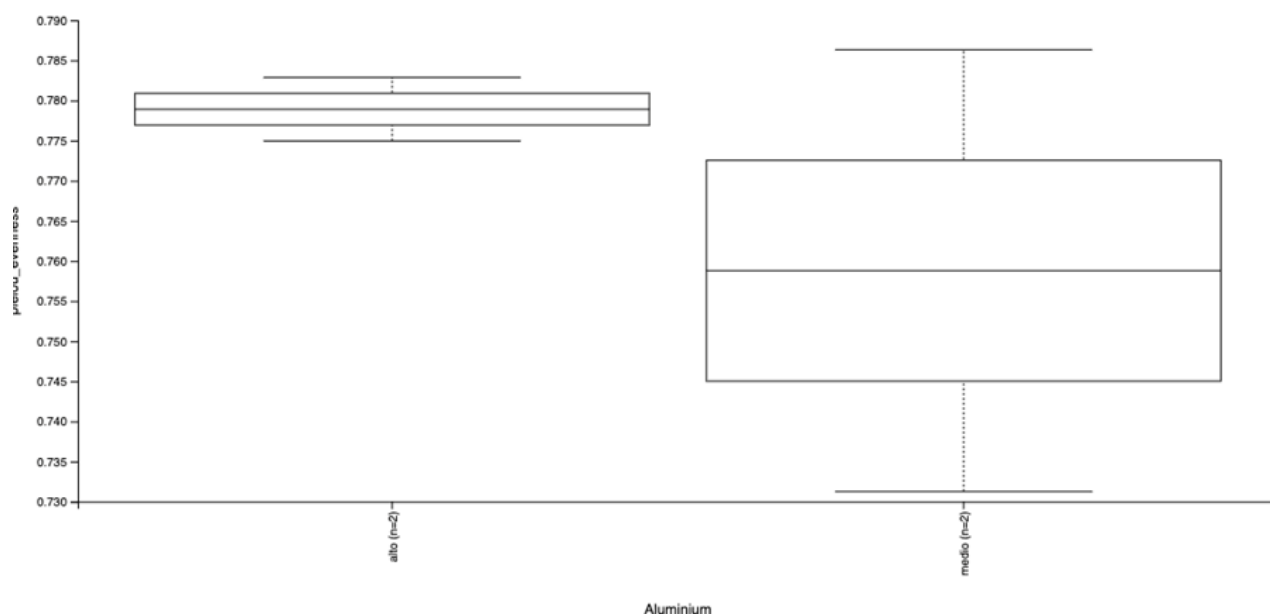


Tabla 15. P- value para esta métrica para la filogenética de Faith en cada uno de los parámetros fisicoquímicos.

<b>Variable de Metada</b>	<b>P-value</b>
<b>Tipo de suelo (cutánico e hiperdistrico)</b>	0.12133525035848367
<b>Condición (Intervenido, no intervenido)</b>	1
<b>pH</b>	0.25924026
<b>Magnesio</b>	0.4065696597405991
<b>% contenido orgánico</b>	0.12133525035848367
<b>Aluminio</b>	1

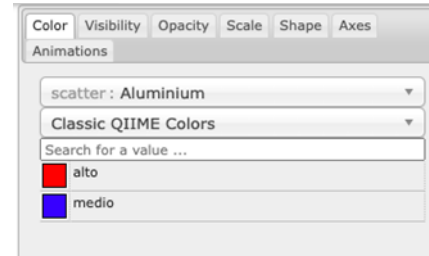
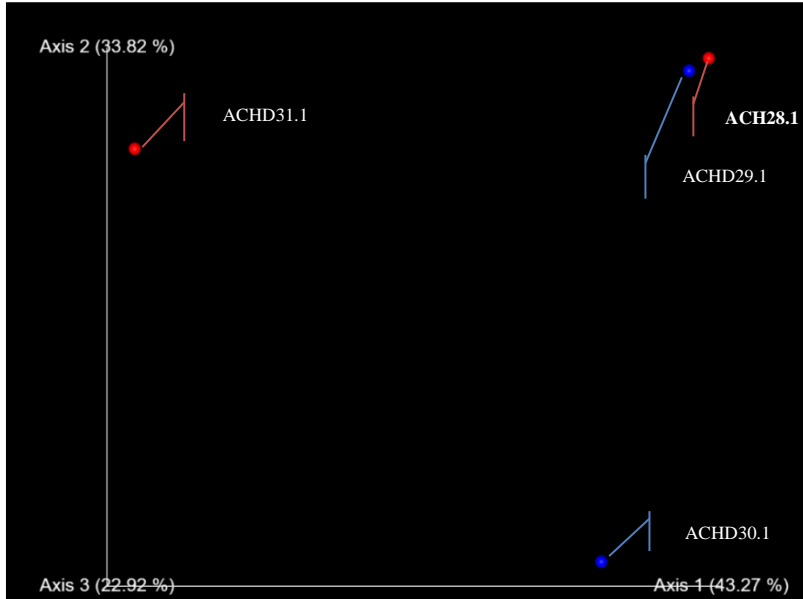
### **Diversidad beta**

Emperor es una herramienta que nos permite trazar el análisis de coordenadas principales de nuestros datos. Para el cálculo de la diversidad beta se utilizó la métrica UniFrac. UniFrac nos permite calcular las diferencias cualitativas (análisis no ponderado) y cuantitativas (análisis ponderado) entre las muestras analizadas, en base a sus relaciones filogenéticas (Lozupone *et al.*, 2007). Para ello, se generó una matriz basada en las distancias filogenéticas y los resultados obtenidos se visualizaron mediante un análisis de coordenadas principales (PCoA) en donde se eligieron las coordenadas que tenían un mayor peso en la distribución de las muestras. Por tanto, explora el grado de divergencia entre las diferentes secuencias que hemos obtenido, al poder observar la distancia métrica entre muestras. Así mismo, podemos comparar muchas muestras simultáneamente utilizando un análisis de coordenadas principales, al igual que con el análisis PCoA, puesto que podemos obtener una matriz de distancias. Básicamente, el concepto UniFrac se refiere al porcentaje de longitud de las ramas filogenéticas únicas para cada muestra. El método UniFrac ponderado o weighted, además tiene en cuenta las abundancias relativas

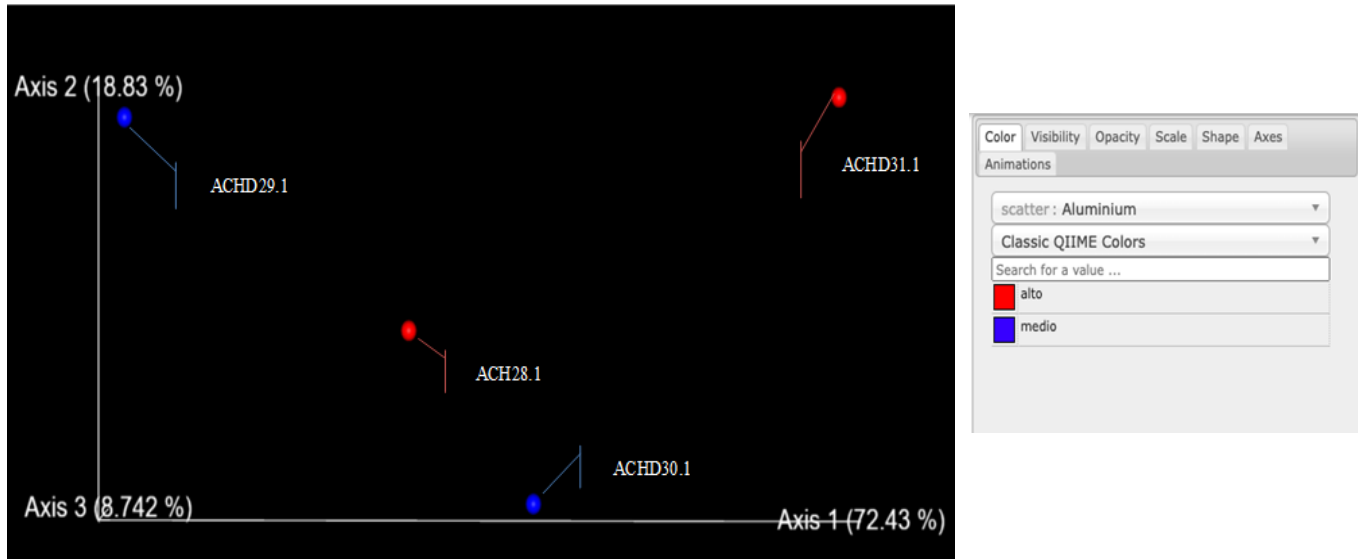


existentes en los rangos taxonómicos, por tanto, enfatiza los grupos dominantes. Por otra parte, el método UniFrac no ponderado o unweighted, solo tiene en cuenta presencia o ausencia, enfatizando más en los grupos menos comunes, de forma general. Los análisis de coordenadas principales (PCoA) muestran cómo se separaron las muestras de cada una de las comunidades de los tipos de suelo y de acuerdo con los parámetros fisicoquímicos en grupos claros.

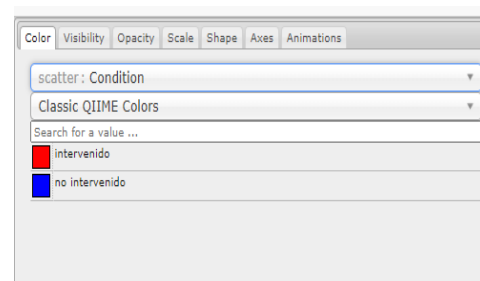
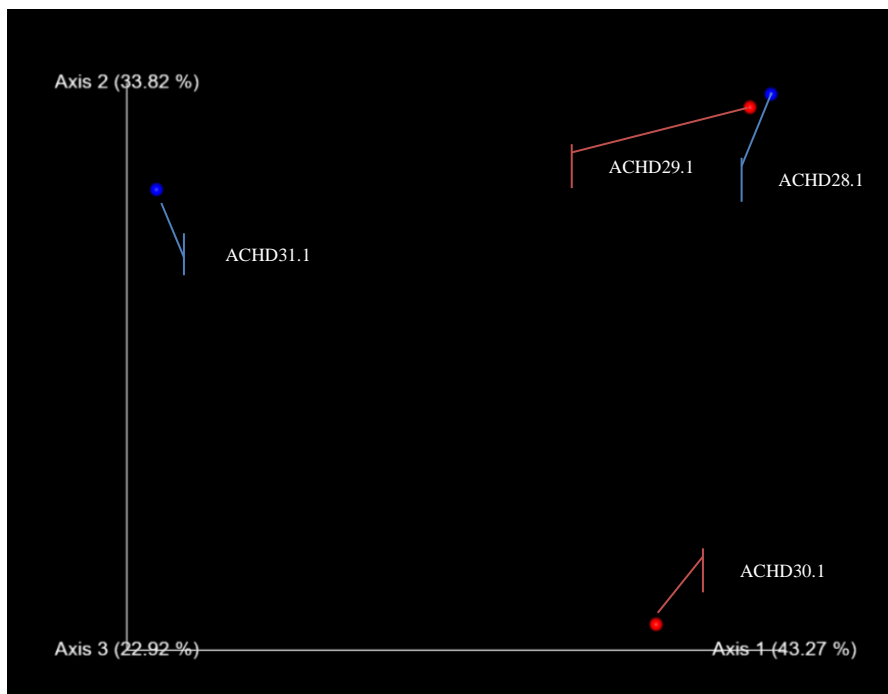
A continuación presentamos el resultado de análisis de coordenadas principales (PCoA) usando Unifrac no ponderado y Unifrac ponderado como medidas de distancia filogenética, en las siguientes figuras de la 12 a la 29 y observa los 4 ejes (Axis 1, Axis 2, Axis 3 y Axis 4) donde se determina la diferencia en datos unifrac ponderado que considera la frecuencia relativa de las características representadas en los datos cualitativos y unifrac no ponderada consiste en la frecuencia relativa de las características cuantitativos.



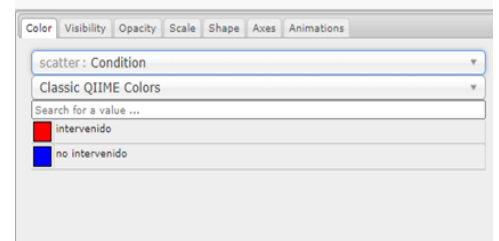
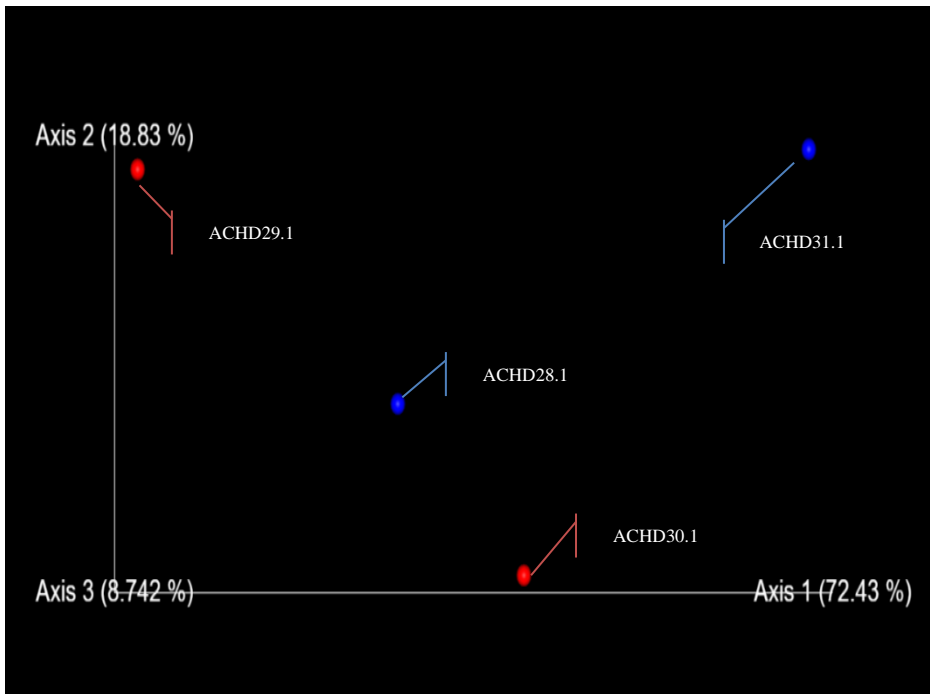
**Figura 12.** Análisis de Coordenadas Principales una medida cualitativa que relaciona la diversidad de las bacterias presente en suelo con la concentración de aluminio. Se observa el PCoA (Análisis de Coordenadas Principales) Unifrac no ponderado cualitativo para el aluminio, el eje Axis 1 mostro 43.27% para la muestra ACHD30.1 y el eje Axis 2 el 33.82% la muestra ACHD31.1 que representa diferencia significativa en la medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de aluminio en proporción al porcentaje de visualización PCoA aquí las muestras se agrupan en núcleos bastante definidos. Por otra parte ambas muestras si mostro distancias filogenéticas en PCoA. El eje Axis 3 de 22.92% no se observó muestras, y por último el eje Axis 4 se visualizó las muestras ACHD28.1, ACHD29.1 la cual no son significativos en medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de aluminio en proporción al porcentaje y no mostro distancias filogenéticas en PCoA.



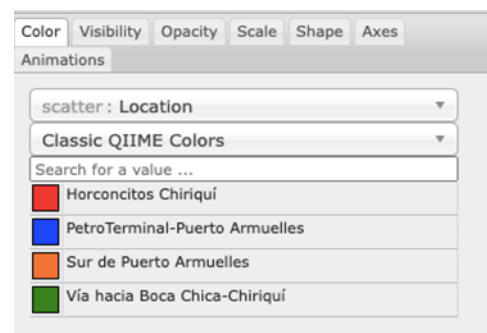
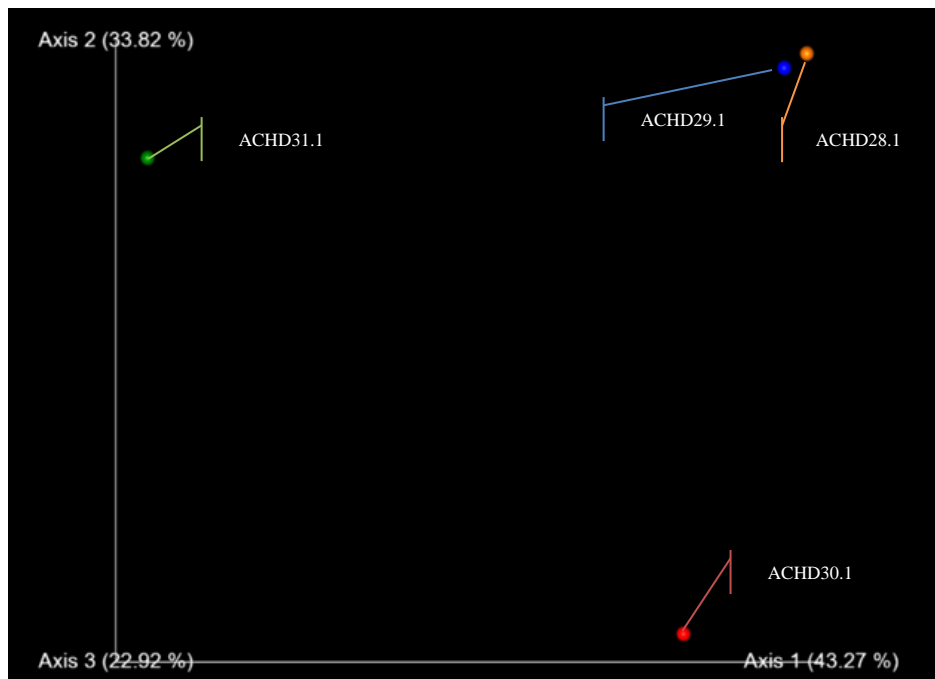
**Figura 13.** Análisis de Coordenadas Principales en medida cuantitativa que relaciona la diversidad de las bacterias presente en suelo con el aluminio. Se visualiza el PCoA (Análisis de Coordenadas Principales) Unifrac ponderado cuantitativo para aluminio. El eje Axis 1 mostro 72.43% para la muestra ACHD30.1, el eje Axis 2 el 18.83% la muestra ACHD29.1 y el eje Axis 3 de 8.74% la muestra ACHD 28.1 las cuales representa diferencia significativa en la medida cuantitativa que considera las abundancias relativas de las muestras y los patrones de agrupamiento; estas tres muestras se visualizó la disimilitud de la comunidad bacteriana para niveles alto y medio de aluminio en proporción al porcentaje de PCoA, marcando distancias filogenéticas entre ellas pero con valores cuantitativo significativos aceptables; estas muestras se agrupan en núcleos bastante definidos. Por otra parte, para las muestras ACHD 31.1 del eje Axis 4 no se observó valores significativos en medida cuantitativa de la disimilitud de la comunidad bacteriana para niveles alto de aluminio en proporción al porcentaje de PCoA



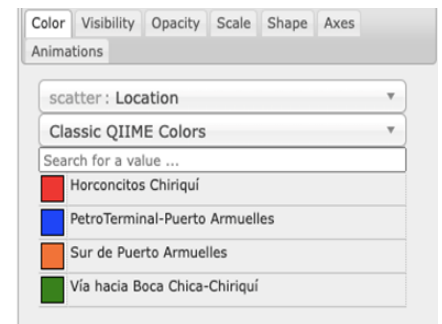
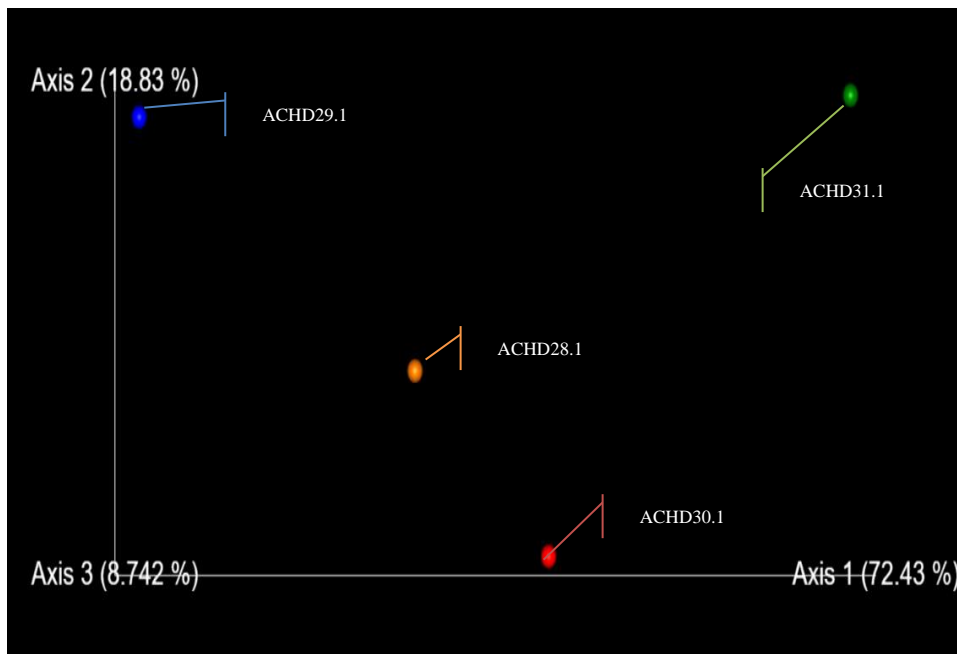
**Figura 14.** Análisis de Coordenadas Principales una medida cualitativa que relaciona la diversidad de las bacterias presente en suelo con la condición (intervenido y no intervenido). Se observa el PCoA (Análisis de Coordenadas Principales) Unifrac no ponderado cualitativo para condición, el eje Axis 1 mostro 43.27% para la muestra ACHD30.1 y el eje Axis 2 el 33.82% la muestra ACHD31.1 que representa diferencia significativa en la medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de aluminio en proporción al porcentaje de visualización PCoA aquí las muestras se agrupan en núcleos bastante definidos. Por otra parte, ambas muestras si mostro distancias filogenéticas en PCoA. El eje Axis 3 de 22.92% no se observó muestras, y por último el eje Axis 4 se visualizó las muestras ACHD28.1, ACHD29.1 la cual no son significativos en medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de aluminio en proporción al porcentaje y no mostro distancias filogenéticas en PCoA



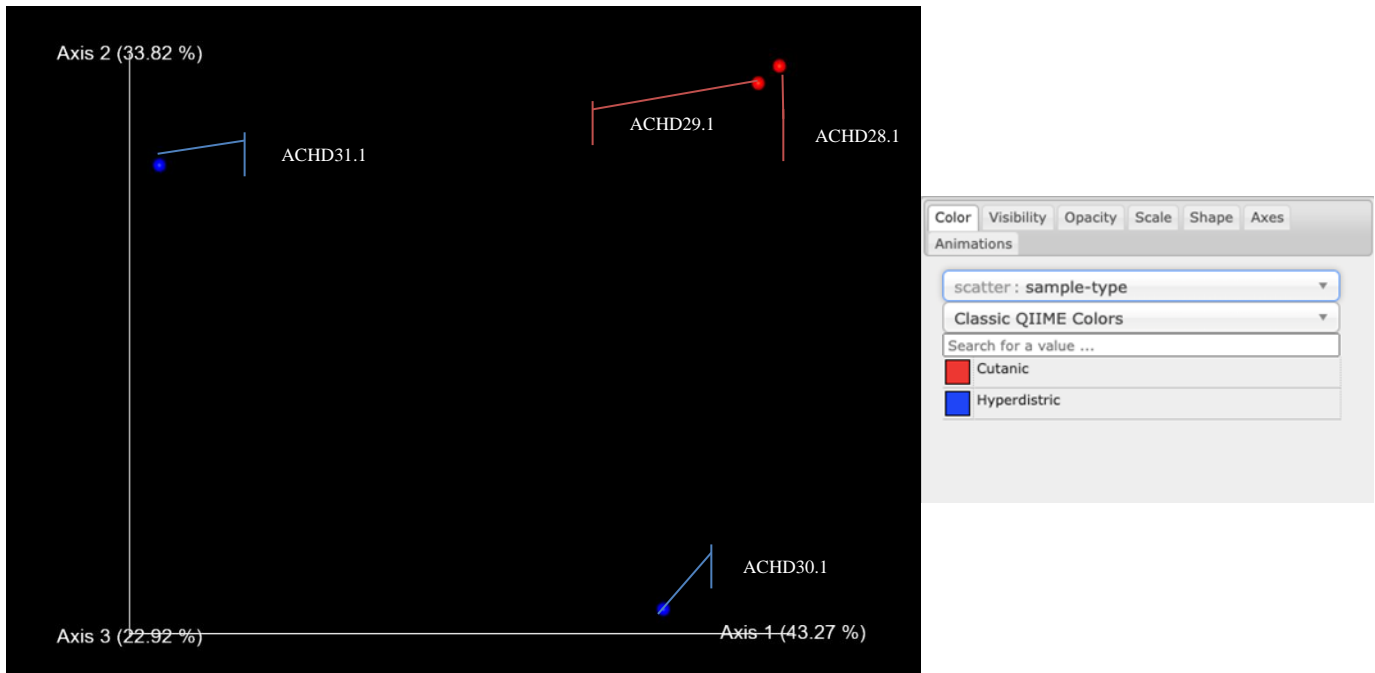
**Figura 15.** Análisis de Coordenadas Principales una medida cuantitativa que relaciona la diversidad de las bacterias presente en suelo con la condición (intervenido y no intervenido). Se visualiza el PCoA (Análisis de Coordenadas Principales) Unifrac ponderado cuantitativo para condición. El eje Axis 1 mostro 72.43% para la muestra ACHD30.1, el eje Axis 2 el 18.83% la muestra ACHD29.1 y el eje Axis 3 de 8.74% la muestra ACHD 28.1 las cuales representa diferencia significativa en la medida cuantitativa que considera las abundancias relativas de las muestras y los patrones de agrupamiento; estas tres muestras se visualizó la disimilitud de la comunidad bacteriana para niveles alto y medio de aluminio en proporción al porcentaje de PCoA, marcando distancias filogenéticas entre ellas pero con valores cuantitativo significativos aceptables; estas muestras se agrupan en núcleos bastante definidos. Por otra parte, para las muestras ACHD 31.1 del eje Axis 4 no se observó valores significativos en medida cuantitativa de la disimilitud de la comunidad bacteriana para niveles alto de condicion en proporción al porcentaje de PCoA



**Figura 16.** Análisis de Coordenadas Principales una medida cualitativa que relaciona la diversidad de las bacterias presente en suelo con la localización. Se observa el PCoA (Análisis de Coordenadas Principales) Unifrac no ponderado cualitativo para localización el eje Axis 1 mostro 43.27% para la muestra ACHD30.1 y el eje Axis 2 el 33.82% la muestra ACHD31.1 que representa diferencia significativa en la medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de localización en proporción al porcentaje de visualización PCoA aquí las muestras se agrupan en núcleos bastante definidos. Por otra parte ambas muestras si mostro distancias filogenéticas en PCoA. El eje Axis 3 de 22.92% no se observó muestras, y por último el eje Axis 4 se visualizó las muestras ACHD28.1, ACHD29.1 la cual no son significativos en medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de localización en proporción al porcentaje y no mostro distancias filogenéticas en PCoA

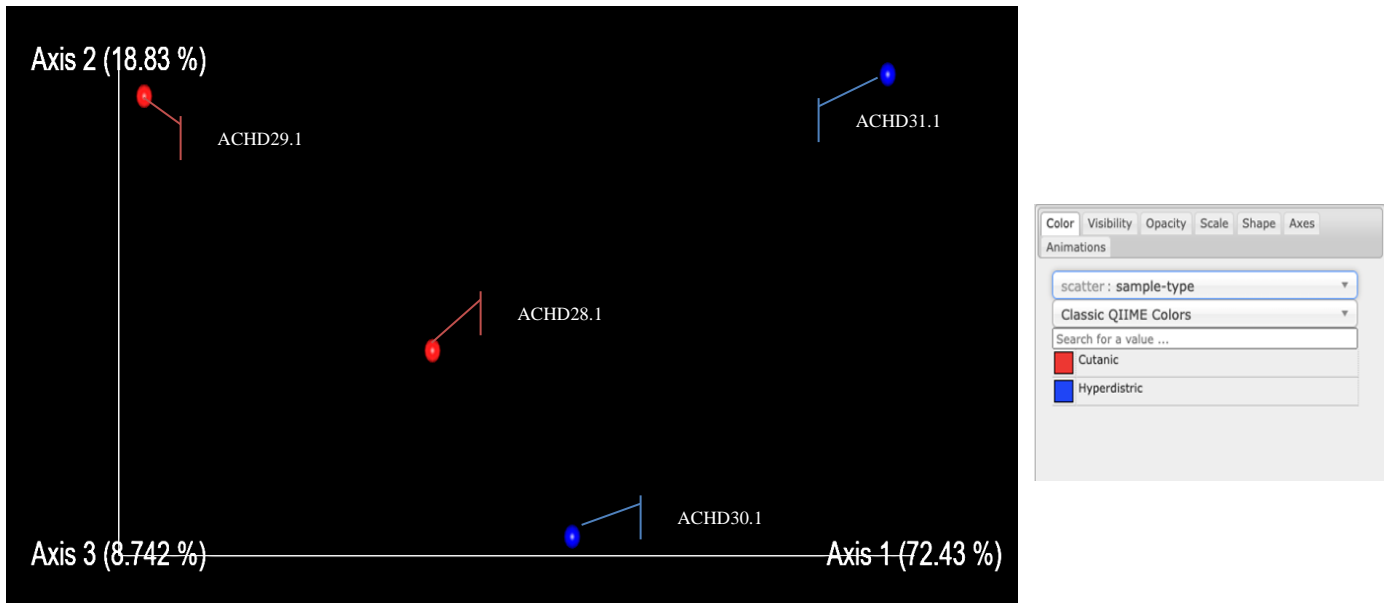


**Figura 17.** Análisis de Coordenadas Principales una medida cuantitativo que relaciona la diversidad de las bacterias presente en suelo con la localización. Se visualiza el PCoA (Análisis de Coordenadas Principales) Unifrac ponderado cuantitativo para localización. El eje Axis 1 mostro 72.43% para la muestra ACHD30.1, el eje Axis 2 el 18.83% la muestra ACHD29.1 y el eje Axis 3 de 8.74% la muestra ACHD 28.1 las cuales representa diferencia significativa en la medida cuantitativa que considera las abundancias relativas de las muestras y los patrones de agrupamiento; estas tres muestras se visualizó la disimilitud de la comunidad bacteriana para niveles alto y medio de localización en proporción al porcentaje de PCoA, marcando distancias filogenéticas entre ellas pero con valores cuantitativo significativos aceptables; estas muestras se agrupan en núcleos bastante definidos. Por otra parte, para las muestras ACHD 31.1 del eje Axis 4 no se observó valores significativos en medida cuantitativa de la disimilitud de la comunidad bacteriana para niveles alto de localización en proporción al porcentaje de PCoA

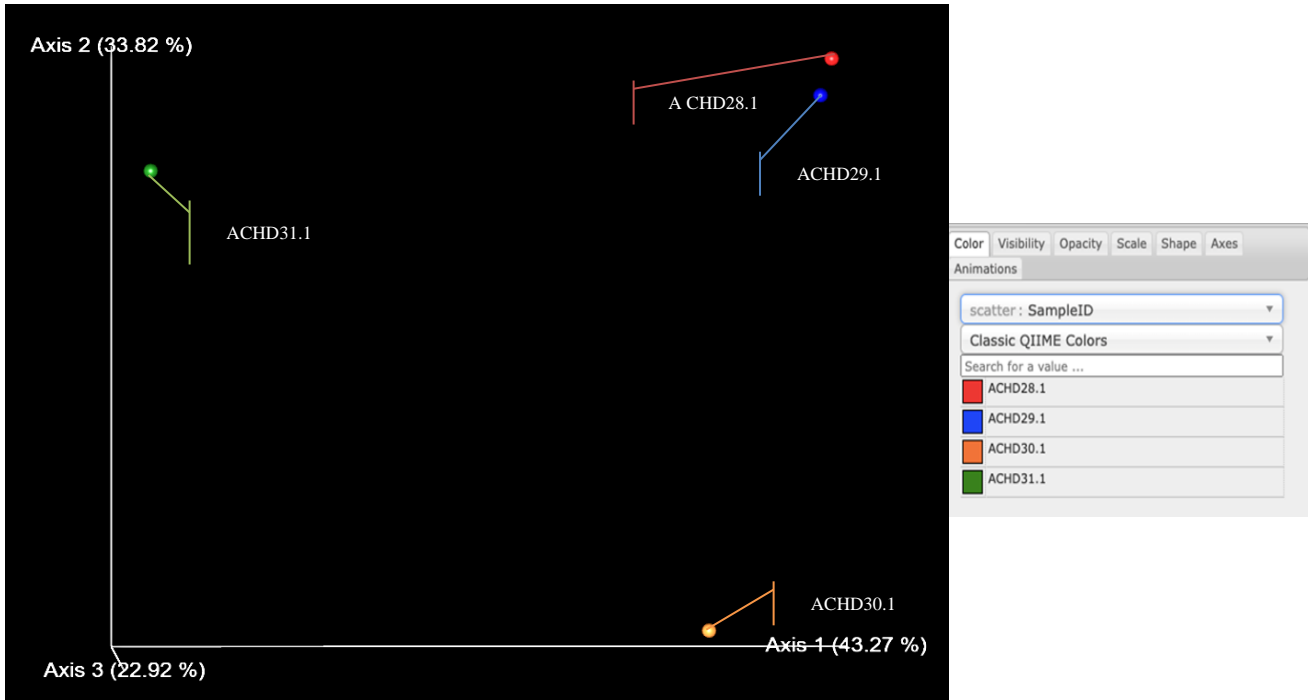


**Figura 18.** Análisis de Coordenadas Principales una medida cualitativa que relaciona la diversidad de las bacterias presente en suelo con los tipos de muestras (cutanico y hiperdistrico). Se observa el PCoA (Análisis de Coordenadas Principales) Unifrac no ponderado cualitativo para tipo de suelo, el eje Axis 1 mostro 43.27% para la muestra ACHD30.1 y el eje Axis 2 el 33.82% la muestra ACHD31.1 que representa diferencia significativa en la medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de tipo de suelo en proporción al porcentaje de visualización PCoA aquí las muestras se agrupan en núcleos bastante definidos. Por otra parte, ambas muestras si mostro distancias filogenéticas en PCoA. El eje Axis 3 de 22.92% no se observó muestras, y por último el eje Axis 4 se visualizó las muestras ACHD28.1, ACHD29.1 la cual no son significativos en medida cualitativa de la disimilitud de la comunidad bacteriana para cutanico y hiperdistrico de tipo de suelo en proporción al porcentaje y no mostro distancias filogenéticas en PCoA

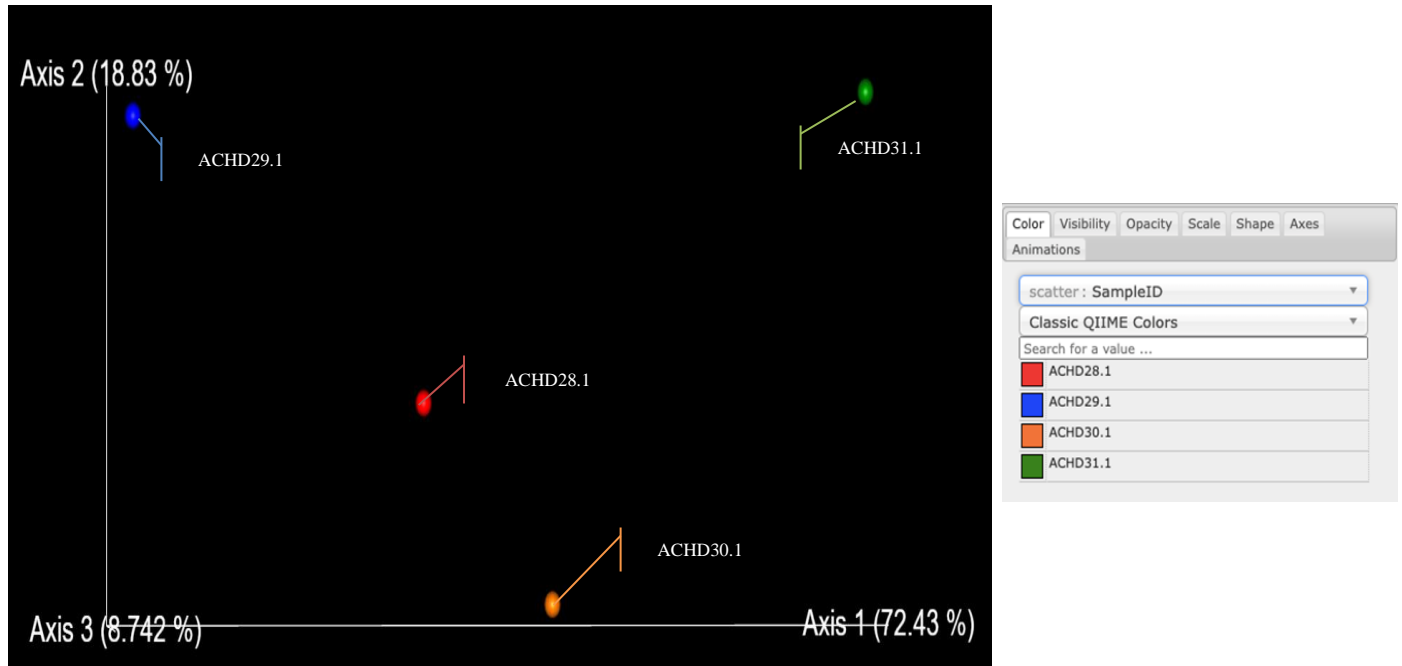




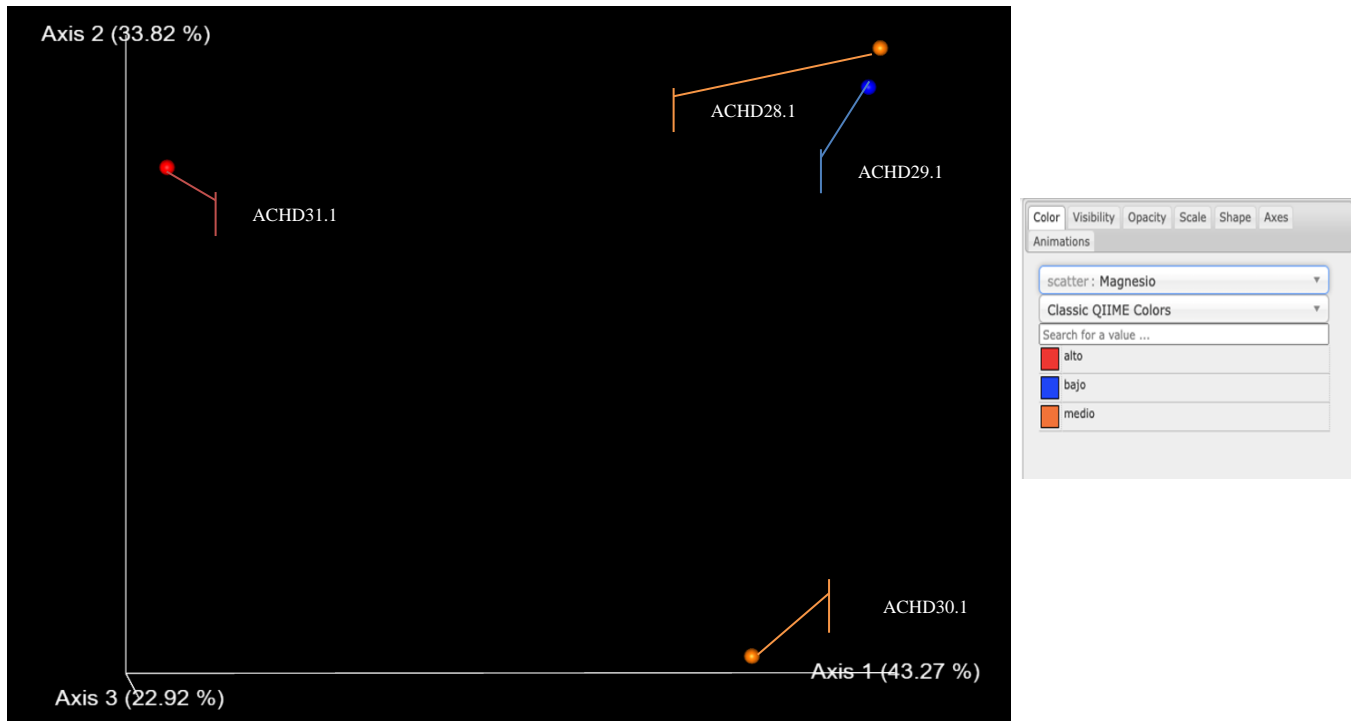
**Figura 19.** Análisis de Coordenadas Principales una medida cuantitativa que relaciona la diversidad de las bacterias presente en suelo con los tipos de muestras (cutanico y hiperdístico). Se visualiza el PCoA (Análisis de Coordenadas Principales) Unifrac ponderado cuantitativo para tipo de suelo. El eje Axis 1 mostro 72.43% para la muestra ACHD30.1, el eje Axis 2 el 18.83% la muestra ACHD29.1 y el eje Axis 3 de 8.74% la muestra ACHD 28.1 las cuales representa diferencia significativa en la medida cuantitativa que considera las abundancias relativas de las muestras y los patrones de agrupamiento; estas tres muestras se visualizó la disimilitud de la comunidad bacteriana para cutanico y hiperdístico de tipo de suelo en proporción al porcentaje de PCoA, marcando distancias filogenéticas entre ellas pero con valores cuantitativo significativos aceptables; estas muestras se agrupan en núcleos bastante definidos. Por otra parte, para las muestras ACHD 31.1 del eje Axis 4 no se observó valores significativos en medida cuantitativa de la disimilitud de la comunidad bacteriana para hiperdístico muestra ACHD 31.1 en proporción al porcentaje de PCoA



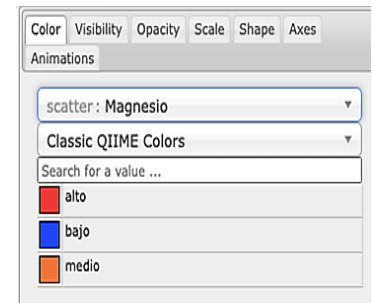
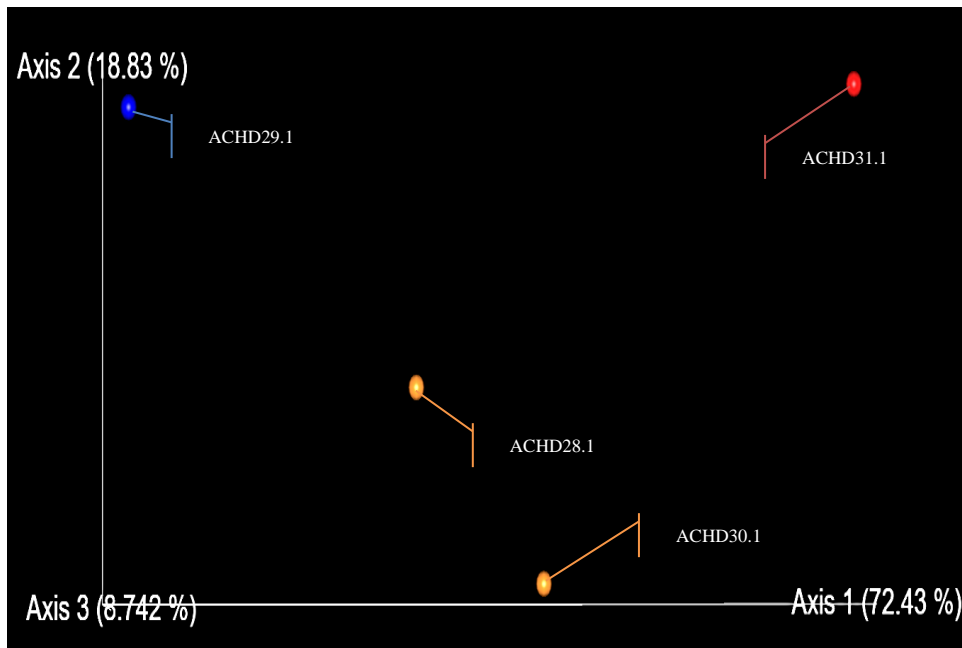
**Figura 20.** Análisis de Coordenadas Principales una medida cualitativa que relaciona la diversidad de las bacterias presente en suelo con el SampleID. Se observa el PCoA (Análisis de Coordenadas Principales) Unifrac no ponderado cualitativo para SampleID, el eje Axis 1 mostro 43.27% para la muestra ACHD30.1 y el eje Axis 2 el 33.82% la muestra ACHD31.1 que representa diferencia significativa en la medida cualitativa de la disimilitud de la comunidad bacteriana para SampleID en proporción al porcentaje de visualización PCoA aquí las muestras se agrupan en núcleos bastante definidos. Por otra parte, ambas muestras si mostro distancias filogenéticas en PCoA. El eje Axis 3 de 22.92% no se observó muestras, y por último el eje Axis 4 se visualizó las muestras ACHD28.1, ACHD29.1 la cual no son significativos en medida cualitativa de la disimilitud de la comunidad bacteriana para cada muestra SampleID en proporción al porcentaje y no mostro distancias filogenéticas en PCoA



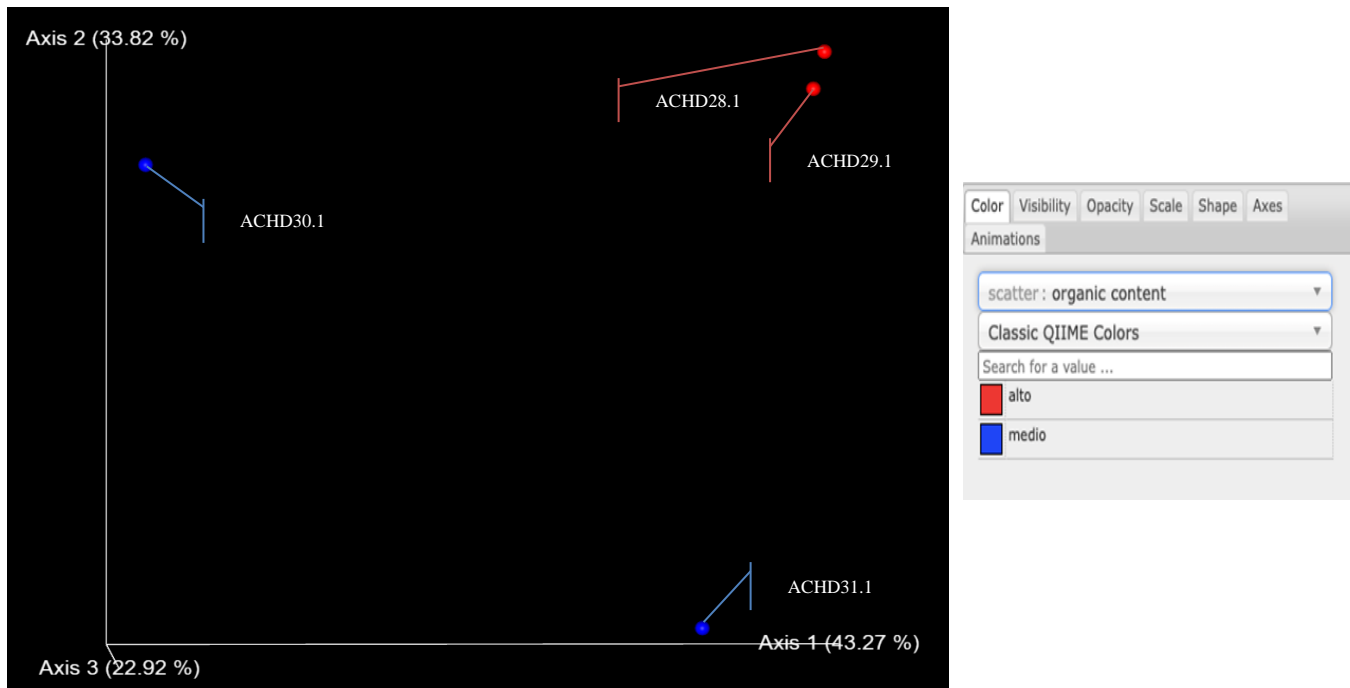
**Figura 21.** Análisis de Coordenadas Principales una medida cuantitativa que relaciona la diversidad de las bacterias presente en suelo con SampleID. Se visualiza el PCoA (Análisis de Coordenadas Principales) Unifrac ponderado cuantitativo para SampleID. El eje Axis 1 mostro 72.43% para la muestra ACHD30.1, el eje Axis 2 el 18.83% la muestra ACHD29.1 y el eje Axis 3 de 8.74% la muestra ACHD 28.1 las cuales representa diferencia significativa en la medida cuantitativa que considera las abundancias relativas de las muestras y los patrones de agrupamiento; estas tres muestras se visualizó la disimilitud de la comunidad bacteriana para SampleID en proporción al porcentaje de PCoA, marcando distancias filogenéticas entre ellas pero con valores cuantitativo significativos aceptables; estas muestras se agrupan en núcleos bastante definidos. Por otra parte, para las muestras ACHD 31.1 del eje Axis 4 no se observó valores significativos en medida cuantitativa de la disimilitud de la comunidad bacteriana para SampleID en proporción al porcentaje de PCoA



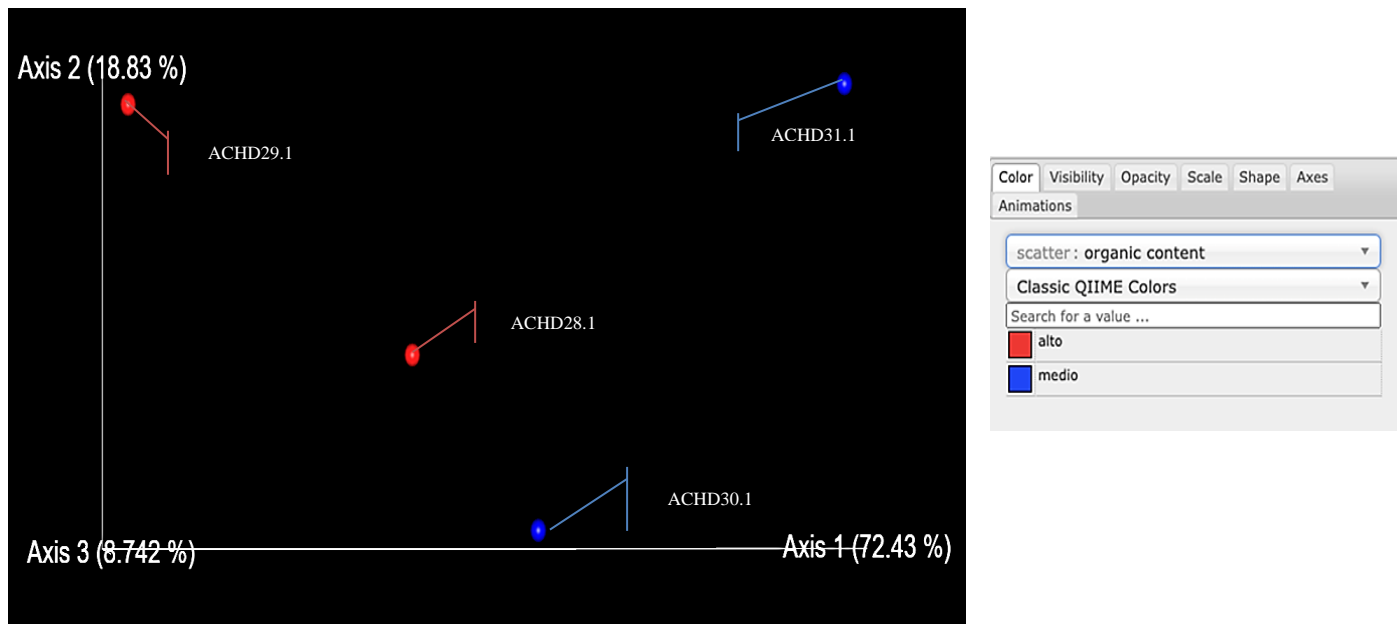
**Figura 22.** Análisis de Coordenadas Principales una medida cualitativa que relaciona la diversidad de las bacterias presente en suelo con la concentración de Magnesio. Se observa el PCoA (Análisis de Coordenadas Principales) Unifrac no ponderado cualitativo para Magnesio, el eje Axis 1 mostro 43.27% para la muestra ACHD30.1 y el eje Axis 2 el 33.82% la muestra ACHD31.1 que representa diferencia significativa en la medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto, medio y bajo de Magnesio en proporción al porcentaje de visualización PCoA aquí las muestras se agrupan en núcleos bastante definidos. Por otra parte, ambas muestras si mostro distancias filogenéticas en PCoA. El eje Axis 3 de 22.92% no se observó muestras, y por último el eje Axis 4 se visualizó las muestras ACHD28.1, ACHD29.1 la cual no son significativos en medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto, medio y bajo de Magnesio en proporción al porcentaje y no mostro distancias filogenéticas en PCoA



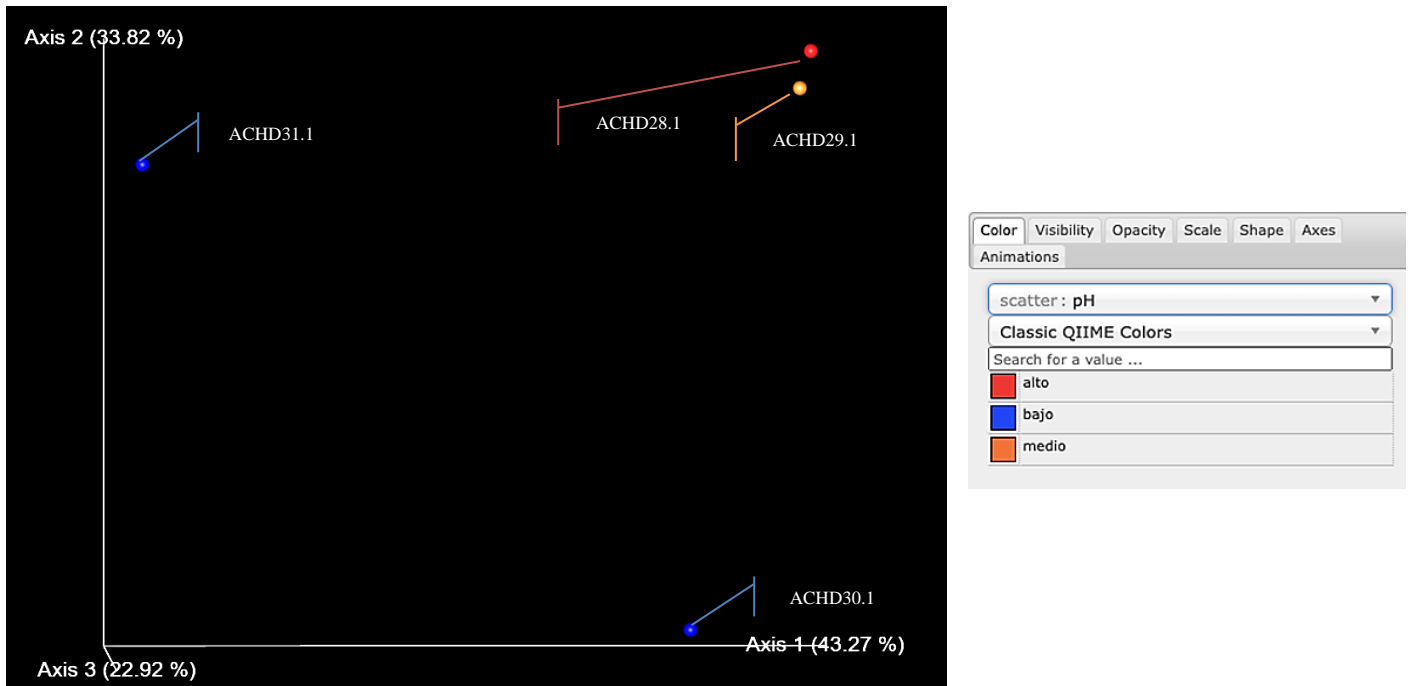
**Figura 23.** Análisis de Coordenadas Principales en medida cuantitativa que relaciona la diversidad de las bacterias presente en suelo con el Magnesio. Se visualiza el PCoA (Análisis de Coordenadas Principales) Unifrac ponderado cuantitativo para Magnesio. El eje Axis 1 mostro 72.43% para la muestra ACHD30.1, el eje Axis 2 el 18.83% la muestra ACHD29.1 y el eje Axis 3 de 8.74% la muestra ACHD 28.1 las cuales representa diferencia significativa en la medida cuantitativa que considera las abundancias relativas de las muestras y los patrones de agrupamiento; estas tres muestras se visualizó la disimilitud de la comunidad bacteriana para niveles alto, bajo y medio de Magnesio en proporción al porcentaje de PCoA, marcando distancias filogenéticas entre ellas pero con valores cuantitativo significativos aceptables; estas muestras se agrupan en núcleos bastante definidos. Por otra parte, para las muestras ACHD 31.1 del eje Axis 4 no se observó valores significativos en medida cuantitativa de la disimilitud de la comunidad bacteriana para niveles alto de Magnesio en proporción al porcentaje de PCoA



**Figura 24.** Análisis de Coordenadas Principales en medida cualitativa que relaciona la diversidad de las bacterias presente en suelo con el contenido orgánico. Se observa el PCoA (Análisis de Coordenadas Principales) Unifrac no ponderado cualitativo para Contenido Organico, el eje Axis 1 mostro 43.27% para la muestra ACHD30.1 y el eje Axis 2 el 33.82% la muestra ACHD31.1 que representa diferencia significativa en la medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de Contenido Organico en proporción al porcentaje de visualización PCoA aquí las muestras se agrupan en núcleos bastante definidos. Por otra parte, ambas muestras si mostro distancias filogenéticas en PCoA. El eje Axis 3 de 22.92% no se observó muestras, y por último el eje Axis 4 se visualizó las muestras ACHD28.1, ACHD29.1 la cual no son significativos en medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de Contenido Organico en proporción al porcentaje y no mostro distancias filogenéticas en PCoA

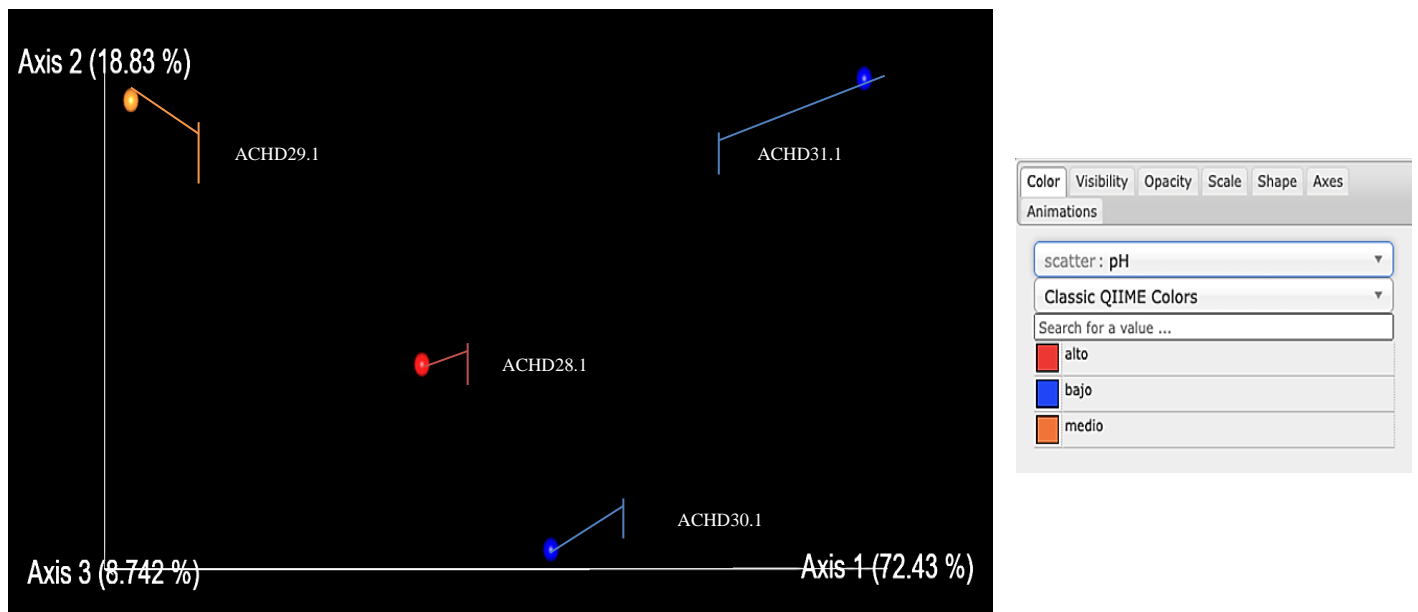


**Figura 25.** Análisis de Coordenadas Principales en medida cuantitativo que relaciona la diversidad de las bacterias presente en suelo con el contenido orgánico. Se visualiza el PCoA (Análisis de Coordenadas Principales) Unifrac ponderado cuantitativo para Contenido Orgánico. El eje Axis 1 mostro 72.43% para la muestra ACHD30.1, el eje Axis 2 el 18.83% la muestra ACHD29.1 y el eje Axis 3 de 8.74% la muestra ACHD 28.1 las cuales representa diferencia significativa en la medida cuantitativa que considera las abundancias relativas de las muestras y los patrones de agrupamiento; estas tres muestras se visualizó la disimilitud de la comunidad bacteriana para niveles alto y medio de Contenido Orgánico en proporción al porcentaje de PCoA, marcando distancias filogenéticas entre ellas pero con valores cuantitativo significativos aceptables; estas muestras se agrupan en núcleos bastante definidos. Por otra parte, para las muestras ACHD 31.1 del eje Axis 4 no se observó valores significativos en medida cuantitativa de la disimilitud de la comunidad bacteriana para nivel medio de Contenido Orgánico en proporción al porcentaje de PCoA

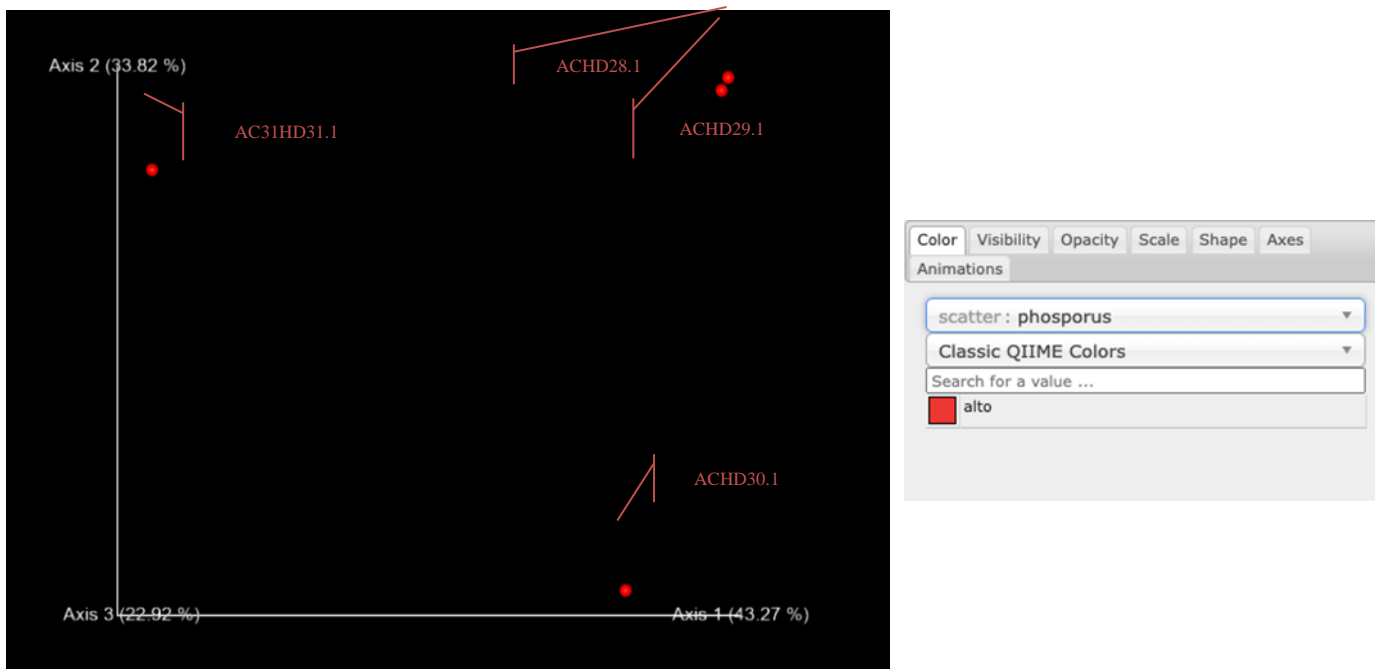


**Figura 26.** Análisis de Coordenadas Principales en medida cualitativa que relaciona la diversidad de las bacterias presente en suelo con el pH. Se observa el PCoA (Análisis de Coordenadas Principales) Unifrac no ponderado cualitativo para pH, el eje Axis 1 mostro 43.27% para la muestra ACHD30.1 y el eje Axis 2 el 33.82% la muestra ACHD31.1 que representa diferencia significativa en la medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de pH en proporción al porcentaje de visualización PCoA aquí las muestras se agrupan en núcleos bastante definidos. Por otra parte, ambas muestras si mostro distancias filogenéticas en PCoA. El eje Axis 3 de 22.92% no se observó muestras, y por último el eje Axis 4 se visualizó las muestras ACHD28.1, ACHD29.1 la cual no son significativos en medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto y medio de pH en proporción al porcentaje y no mostro distancias filogenéticas en PCoA

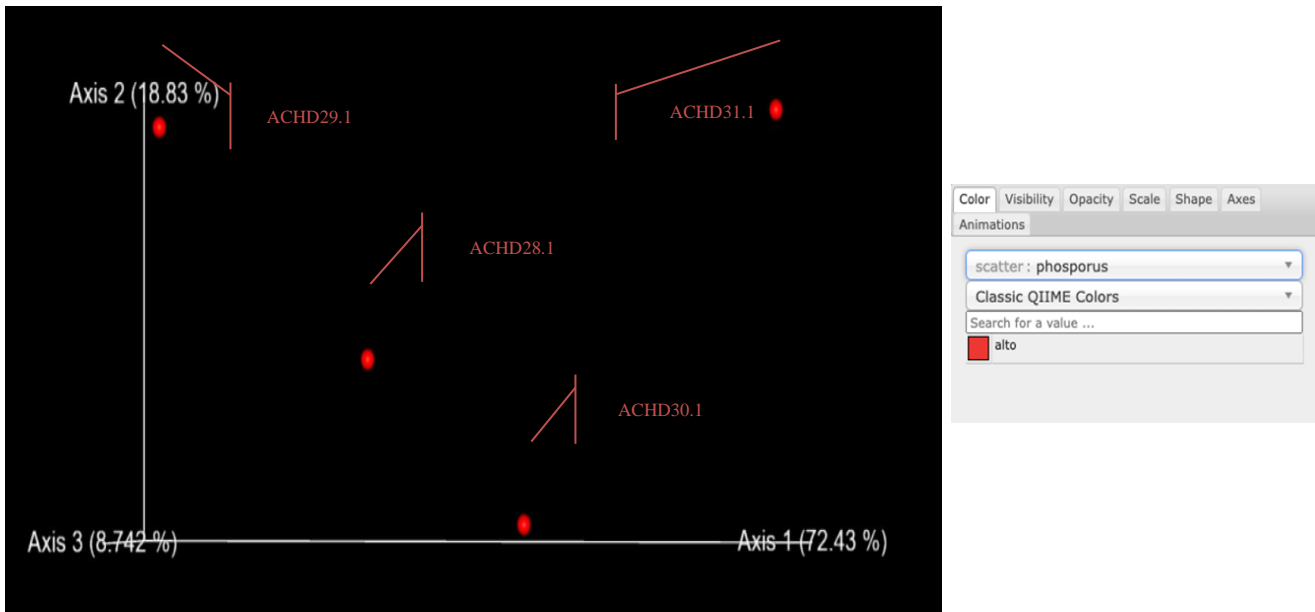




**Figura 27.** Análisis de Coordenadas Principales en medida cuantitativo que relaciona la diversidad de las bacterias presente en suelo con el pH. Se visualiza el PCoA (Análisis de Coordenadas Principales) Unifrac ponderado cuantitativo para pH. El eje Axis 1 mostro 72.43% para la muestra ACHD30.1, el eje Axis 2 el 18.83% la muestra ACHD29.1 y el eje Axis 3 de 8.74% la muestra ACHD 28.1 las cuales representa diferencia significativa en la medida cuantitativa que considera las abundancias relativas de las muestras y los patrones de agrupamiento; estas tres muestras se visualizó la disimilitud de la comunidad bacteriana para niveles alto, bajo y medio de pH en proporción al porcentaje de PCoA, marcando distancias filogenéticas entre ellas pero con valores cuantitativo significativos aceptables; estas muestras se agrupan en núcleos bastante definidos. Por otra parte, para las muestras ACHD 31.1 del eje Axis 4 no se observó valores significativos en medida cuantitativa de la disimilitud de la comunidad bacteriana para nivel bajo de pH en proporción al porcentaje de PCoA



**Figura 28.** Análisis de Coordenadas Principales una medida cualitativa que relaciona la diversidad de las bacterias presente en suelo con el fosforo. Se observa el PCoA (Análisis de Coordenadas Principales) Unifrac no ponderado cualitativo para fosforo, el eje Axis 1 mostro 43.27% para la muestra ACHD30.1 y el eje Axis 2 el 33.82% la muestra ACHD31.1 que representa diferencia significativa en la medida cualitativa de la disimilitud de la comunidad bacteriana para niveles alto de fosforo en proporción al porcentaje de visualización PCoA aquí las muestras se agrupan en núcleos bastante definidos. Por otra parte, ambas muestras si mostro distancias filogenéticas en PCoA. El eje Axis 3 de 22.92% no se observó muestras, y por último el eje Axis 4 se visualizó las muestras ACHD28.1, ACHD29.1 la cual no son significativos en medida cualitativa de la disimilitud de la comunidad bacteriana para nivel alto de fosforo en proporción al porcentaje y no mostro distancias filogenéticas en PCoA



**Figura 29.** Análisis de Coordenadas Principales una medida cuantitativa que relaciona la diversidad de las bacterias presente en suelo con el fosforo. Se visualiza el PCoA (Análisis de Coordenadas Principales) Unifrac ponderado cuantitativo para fosforo. El eje Axis 1 mostro 72.43% para la muestra ACHD30.1, el eje Axis 2 el 18.83% la muestra ACHD29.1 y el eje Axis 3 de 8.74% la muestra ACHD 28.1 las cuales representa diferencia significativa en la medida cuantitativa que considera las abundancias relativas de las muestras y los patrones de agrupamiento; estas tres muestras se visualizó la disimilitud de la comunidad bacteriana para niveles alto de fosforo en proporción al porcentaje de PCoA, marcando distancias filogenéticas entre ellas pero con valores cuantitativo significativos aceptables; estas muestras se agrupan en núcleos bastante definidos. Por otra parte, para las muestras ACHD 31.1 del eje Axis 4 no se observó valores significativos en medida cuantitativa de la disimilitud de la comunidad bacteriana para nivel alto de fosforo en proporción al porcentaje de PCoA

Por otro lado, las pruebas estadísticas de PERMANOVA para diversidad beta, se realizaron con el comando qiime diversity beta-group-significance.

QIIME2 utilizó análisis permutacional de varianza (PERMANOVA) para validar todas las pruebas de diversidad beta. Nuevamente, este es un análisis multivariante no paramétrico que compara grupos de objetos bajo la hipótesis nula de que la dispersión de cada grupo es equivalente para todos los grupos que se comparan; rechazar la hipótesis nula significaría que los grupos están dispersos de manera diferente. Para hacerlo, sería necesario obtener un valor P significativo (menos de 0,05).

La prueba permutacional de análisis de varianza multivariada (PERMANOVA) (Grafico 3) indica que no hay una diferencia en la composición de taxa entre los tipos de suelo ( $p=0.666$ ; con 999 permutaciones) (Tabla, 16, 17). Así mismo, la prueba estadística sugiere que no existe una diferencia significativa en la composición taxonómica entre la condición de los suelos (intervenido y no intervenido), ( $p=0.341$ ; con 999 permutaciones) (Grafico, 4) (Tabla, 18, 19).

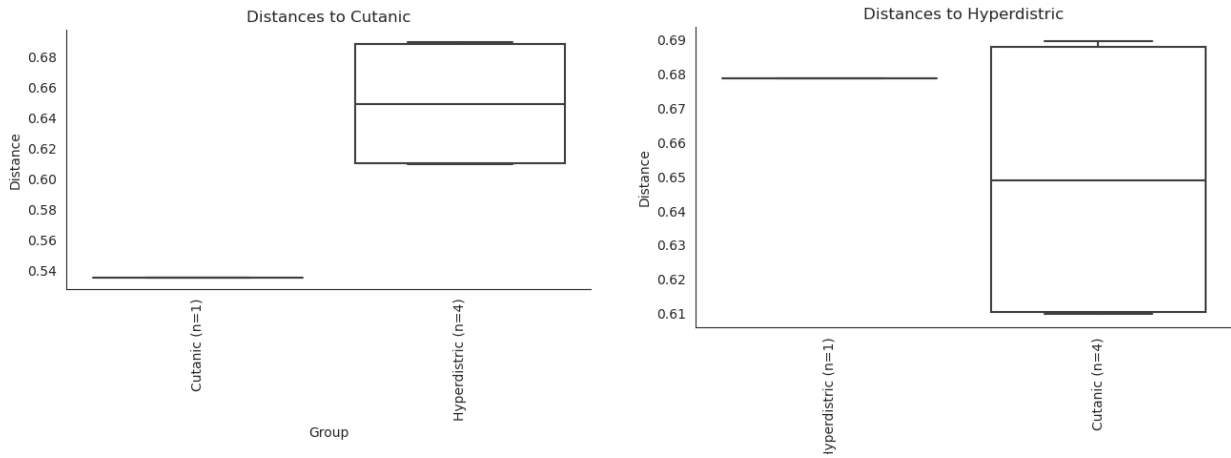
Tabla 16. **Permanova Tipo de Muestra**

### Overview

	PERMANOVA results
method name	PERMANOVA
test statistic name	pseudo-F
sample size	4
number of groups	2
test statistic	1.26495
p-value	0.341
number of permutations	999

## Group significance plots

[Download raw data as TSV](#)



**Grafico 3. Permanova Tipo de Muestra**

Tabla 17. Permanova Tipo de Muestra

## Pairwise permanova results

[Download CSV](#)

		Sample size	Permutations	pseudo-F	p-value	q-value
Group 1	Group 2					
Cutanic	Hyperdistic	4	999	1.264951	0.3	0.3

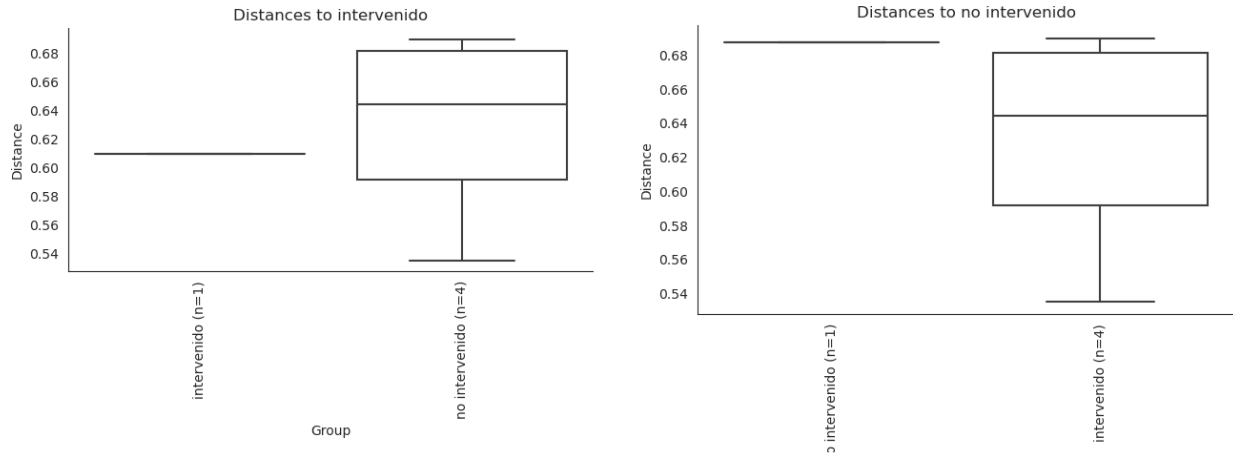
Tabla 18. Resultado Permanova Condición

## Overview

PERMANOVA results	
method name	PERMANOVA
test statistic name	pseudo-F
sample size	4
number of groups	2
test statistic	0.890069
p-value	0.666
number of permutations	999

## Group significance plots

[Download raw data as TSV](#)



**Grafico 4. Permanova Condición**

**Tabla 19. Permanova Condición**

## Pairwise permanova results

[Download CSV](#)

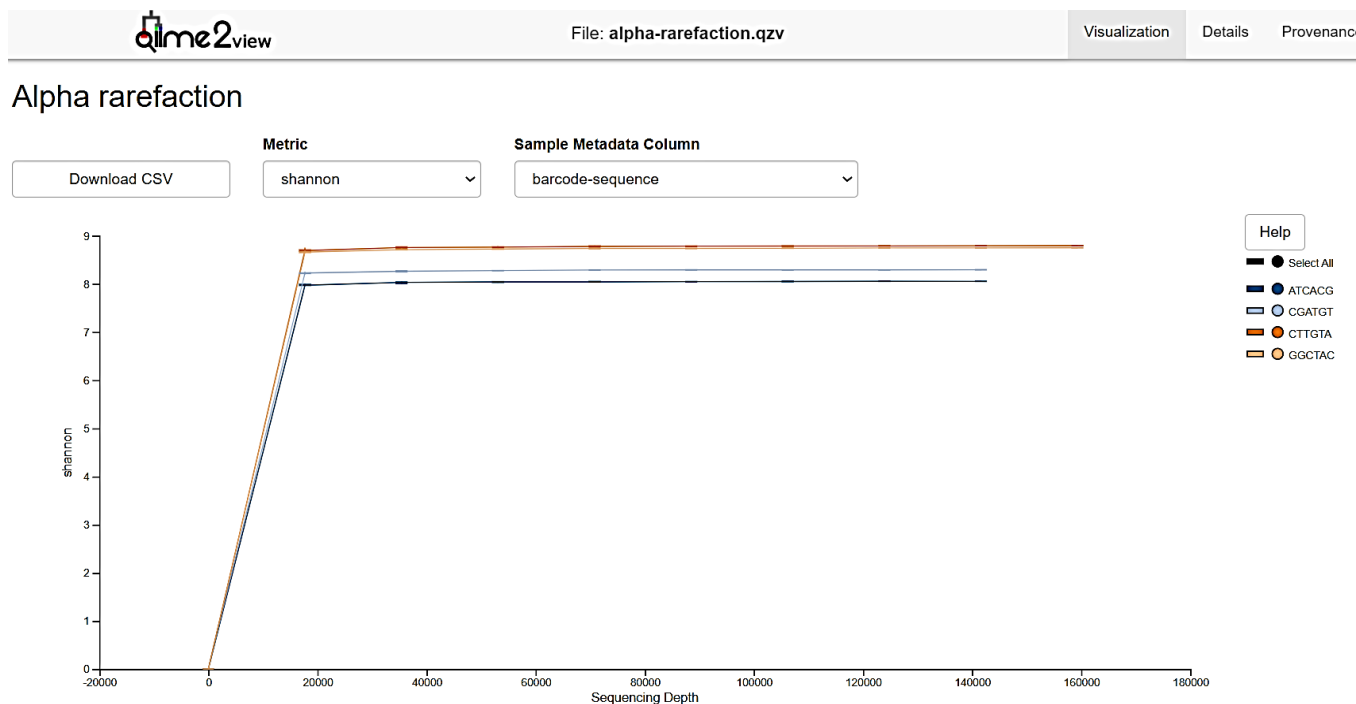
Group 1	Group 2	Sample size	Permutations	pseudo-F	p-value	q-value
intervenido	no intervenido	4	999	0.890069	0.651	0.651

## Alfa Rarefacción

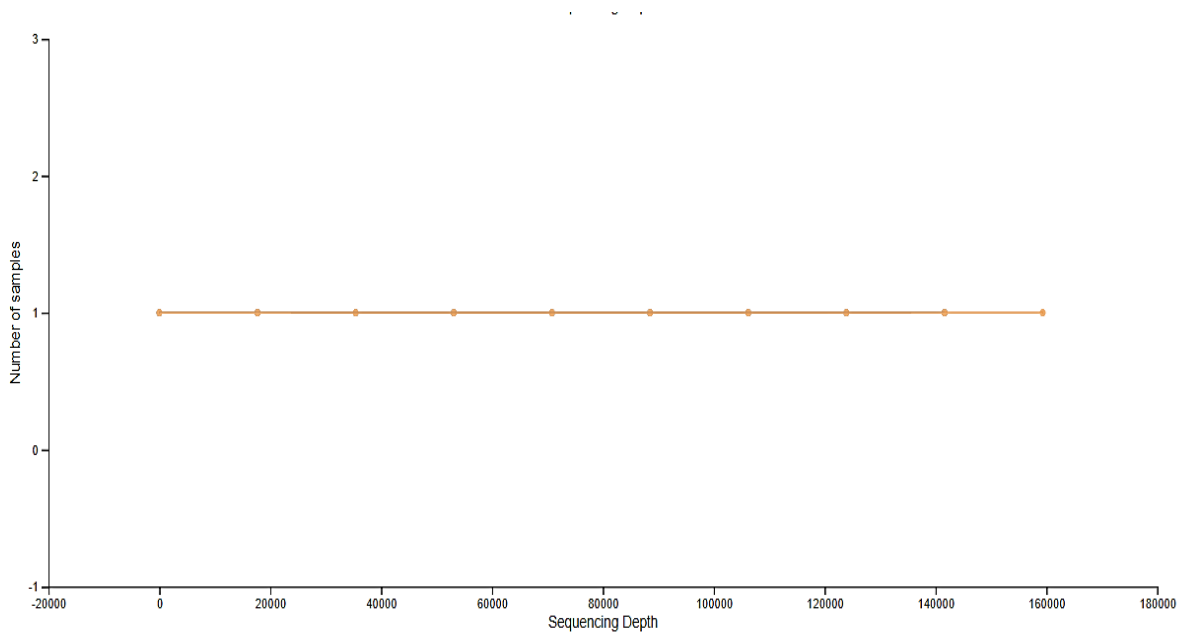
Específicamente la rarefacción es el proceso de generación de la relación entre el número de especies vs el número de individuos en una o más muestras (Stevens 2009). Esta corrección por el número de individuos nos permite la comparación directa de la riqueza de todas las muestras que inicialmente tenían diferente tamaño.

Esta sección de resultado nos muestra la diversidad alfa en función de la profundidad de muestreo utilizando el visualizador (Grafico, 5). Este análisis nos calcula 3 métricas de diversidad alfa a múltiples profundidades de muestreo. Cada paso de profundidad de muestreo, se generó 10 puntos en el gráfico y se calculó las métricas de diversidad para los 4 cuatros muestras en el gráfico. Los valores de diversidad promedio se trazaron para cada muestra en

cada profundidad de muestreo uniforme, y las muestras se agrupo de acuerdo con los metadatos como visualización resultante. Los metadatos de muestra se proporcionan con el parámetro. qiime diversity alpha-rarefaction--p-min-depth--p-max-depth--p-iterations--m-metadata-fil



**Grafico 5.** Columna SampleMetadata Barcode sequence – Metric Shannon. En esta visualización se determina si la riqueza de las 4 muestras se ha observado o secuenciado completamente la cual da como resultado que las líneas de forma nivelada (es decir, acercarse a una pendiente de cero) a cierta profundidad de muestreo a lo largo del eje x, eso significa que las muestras estudiadas es efectiva para la observación de características debido a la recopilación de suficientes secuencias de dichas muestras y cero error de secuencia en los datos es decir que la riqueza de las muestras se han observado o secuenciado completamente



**Grafico 6.** Rarefacción Alfa. Esta visualización de resultado nos muestra la profundidad de muestreo. Si una profundidad de muestreo dada es mayor que la frecuencia total de una muestra (es decir, el número de secuencias que se obtuvieron para la muestra), no es posible calcular la métrica de diversidad para la muestra a profundidad de muestreo.

El gráfico superior (Grafico, 6) de esta visualización es importante cuando se agrupan muestras por metadatos. Ilustra el número de muestras que permanecen en cada grupo cuando la tabla de características se enrarece a cada profundidad de muestreo. Si una profundidad de muestreo dada es mayor que la frecuencia total de una muestra (es decir, el número de secuencias que se obtuvieron para la muestra), no es posible calcular la métrica de diversidad para la muestra a profundidad de muestreo. Si muchas de las muestras en un grupo tienen frecuencias totales más bajas que, la diversidad promedio presentada para ese grupo en la gráfica superior no será confiable porque se habrá calculado en relativamente pocas muestras. Por lo tanto, al agrupar muestras por metadatos, es esencial mirar la gráfica inferior para asegurarse de que los datos presentados en la gráfica superior sean confiables.



## Análisis Taxonómico

En las siguientes secciones (Figura, 30) exploramos la composición taxonómica de las muestras y, de nuevo, la relacionaremos con los metadatos de muestra. El primer paso en este proceso es asignar taxonomía a las secuencias en nuestro artefacto QIIME 2. **Confidencialidad 0.90**

Feature ID	Taxon	Confidence
9b34325ae0c300c4e0c2c48d526c1	k__Bacteria	0.9708429377780861
9b6824b1e8380844294c13c17054a	k__Bacteria	0.9690055165943757
7043c9ee1dab3ee00c6401f72696e	k__Bacteria; p__Actinobacteria; c__Acidimicrobia; o__Acidimicrobiales; f__EB1017; g__s__	0.997279175222591
0896b4668b6347a1adbb96d42e91175	k__Bacteria; p__Actinobacteria; c__Actinobacteria; o__Actinomycetales	0.7246753383457767
780c85cc415268a1baec452b2604394	k__Bacteria; p__Actinobacteria; c__Actinobacteria; o__Actinomycetales	0.7392891003405023
a38e84f1a7533534a77a844c31f4847	k__Bacteria; p__Actinobacteria; c__Actinobacteria; o__Actinomycetales	0.750669241940217
9e73120ad74d39f14079200448891b	k__Bacteria; p__Bacteroidetes	0.7470435442658225
1d369f68dafe926419e963d585240	k__Bacteria; p__Chloroflexi; c__TK17; o__f__g__s__	0.7500410231842226
d2b090d4ce328866915e26b219e60f	k__Bacteria; p__Firmicutes	0.7434050594255005
4d5d784a310051c7c5c19a0fc18817ea	k__Bacteria; p__Firmicutes; c__Bacilli; o__Bacillales; f__Bacillaceae	0.7759062819638024
20a5732341a08a18d3a006100df12784	k__Bacteria; p__Proteobacteria	0.7222765387252399
310e88be4e66581b0a85aa291cd15e89	k__Bacteria; p__Proteobacteria	0.7222717876712511
77b4d2e13ed1eca5a71481f21d84852f	k__Bacteria; p__Proteobacteria	0.959509435310579
960747dc66f7f09d9e979e444e6e585	k__Bacteria; p__Proteobacteria	0.7014517609273534
9aacfee190119582ac05a4a3484bc08	k__Bacteria; p__Proteobacteria	0.7241686455828479
9ed3fb2f40ab595e4e31abc042378	k__Bacteria; p__Proteobacteria	0.933948754756381
a0ab738bd64876903035c7d37bd9d19	k__Bacteria; p__Proteobacteria	0.7127444883049152
bd34f980e945d9f35ac2c4596d8e10	k__Bacteria; p__Proteobacteria	0.7189228677163034
bdad3371499c02d88043dcb3aa4cc872	k__Bacteria; p__Proteobacteria	0.7490594879210276
d87d1c5eab754936aef45da5c5e7a2	k__Bacteria; p__Proteobacteria	0.7155938645058743
e874cb4054c5db1acb1d1bd7b98fae	k__Bacteria; p__Proteobacteria	0.9050795381956759
959a5527da2dba1023b8f86e99e36	k__Bacteria; p__Proteobacteria	0.70342761421114516
9f67ab16d5e18b7eb30c2e80990c298	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria	0.9634364161220511
e8aa00d9e03a9c28742a847595da199	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria	0.956263191526681
514900507d35db28420af9f3c95a9	k__Bacteria; p__Verrucomicrobia	0.742014079693015
113a8f09f75767a78c987370b4d4455	k__Bacteria; p__WS6; c__SC72; o__A-2AF; f__g__s__	0.7334975012801443
2c2157154dca9180988a1894c8a153	k__Bacteria; p__WS6; c__SC72; o__A-2AF; f__g__s__	0.7399055776175787

**Figura 30.** Composición taxonómica de las muestras. El análisis del microbiota de las muestras de suelo Acrisoles Haplico se evaluó a los niveles taxonómicos de dominio, phylum, clase, orden, familia, género y especie. La asignación taxonómica a nivel de phylum reveló un total de 8 phylum, 11 clase, 11 orden, 11 familia, 11 género y especie para los cuatro grupos experimentales, aunque solo muestra algunas identificadas y otras no. Si nos fijamos se observa.

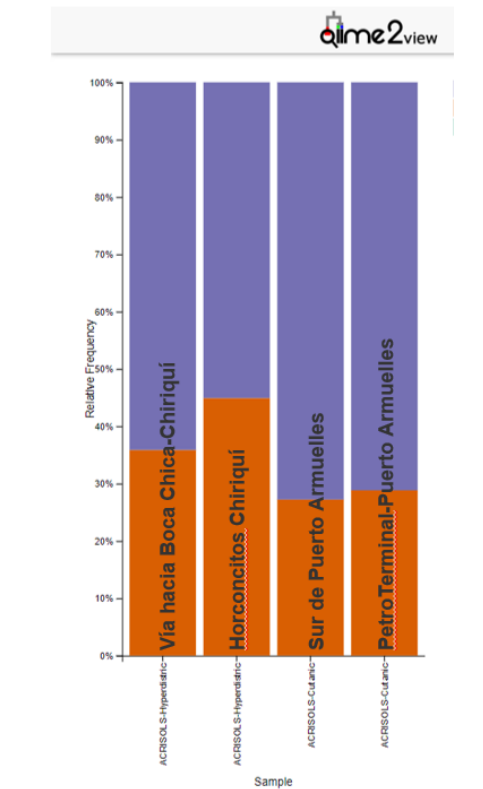


Tabla 20. Tabla de valores para dominio

index,	Unassigned,	k Archaea,	k Bacteria,
ACHD28.1,	48161.0	2.0,	118878.0
ACHD29.1,	44363.0,	0.0,	118583.0
ACHD30.1,	67887.0,	0.0,	83243.0
ACHD31.1,	55902.0,	0.0,	99917.0

### Grafico 7. Taxonomía Nivel 1

El análisis taxonómico del microbiota de suelo de los cuatros muestras secuenciadas mostró un total de 55 % de bacterias y 45% taxones sin asignar para muestras ACHD 30.1, la muestras ACHD 31.1 62% de bacteria y 38% sin asignar, muestra ACHD 28.1 tiene 70% de bacteria y 30% taxones sin asignar, muestra ACHD 29.1 tiene 71% de bacteria y 29 % sin asignar. Se observa la asignación taxonómica en función del Dominio. Un total de 2 dominio fueron identificados (Bacteria y Archeae). Se visualizo que la principal categoría determinada es Bacteria para todas las muestras y solo una muestra ACHD 28.1 marco presencia de Archaea de suelo Acrisoles Háptico.

Hover over the plot to learn more

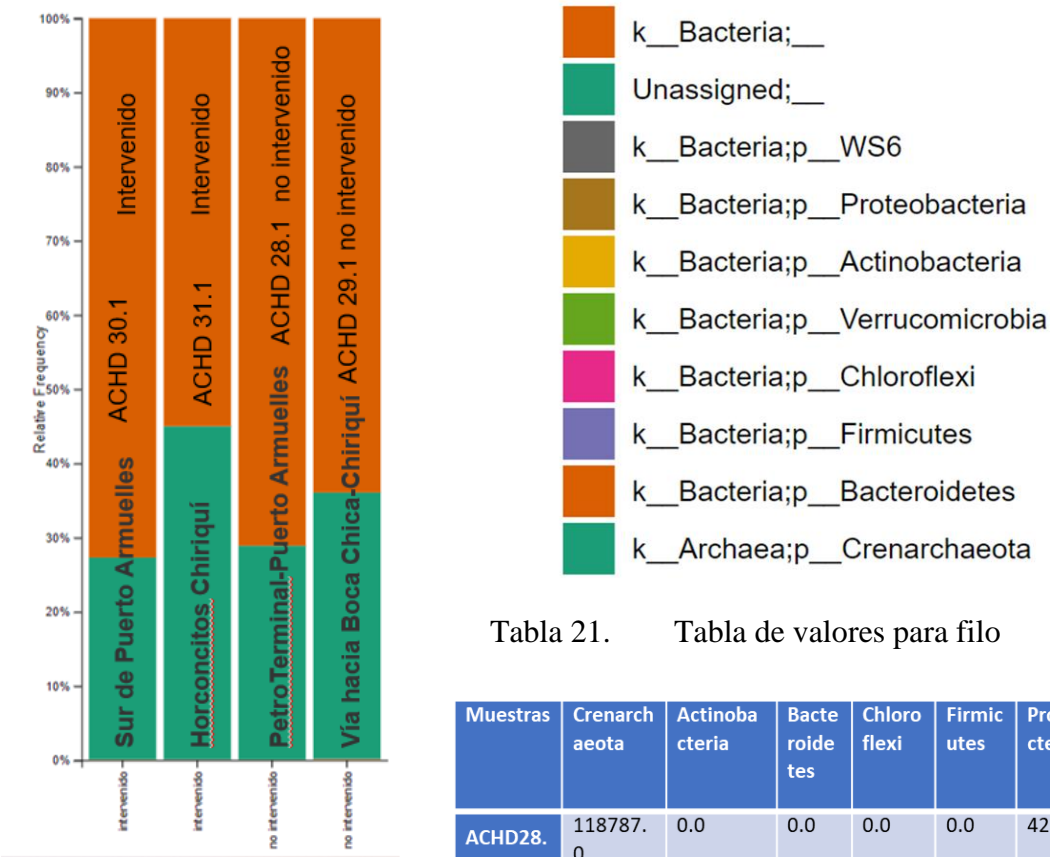


Tabla 21. Tabla de valores para filo

Muestras	Crenarchaeota	Actinobacteria	Bacteroidetes	Chloroflexi	Firmicutes	Proteobacteria	Verrucomicrobia	WS6
ACHD28.1	118787.0	0.0	0.0	0.0	0.0	42.0	0.0	49.0
ACHD29.1	118456.0	33.0	0.0	10.0	2.0	19.0	13.0	50.0
ACHD30.1	83118.0	2.0	2.0	0.0	2.0	21.0	0.0	98.0
ACHD31.1	99627.0	85.0	0.0	0.0	0.0	80.0	0.0	25.0

### Grafico 8. Taxonomía Nivel 2

El análisis taxonómico del microbiota de suelo de los cuatro muestras secuenciadas mostró un total de 55 % de bacterias y 45% taxones sin asignar para muestras ACHD 30.1, muestras ACHD 31.1 62% de bacteria y 38% sin asignar, muestra ACHD 28.1 tiene 70% de bacteria y 30% taxones sin asignar, muestra ACHD 29.1 tiene 71% de bacteria y 29 % sin asignar. Se observa la asignación taxonómica en función de phylum. Un total de 8 phylum fueron identificados (Actinobacteria, Bacteroidetes, Firmicutes, Proteobacteria, Verrucomicrobia, Chloroflexi, Crenarchaeota, WS6).

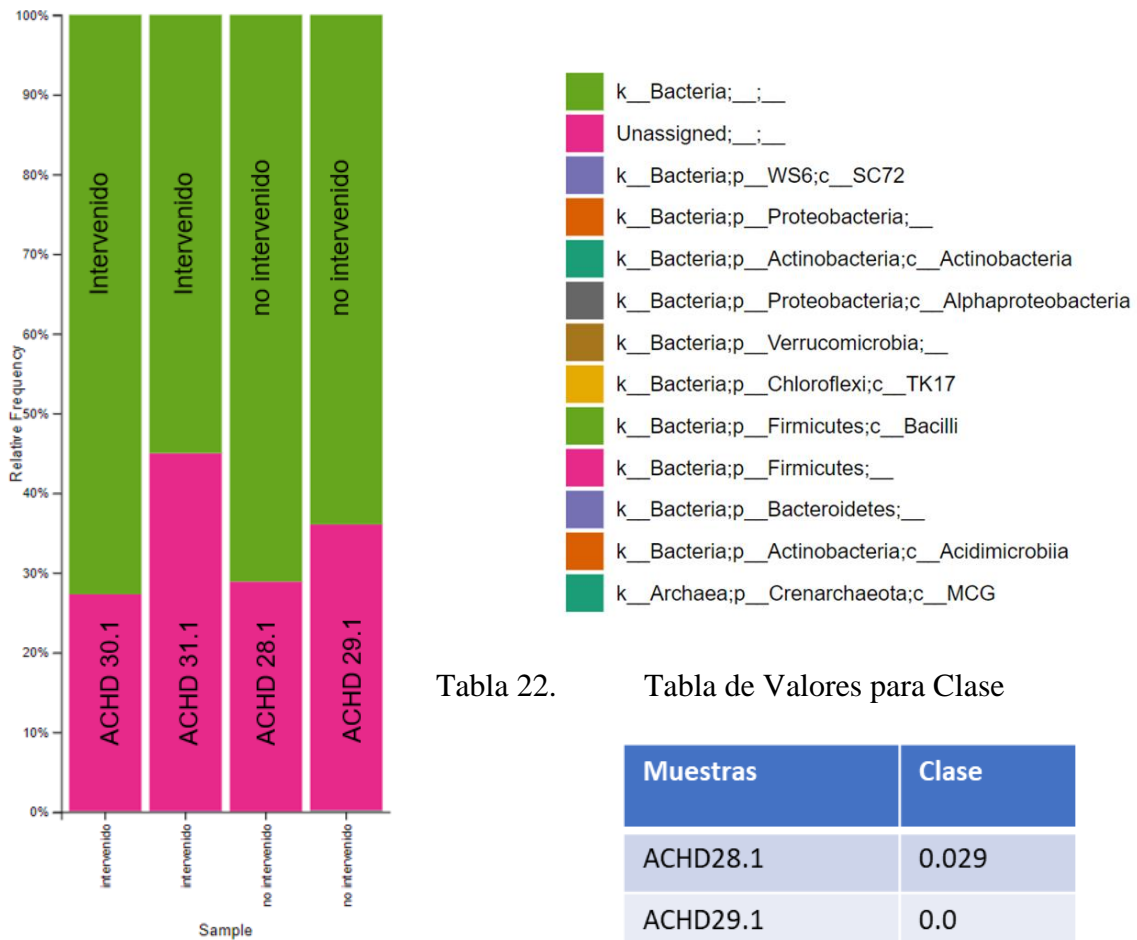
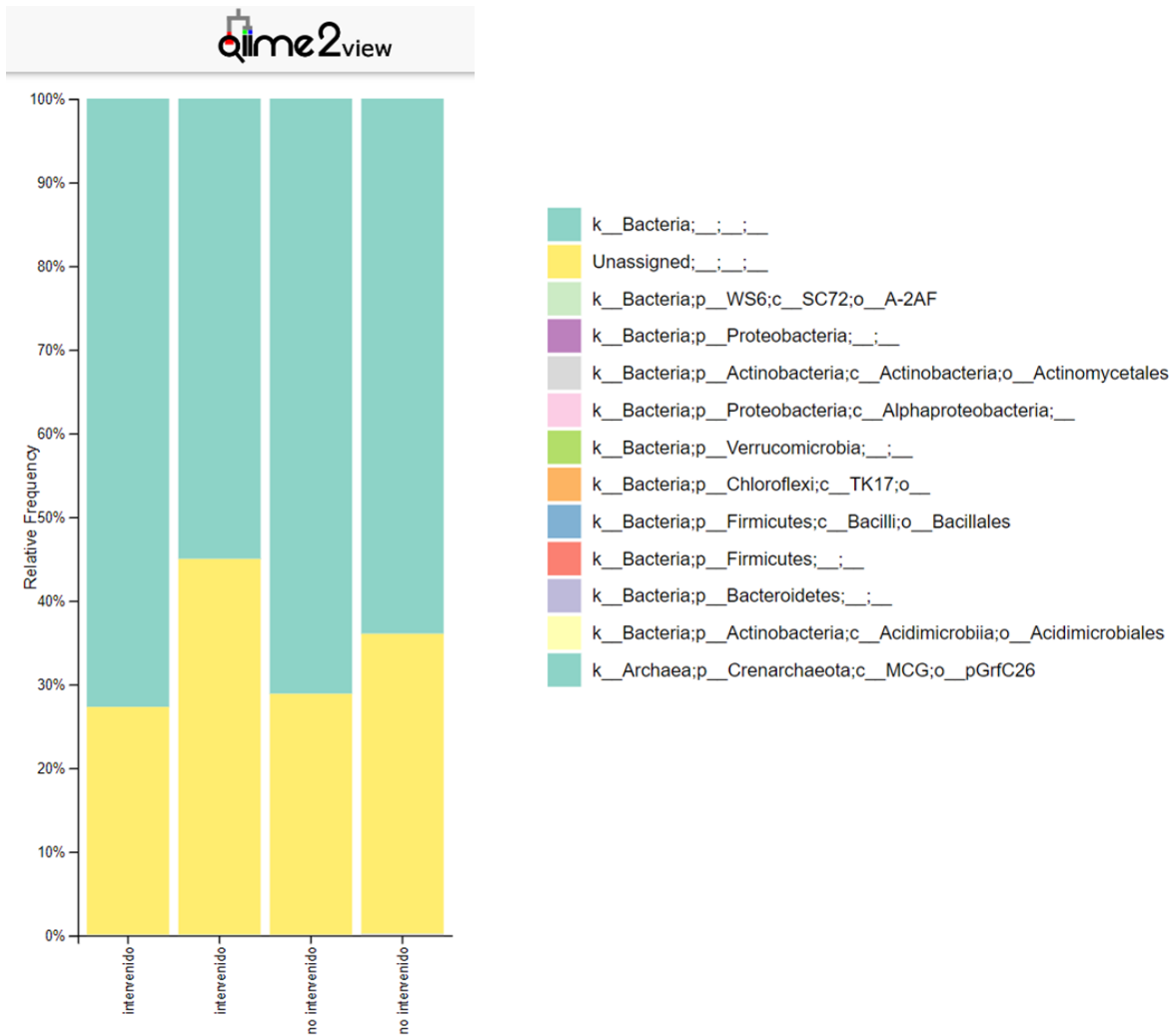


Tabla 22. Tabla de Valores para Clase

Muestras	Clase
ACHD28.1	0.029
ACHD29.1	0.0
ACHD30.1	0.065
ACHD31.1	0.080

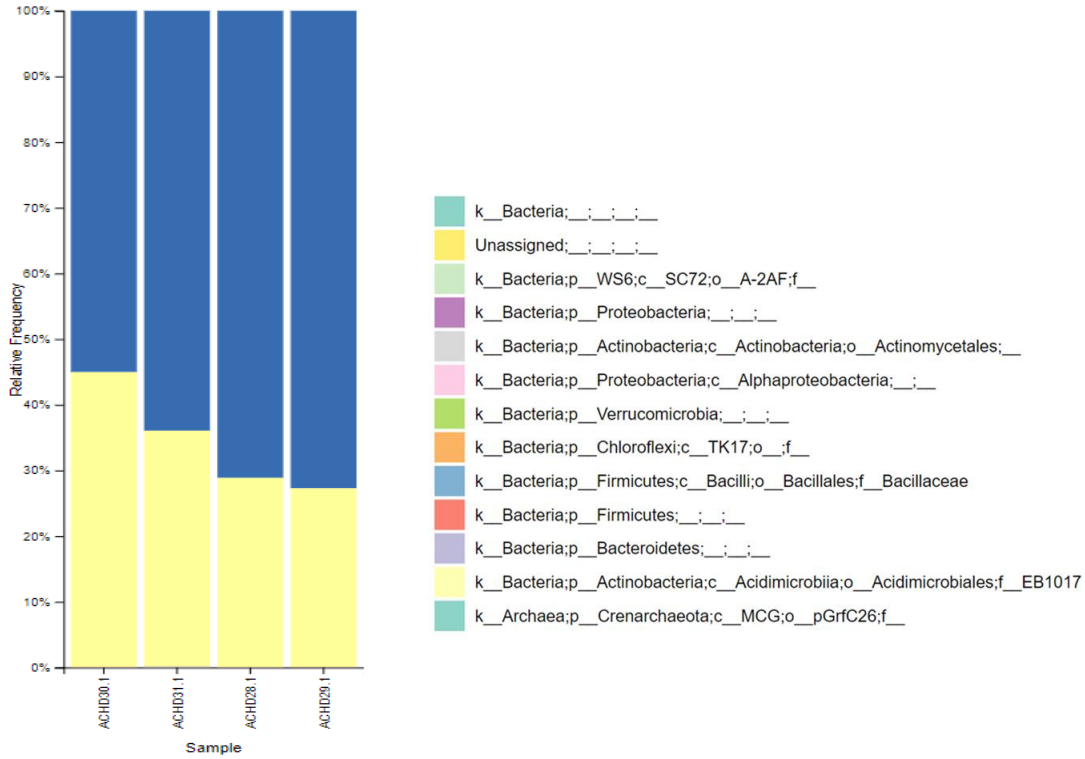
### Grafico 9. Taxonomía Nivel 3

El análisis taxonómico del microbiota de suelo de los cuatro muestras secuenciadas mostró un total para muestras ACHD 30.1 de 55 % de bacterias y 45% taxones sin asignar, muestras ACHD 31.1 62% de bacteria y 38% sin asignar, muestra ACHD 28.1 tiene 70% de bacteria y 30% taxones sin asignar, muestra ACHD 29.1 tiene 71% de bacteria y 29 % sin asignar. Se observa la asignación taxonómica en función de clase. Un total de 11 clases algunas identificadas otras no (Actinobacteria, Alphaproteobacteria, TK17, Bacilli, Acidimicrobiia, MCG y 4 taxon de clase no identificadas). Observamos que la diferencia de clase visualizado es tan baja que no se puede dar porcentaje.



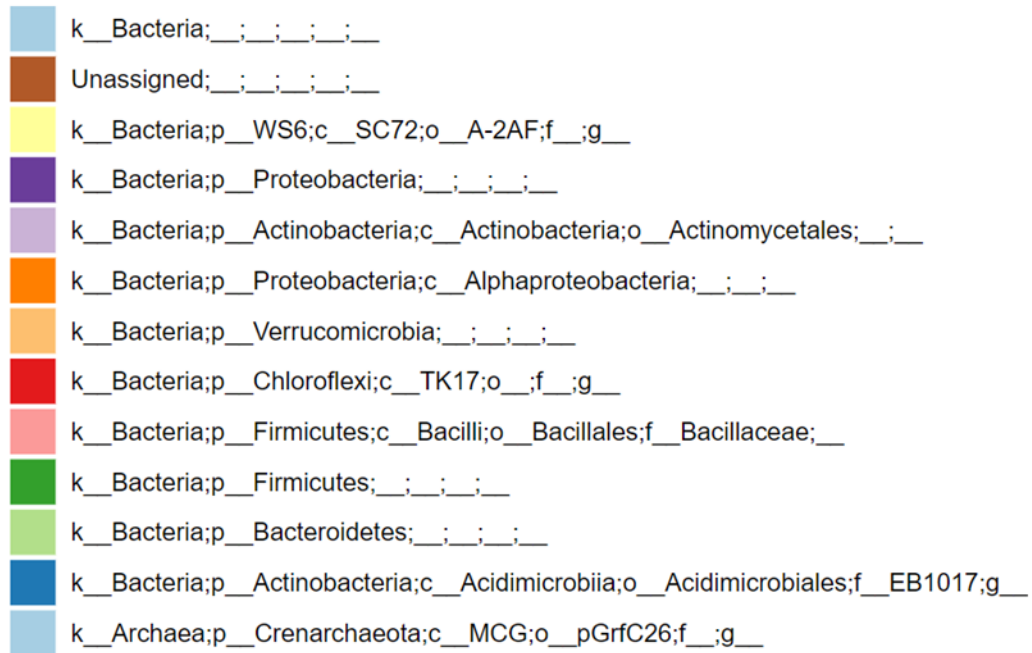
#### **Grafico 10. Taxonomía Nivel 4**

El análisis taxonómico del microbiota de suelo de los cuatros muestras secuenciadas mostró un total para muestras ACHD 30.1 de 55 % de bacterias y 45% taxones sin asignar, muestras ACHD 31.1 62% de bacteria y 38% sin asignar, muestra ACHD 28.1 tiene 70% de bacteria y 30% taxones sin asignar, muestra ACHD 29.1 tiene 71% de bacteria y 29 % sin asignar. Se observa la asignación taxonómica en función de orden. Un total de 11 orden algunos identificados y otros no (A-2AF, Actinomycetales, Bacillales, Acidimicrobiales, GrfC26 y 6 no identificado). Observamos que la diferencia de orden visualizado es tan baja que no se puede dar porcentaje.



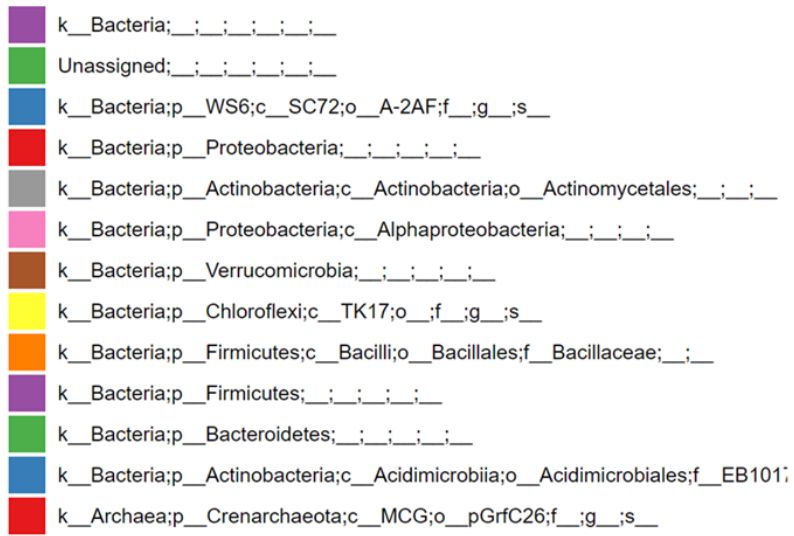
**Grafico 11. Taxonomía Nivel 5**

El análisis taxonómico del microbiota de suelo de los cuatros muestras secuenciadas se observa la asignación taxonómica en función de familia. Un total de 11 familia fueron identificadas (Bacillaceae, EB1017 y 9 no identificada). Observamos que la diferencia de familia visualizada es tan baja que no se puede dar porcentaje.



**Grafico 12. Taxonomía Nivel 6**

El análisis taxonómico del microbiota de suelo de los cuatros muestras secuenciadas se observa la asignación taxonómica en función de género. Un total de 11 género fueron identificados. Observamos que la diferencia de género visualizado es tan baja que no se puede dar porcentaje.



**Grafico 13. Taxonomía Nivel 7**

El análisis taxonómico del microbiota de suelo de los cuatros muestras secuenciadas se observa la asignación taxonómica en función de especie. Un total de 11 especie fueron identificados. Observamos que la diferencia de especie visualizado es tan baja que no se puede dar porcentaje.



# **Discusión**

## Discusión

En el campo de la metagenómica el análisis del microbioma y microbiota ha ganado una notable importancia en los últimos años. Cada vez existen más estudios que exploran el microbiota de los suelos y su relación con caracteres de interés para los sectores de la biología. La metagenómica es el reciente método de estudio de omicas que se está introduciendo en los últimos años en el sector de la Genética, Biología Molecular, Microbiología para identificar nuevas y conocidas especies procariota no cultivadas en tipos de suelo. Existen diversos estudios sobre cómo las condiciones fisicoquímicas por cambios naturales o del hombre afectan las comunidades bacterianas presentes en el suelo. Este estudio constituye unas de las primeras investigaciones de metagenómica en suelos Acrisoles Háptico de Panamá. El objetivo principal es la relación abundante o escasa de diversidad del microbiota (procariota) en suelo Acrisoles Háplico cutanico y hiperdistrico y su posible efecto con los parámetros fisicoquímicos (Van Reeuwijk, 2002).

Por otra parte, los resultados obtenidos de la secuenciación metagenómica de amplicón 16S ARNr serán clave para aportar nuevos registros, ya que no hay referencia del microbiota en suelo Acrisoles Hápticos cutanico y hiperdistrico en Panamá que haya utilizado el método de secuenciación de segunda generación para amplificar esta secuencia de genes. Por ese motivo, es importante este estudio, ya que nos va a permitir conocer la comunidad microbiana y valorar cómo los factores fisicoquímicos afectan a la composición, abundancia y diversidad del microbiota de suelo.

En cuanto a la alfa diversidad, se refiere a la diversidad filogenética de Faith, la cual muestra la suma de las longitudes de las ramas del árbol filogenético que conecta todas las especies en las muestras que nos lleva al índice de Pielou que es la forma de medir cómo las especies se distribuyen uniformemente en una comunidad. La uniformidad en las abundancias (Pielou evenness) la visualización de `Evenness_statistics.qzv`.

Se observaron diferencias significativas de riqueza y diversidad en las muestras. Dado que el grupo de suelos Acrisoles Háplico mostró un mayor índice de diversidad  $H$  y, por tanto, una mayor riqueza de especies ( $H = 2,39$ ) para cuantificar la biodiversidad específica. La prueba de Kruskal-Wallis manifestó que ninguna de las diferencias entre grupos de suelos fue significativa con un  $p\text{-value} = 0,1213$ . Este método se utilizó para poder ver si nuestros resultados de riqueza

y abundancia eran significativos, a nivel de alfa diversidad. La prueba de Kruskal-Wallis es un enfoque paramétrico, ya que nuestros datos cumplían el supuesto de normalidad para poder aplicar un análisis de las varianzas (PERMANOVA).

En el análisis del índice de beta diversidad según las distancias de Bray-Curtis concluyó que ninguna de las comparaciones entre grupos era estadísticamente significativa según la composición de las comunidades microbiana. Visualizamos los resultados con la matriz de distancia calculada a partir del índice de disimilitud de Bray-Curtis por el método de análisis de coordenadas principales PCoA, y finalmente graficamos los resultados. El análisis de coordenadas principales se utiliza para observar la similitud o disimilitud entre muestras. Cuando más similares sean las muestras respecto a sus características, más cerca entre sí estarán en la gráfica de ordenación de PCoA. Para ver la significancia de la beta diversidad se realizó un PERMANOVA. El PERMANOVA nos permitió medir si los grupos de muestras eran significativamente diferentes entre sí mediante una prueba estadística centrada en las permutaciones. La prueba permutacional de análisis de varianza multivariada (PERMANOVA) indico que no hay una diferencia en la composición de taxa entre los tipos de suelo ( $p=0.666$ ; con 999 permutaciones). Así mismo, la prueba estadística sugiere que no existe una diferencia significativa en la composición taxonómica entre la condición de los suelos (intervenido y no intervenido ( $p=0.341$ ; con 999 permutaciones).

En este estudio, el análisis taxonómico del microbiota de suelo mostró un total de 55 % de bacterias para las muestras ACHD 30.1, 62% de bacteria para las muestras ACHD 31.1, 70% para la muestra ACHD 28.1, 71% para la muestra ACHD 29.1. Se observaron un total de 4,298 perteneciente al dominio Bacteria; 4 bacterias del phylum Actinobacteria de la clase Acidimicrobiia y orden Actinomycetales. 1 bacteria del phylum Bacteroidetes, Chloroflexi; 2 bacteria del phylum Firmicutes la cual 1 pertenece a la clase Bacilli, orden Bacillales y familia Bacillaceae. 14 bacterias del phylum Proteobacteria la cual 2 pertenece a la clase Alphaproteobacteria. 1 bacteria del phylum Verrucomicrobia, 12 bacterias perteneciente al phylum WS6 ahora conocidas como Dojkabacteria es una clase candidata de bacterias recientemente propuesta, previamente conocido como WS6, 1 Archaea del phylum Crenarchaeota y 1,821 sin asignar.

En esta investigación los resultados de la secuenciación metagenómica de Illumina de las muestras ACHD28, ACHD29, ACHD30, ACHD31. Son comparables con el trabajo realizado por (Kheng Goh, et al., 2020) quienes analizaron la composición y diversidad de las comunidades procariota y eucariota del suelo en un lugar específico de Malasia con condición físicas muy distinta a Panamá, pero ellos determinando los factores biológicos y fisicoquímicos del suelo. Dando como resultados que el suelo de Blenheim con un pH alto y calcio tiene una mayor diversidad de procariotas y eucariotas en comparación con el suelo Bernam. Obteniendo también en el suelo de Blenheim grandes abundancias de raros taxones bacterianos metabólicamente diversos y versátil, taxones bacterianos que aumentaron con la introducción de agentes biocontrol, bacterias potencialmente supresoras de enfermedades y flagelados bacterivoros. Por el contrario, el suelo de Bernam se caracterizó predominantemente por taxones bacterianos potencialmente inducibles por enfermedades. Una combinación de elementos abióticos y bióticos podría ser esencial para impulsar el microbioma del suelo supresor de enfermedades hacia Ganoderma BSR en suelo Blenheim.

A comparación del estudio de suelo de (Pitblado, et al., 2021) con esta investigación el análisis los datos obtenidos de un suelo a largo plazo para determinar los efectos de la remoción de materia orgánica sobre los factores abióticos del suelo, y el impacto resultante sobre la diversidad bacteriana 10 años después de la reforestación. Al analizar la diversidad beta, encontramos que tanto la ubicación geográfica como la profundidad del suelo estaban asociadas con diferencias en la diversidad entre las comunidades bacterianas del suelo. Un análisis estadístico adicional también reveló relaciones significativas entre la profundidad del suelo y los factores abióticos. Se asoció un mayor contenido de carbono orgánico, nitrógeno y humedad en el suelo con las muestras tomadas de la capa superior del suelo orgánico, y los niveles de pH del suelo fueron más ácidos en las muestras de suelo orgánico en comparación con las muestras de suelo mineral. Los análisis de diversidad alfa y abundancia taxonómica indicaron que la distribución de filos bacterianos difería entre regiones geográficas, con una diversidad significativamente menor en las comunidades del suelo de la Columbia Británica. No encontramos que la remoción de materia orgánica impacte consistentemente los niveles de carbono orgánico del suelo, contenido de nitrógeno, contenido de humedad o pH. De manera similar, los modelos de regresión lineal para cada región indicaron asociaciones mínimas entre los factores abióticos del suelo y la diversidad alfa bacteriana. En general, estos hallazgos

brindan información sobre la asociación entre la composición de la comunidad bacteriana y los factores abióticos del suelo en una variedad de regiones geográficas, y el impacto de la deforestación en estas relaciones.

Las condiciones de los factores abiótico y biótico de las muestras estudiadas dieron como resultado las presencias de valor no significativo en las diversidad alfa y beta en grupo de suelos Acrisoles Háplico se mostró un mayor índice de diversidad y, por tanto, una mayor abundancia y riqueza de comunidades microbiana de especies determinando así la frecuencias cuantitativas y cualitativas de las comunidades microbiana tomando en cuenta los dos tipos de suelo Acrisoles Háplico Cutanico y Hyperdistrico, condición ( intervenido y no intervenido), aluminio, pH, magnesio, localidad, contenido orgánico. Siendo un total de 829,398 secuencias filtrada.

Aunque los resultados obtenidos no han mostrado diferencias significativas en relación a la diversidad bacteriana y los factores de fisicoquímico de pH, condición, tipo de suelo, aluminio, contenido orgánico, magnesio y localización no se descarta una futura publicación científica, ya que esta sería, a fecha de hoy, el primer estudio científico que describe la metagenómica de suelo en Panamá a base de secuenciación de segunda generación y análisis bioinformático de la data, un método novedoso para la diversidad procariota en suelo de Panamá con la que puede llevarse a cabo trabajos interesante sobre biblioteca genómica, descubrimiento de genes novedosos con importancia médica, farmacéutica y de industria.

# **Capitulo IV**

## Conclusiones

- Las bacterias estuvieron presentes en porcentaje alto a comparación de las archeas, debido a las condiciones del suelo y las localidades de Panamá.
- El análisis de Qiime2 para las diversidades alfa presente en la muestra demostró el alto índice de similitud para los métodos métrico de índice de diversidad de Shannon y el método de Krustal Wallis de características observadas, diversidad filogenética de las muestras y uniformidad en las comunidades presente. La diversidad beta mostro un alto índice de disimilitud para el método distancia UniFrac no ponderada Cualitativa en los suelos ACRIOLS HAPLICO hyperdistrico (intervenido y no intervenido) siendo las muestras ACHD30.1, ACHD31.1. y distancia uniFrac ponderada cuantitativo para suelo ACRIOLS HAPLICO cutanico y hyperdistrico (intervenido y no intervenido) para las muestras ACHD28.1, ACHD29.1, ACHD30.1 en todas las categorías (ID de muestras, tipo de muestra, pH, aluminio, magnesio, contenido orgánico, localización, condición.)
- Al no cumplir los supuestos de normalidad, se eligió el método de Kruskal Wallis para realizar las comparaciones por parejas y ver si existían diferencias significativas de alfa diversidad. ninguno de los contrastes mostró diferencias significativas ( $p\text{-value} < 0,05$ ) entre los grupos experimentales analizados en el presente estudio, aunque se observa una tendencia no significativa en las métricas de variabilidad.
- La secuenciación de amplicón 16S ARNr de extremos emparejados en la plataforma ilumina y los cebadores 515F-806R se dirigen a la región V3-V4 del ARNr de 16s SSU es el método utilizado para amplificar procariotas (bacterias y archeas), Las secuenciaciones de amplicon utilizando para la identificación de ASV en Novogen resulto una cantidad de referencias que genera 50,000–100,000 pb.
- DADA2 utilizo un algoritmo de eliminación de ruido que genera ASV como más secuencias exactamente representativas (Callahan et al., 2016).
- Debido a los factores Físico- químicos presente en el suelo de grupo Acrisol y clasificación Haplico mediante el análisis de Qiime2 no se determinó la presencia de Archeae, y las presencias de la comunidad microbiana fueron presentes en ACHD28, ACHD29, ACHD30, ACHD31.

## Recomendación

- No contamos con disponibilidad local de kit de extracción de ADN de suelo.
- Suelos de bajo rendimiento de ADN, ya que no podemos detectar
- Algunas de las muestras de los suelos seleccionados están ubicadas en localidades distantes y de difícil acceso lo cual requiere una logística de transporte y disponibilidad de recursos.
- Los servicios de secuenciación metagenómica de amplicon novogénico proporcionaron una detección eficiente de variantes y organismos objetivo, describen y comparan la diversidad de múltiples entornos complejos.
- Las métricas de diversidad son sensibles a los errores introducidos durante la PCR y secuenciación por la cual recomendamos este proceso de una manera eficaz.



# **Bibliográficas**

Amann RI, L. W. S. K., 1995. Identificación filogenética y detección in situ de células microbianas individuales sin cultivo.. *Microbiol Rev* , 59(5), pp. 143-169.

Anon., s.f. e.

Bruner, G. y otros, 2013. Bacterial Diversity and Physicochemical Stratification of the Water Column during the Rainy Season in Las Cumbres Lake, Panama. *The Internet Journal of Microbiology*, 12(1), pp. 2-3.

Burke, M. C. & Darling, A. E., 2016. Un metodo para la secuenciacion siendo asi de facil analisis de comunidades microbianas no cultivables. *PeerJ*, pp. 35-37.

Caporaso , J. G. & Lauber, C. L., 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA*, 1(108), pp. 4516-4522.

Daniel, R., 2005. The metagenomics of soil. *Nature Reviews Microbiology*, 3(6), p. 470–478.

Daniel, R., 2005. THE METAGENOMICS OF SOIL. *Nature Publishing Group*, 3(4), p. 470.

E. Paul, & E. Clark, 1989. *Soil microbiology and biochemistry*, San Diego: Academic Press.

FAO, y otros, 2020. *State of knowledge soil biodiversity*, Roma: FAO.

Fleischmann R, A. M. W. O. C. R. K. E., 1995. Secuenciación aleatoria de genoma completo y ensamblaje de haemophilus influenzae rd. *Science*, 269(5), p. 496–512.

Hartemink, A. E., 2016. Chapter Two - The definition of soil since the early 1800s. *Academic Press*, pp. 112-118.

Hernández-León, R., Velázquez-Sepúlveda, I., Orozco-Mosqueda, M. C. & Santoyo, G., 2010. Metagenómica de suelos: Grandes desafíos y nuevas oportunidades biotecnológicas. *Phyton*, 8(76), p. 136.

IUSS Working Group WRB, 2014. *World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps*. World Soil Resources Reports No. 106 ed. Rome, Italy: FAO.

John C. Wooley, A. G. I. F., 2010. A Primer on Metagenomics. *journals.plos.org*, 5(2), pp. 10-13.

Joppa, . L. . N., Roberts, D. L., Myers, N. & Pimm, S. L., 2011. Biodiversity hotspots House most undiscovered plant species. *Proceedings of the National Academy of Sciences of the United States of America*, 32(108), pp. 1-5.

Kembel, y otros, 2011. La diversidad filogenetica de los metagenomas. *PloSuno*, 8(6), p. e23214.

- Laserna Mendieta, E. J. & Claesson, M. J., 2016. CREACION DE LIBRERIAS MEDIANTE AMPLIFICACION DEL ARN RIBOSOMAL 16S PARA LA CARACTERIZACION DE LA MICROBIOTA. *Microbiologia*, pp. 70-75.
- Marchesi , J. R. & Ravel, J., 2015. The Vocabulary of microbiome Research: A Proposal. *Microbiome*, pp. 3-31.
- McBratney, A. B., Field, D. J. & Koch, A., 2014. The dimensions of soil security. *Geoderma* 213, pp. 203-213.
- Nogales, B., 2005. La microbiología del suelo en la era de la biología molecular: descubriendo la punta del iceberg. *Ecosistema*, pp. 41-51.
- Pace , N. R., 1997. A molecular view of microbial diversity and the biosphere. *Science*, pp. 734-740.
- Qian, X.-B. y otros, 2020. A guide to human microbiome research: study design, sample collection, and bioinformatics analysis. *Chinese medical journal*, 133(15), pp. 1844-1852.
- R. Amundson, A. B. , J. H. , C. O. , A. S. , D. S., 2015. El suelo y la seguridad humana en el siglo XXI. *Ciencia*, 348(6), pp. 1261071-1–1261071-6.
- Rappé MS, G. S., 2003. La mayoría microbiana no cultivada. *Annu Rev Microbiol* , 57(6), p. 369–394..
- Reyes, I. & Alexis Valery, 2007. Efecto de la Fertilidad del Suelo sobre la Microbiota y la Promoción del Crecimiento Del Maíz (zea mays L.) con Azotobacter spp.. *Bioagro*, pp. 5-9.
- Richter, D. D. & M. D., 1995. How deep is soil?. *Bioscience*, 45(1), p. 600–609.
- Rusch, D. B. y otros, 2007. La expedición sorcerer II Global Ocean Sampling: Atlántico noroeste a través del Pacífico tropical oriental. *PLoS Biol*, 2 Abril, Issue 5, p. e77.
- S. Hofmockel, J. K. J. K., 2018. The soil microbiome — from metagenomics to metaphenomics. *sciencedirect.com*, 43(4), p. 162–168.
- Targulian, V. O. & Krasilnikov, P., 2007. Soil system and pedogenic processes: self-organization,. *Catena* 71, p. 373–381.
- Torsvik, V. F. S. R.-A. & Ø., 2002. Microbial diversity and function in soil: from genes to ecosystems.. *Curr. Opin. Microbiol.*, 5(3), pp. 240-245.
- Venter J, R. K. H. J. H. A. R. D. E. J. W. D. P. I. N. K. N. W. F. D. L. S. K. A. L. M. N. K. W. O. P. J. H. J. P. R. B.-T. H. P. C. R. Y. S. H., 2004. Environmental genome shotgun sequencing

of the Sargasso Sea. *US National Library of Medicine National Institutes of Health*, 304(5667), pp. 66-74.

Woese, C. R. & Fox, G. E., 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci USA*, 11(74), pp. 5088-5090.

Wooley, J. C., Godzik, A. & Friedberg, I., 2010. A Primer on Metagenomics. *PLoS Computational*, 6(2), p. 1.

Young, V. B., 2017. The Role of the Microbiome in Human Health and Disease; An Introduction for Clinicians. *BMJ*, p. 35

Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C (2014) Binning metagenomic contigs by coverage and composition. *Nat Methods* 11:1144–1146

Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM et al (2011) Enterotypes of the human gut microbiome. *Nature* 473:174–180  
Asnicar F, Weingart G, Tickle TL, Huttenhower C, Segata N (2015) Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* 3:e1029

Asshauer KP, Wemheuer B, Daniel R, Meinicke P (2015) Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* 31:2882–2884

Bai Y, Müller DB, Srinivas G, Garrido-Oter R, Potthoff E, Rott M, Dombrowski N, Münch PC, Spaepen S, Remus-Emsermann M et al (2015) Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* 528:364–369

Bastian M, Heymann S, and Jacomy M (2009). Gephi: an open source software for exploring and manipulating networks. In: Third international AAAI conference on weblogs and social media. Beckers B, Op De Beeck M, Weyens N, Boerjan W, Vangronsveld J (2017) Structural variability and niche differentiation in the rhizosphere and endosphere bacterial microbiome of field-grown poplar trees. *Microbiome* 5:25

Bertrand D, Shaw J, Kalathiyappan M, Ng AHQ, Kumar MS, Li C, Dvornicic M, Soldo JP, Koh JY, Tong C et al (2019) Hybrid metagenomic assembly enables high-resolution analysis of resistance determinants and mobile elements in human microbiomes. *Nat Biotechnol* 37:937–944

Bishara A, Moss EL, Kolmogorov M, Parada AE, Weng Z, Sidow A, Dekas AE, Batzoglou S, Bhatt AS (2018) High-quality genome sequences of uncultured microbes by assembly of read clouds. *Nat Biotechnol* 36:1067–1075

Blin K, Weber T, Lee SY, Medema MH, Pascal Andreu V, de los Santos ELC, Del Carratore F (2018) The antiSMASH database version 2: a comprehensive resource on secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res* 47:D625–D630

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F et al (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37:852–857

Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloë-Fadrosh EA et al (2017) Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 35:725–731

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–583

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI et al (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336

Carini P, Marsden PJ, Leff JW, Morgan EE, Strickland MS, Fierer N (2016) Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nat Microbiol* 2:16242

Carrión VJ, Perez-Jaramillo J, Cordovez V, Tracanna V, de Hollander M, Ruiz-Buck D, Mendes LW, van Ijcken WFJ, GomezExposito R, Elsayed SS et al (2019) Pathogen-induced activation of disease-suppressive functions in the endophytic root microbiome. *Science* 366:606–612

Charalampous T, Kay GL, Richardson H, Aydin A, Baldan R, Jeanes C, Rae D, Grundy S, Turner DJ, Wain J et al (2019) Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat Biotechnol* 37:783–792

Chen Q, Jiang T, Liu Y-X, Liu H, Zhao T, Liu Z, Gan X, Hallab A, Wang X, He J et al (2019) Recently duplicated sesterterpene (C25) gene clusters in *Arabidopsis thaliana* modulate root microbiota. *Sci China Life Sci* 62:947–958

Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, Tramontano M, Driessen M, Hercog R, Jung F-E et al (2017) Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol* 35:1069–1076

Csardi G, Nepusz T (2006) The igraph software package for complex network research. *InterJ Complex Syst* 1695:1–9 REVIEW Yong-Xin Liu et al. © The Author(s) 2020 Protein & Cell

de Goffau MC, Lager S, Sovio U, Gaccioli F, Cook E, Peacock SJ, Parkhill J, Charnock-Jones DS, Smith GCS (2019) Human placenta has no microbiome but can contain potential pathogens. *Nature* 572:329–334

De Muinck EJ, Trosvik P, Gilfillan GD, Hov JR, Sundaram AYM (2017) A novel ultra high-throughput 16S rRNA gene amplicon sequencing library preparation method for the Illumina HiSeq platform. *Microbiome* 5:68 Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461

Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10:996–998

Edgar RC, Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 31:3476–3482

Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, Eisen JA, Sundaresan V (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci USA* 112:E911–E920

Edwards JA, Santos-Medellín CM, Liechty ZS, Nguyen B, Lurie E, Eason S, Phillips G, Sundaresan V (2018) Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. *PLoS Biol* 16:e2003862

Fan K, Delgado-Baquerizo M, Guo X, Wang D, Wu Y, Zhu M, Yu W, Yao H, Zhu Y-g, Chu H (2019) Suppressed N fixation and diazotrophs after four decades of fertilization. *Microbiome* 7:143

Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV et al (2008) The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 26:541–547

Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, Lipson KS, Knight R, Caporaso JG, Segata N et al (2018) Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* 15:962–968

Fresia P, Antelo V, Salazar C, Giménez M, D'Alessandro B, Afshinnekoo E, Mason C, Gonnet GH, Iraola G (2019) Urban metagenomics uncover antibiotic resistance reservoirs in coastal beach and sewage waters. *Microbiome* 7:35

Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150–3152

Galkin F, Aliper A, Putin E, Kuznetsov I, Gladyshev VN, Zhavoronkov A (2018) Human microbiome aging clocks based on deep learning and tandem of permutation feature importance and accumulated local effects. *bioRxiv* 507780

Gao L, Xu T, Huang G, Jiang S, Gu Y, Chen F (2018) Oral microbiomes: more and more importance in oral cavity and whole body. *Protein Cell* 9:488–500

Gonzalez A, Navas-Molina JA, Kosciulek T, McDonald D, VázquezBaeza Y, Ackermann G, DeReus J, Janssen S, Swafford AD, Orchanian SB et al (2018) Qiita: rapid, web-enabled microbiome meta-analysis. *Nat Methods* 15:796–798

Goodman AL, Kallstrom G, Faith JJ, Reyes A, Moore A, Dantas G, Gordon JI (2011) Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci USA* 108:6252–6257

Grüning B, Dale R, Sjödin A, Chapman BA, Rowe J, Tomkins-Tinch CH, Valieris R, Köster J, The Bioconda T (2018) Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat Methods* 15:475–476

Guo X, Zhang X, Qin Y, Liu Y-X, Zhang J, Zhang N, Wu K, Qu B, He Z, Wang X et al (2020) Host-associated quantitative abundance profiling reveals the microbial load variation of root microbiome. *Plant Commun* 1:100003

Huang AC, Jiang T, Liu Y-X, Bai Y-C, Reed J, Qu B, Goossens A, Nützmann H-W, Bai Y, Osbourn A (2019) A specialized metabolic network selectively modulates Arabidopsis root microbiota. *Science* 364:eaau6389

Huang P, Zhang Y, Xiao K, Jiang F, Wang H, Tang D, Liu D, Liu B, Liu Y, He X et al (2018) The chicken gut metagenome and the modulatory effects of plant-derived benzyloquinoline alkaloids. *Microbiome* 6:211

Huson DH, Beier S, Flade I, Górská A, El-Hadidi M, Mitra S, Ruscheweyh H-J, Tappu R (2016) MEGAN community edition— interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput Biol* 12:e1004957

Hyatt D, LoCascio PF, Hauser LJ, Uberbacher EC (2012) Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* 28:2223–2230

Ji P, Zhang Y, Wang J, Zhao F (2017) MetaSort untangles metagenome assembly by reducing microbial community complexity. *Nat Commun* 8:14306

Jiang X, Li X, Yang L, Liu C, Wang Q, Chi W, Zhu H (2019) How microbes shape their communities? A microbial community model based on functional genes. *Genom Proteom Bioinf* 17:91–105

Jiao S, Liu Z, Lin Y, Yang J, Chen W, Wei G (2016) Bacterial communities in oil contaminated soils: biogeography and cooccurrence patterns. *Soil Biol Biochem* 98:64–73

Jin T, Wang Y, Huang Y, Xu J, Zhang P, Wang N, Liu X, Chu H, Liu G, Jiang H et al (2017) Taxonomic structure and functional association of foxtail millet root microbiome. *Giga Sci* 6:1–12

Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30

Kanehisa M, Sato Y, Morishima K (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 428:726–731

Kang DD, Froula J, Egan R, Wang Z (2015) MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3:e1165

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2012) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencingbased diversity studies. *Nucleic Acids Res* 41:e1–e1

Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, Gonzalez A, Kosciolek T, McCall L-I, McDonald D et al (2018) Best practices for analysing microbiomes. *Nat Rev Microbiol* 16:410–422

Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R, Kelley ST (2011) Bayesian community-wide culture-independent microbial source tracking. *Nat Methods* 8:761

A practical guide to amplicon and metagenomic analysis of microbiome data REVIEW © The Author(s) 2020 Protein & Cell Kurtz ZD, Müller CL, Miraldi ER, Littman DR, Blaser MJ, Bonneau RA (2015) Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol* 11:e1004226



Lagier J-C, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, Levasseur A, Rolain J-M, Fournier P-E, Raoult D (2018) Culturing the human microbiota and culturomics. *Nat Rev Microbiol* 16:540–550

Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Vega Thurber RL, Knight R et al (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 31:814

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359

Letunic I, Bork P (2019) Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47:W256–W259

Levy A, Salas Gonzalez I, Mittelviehhaus M, Clingenpeel S, Herrera Paredes S, Miao J, Wang K, Devescovi G, Stillman K, Monteiro F et al (2018) Genomic features of bacterial adaptation to plants. *Nat Genet* 50:138–150

Li D, Liu C-M, Luo R, Sadakane K, 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

Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, Arumugam M, Kultima JR, Prifti E, Nielsen T et al (2014) An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* 32:834–841

Liu C, Zhou N, Du M-X, Sun Y-T, Wang K, Wang Y-J, Li D-H, Yu H-Y, Song Y, Bai B-B et al (2020) The mouse gut microbial Biobank expands the coverage of cultured bacteria. *Nat Commun* 11:79

Liu Y-X, Qin Y, Bai Y (2019) Reductionist synthetic community approaches in root microbiome research. *Curr Opin Microbiol* 49:97–102

Liu Y-X, Qin Y, Guo X, Bai Y (2019) Methods and applications for microbiome data analysis. *Hereditas (Beijing)* 41:1–18

Louca S, Parfrey LW, Doebeli M (2016) Decoupling function and taxonomy in the global ocean microbiome. *Science* 353:1272–1277

Mahnert A, Moissl-Eichinger C, Zojer M, Bogumil D, Mizrahi I, Rattei T, Martinez JL, Berg G (2019) Man-made microbial resistances in built environments. *Nat Commun* 10:968

Marchesi JR, Ravel J (2015) The vocabulary of microbiome research: a proposal. *Microbiome* 3:31

McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P (2011) An improved Greengenes taxonomy with explicit ranks for

ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6:610 Members BDC (2019) Database resources of the BIG data center in 2019. *Nucleic Acids Res* 47:D8–D14

Metcalf JL, Xu ZZ, Weiss S, Lax S, Van Treuren W, Hyde ER, Song SJ, Amir A, Larsen P, Sangwan N et al (2016) Microbial community assembly and metabolic function during mammalian corpse decomposition. *Science* 351:158–162

Metsky HC, Siddle KJ, Gladden-Young A, Qu J, Yang DK, Brehio P, Goldfarb A, Piantadosi A, Wohl S, Carter A et al (2019) Capturing sequence diversity in metagenomes with comprehensive and scalable probe design. *Nat Biotechnol* 37:160–168

Mikheenko A, Saveliev V, Gurevich A (2016) MetaQUAST: evaluation of metagenome assemblies. *Bioinformatics* 32:1088–1090

Mitchell AL, Almeida A, Beracochea M, Boland M, Burgin J, Cochrane G, Crusoe MR, Kale V, Potter SC, Richardson LJ et al (2020)

MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Res* 48:D570–D578 Moss EL, Maghini DG, and Bhatt AS (2020) Complete, closed bacterial genomes from microbiomes using nanopore sequencing. *Nat Biotechnol*

Mu D-S, Liang Q-Y, Wang X-M, Lu D-C, Shi M-J, Chen G-J, Du Z-J (2018) Metatranscriptomic and comparative genomic insights into resuscitation mechanisms during enrichment culturing. *Microbiome* 6:230 Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ (2014)

IQTREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274

Ning K, Tong Y (2019) The fast track for microbiome research. *Genom Proteom Bioinf* 17:1–3 Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834

Oksanen J, Kindt R, Legendre P, O’Hara B, Stevens MHH, Oksanen MJ, Suggests M (2007) The vegan package. *Commun Ecol Pack* 10:631–637

Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30:3123–3124

Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A, Ghensi P et al (2019)

Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell* 176:649–662. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017)

Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 14:417–149

Pedersen HK, Forslund SK, Gudmundsdottir V, Petersen AØ, Hildebrand F, Hyötyläinen T, Nielsen T, Hansen T, Bork P, Ehrlich SD et al (2018) A computational framework to integrate highthroughput ‘-omics’ datasets for the identification of potential mechanistic links. *Nat Protoc* 13:2781–2800

Proctor LM, Creasy HH, Fettweis JM, Lloyd-Price J, Mahurkar A, Zhou W, Buck GA, Snyder MP, Strauss JF, Weinstock GM et al (2019) The integrative human microbiome project. *Nature* 569:641–648

Qian X, Liu Y-X, Ye X, Zheng W, Lv S, Mo M, Lin J, Wang W, Wang W, Zhang X et al (2020) Gut microbiota in children with juvenile idiopathic arthritis: characteristics, biomarker identification, and usefulness in clinical prediction. *BMC Genom* 21:286  
Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T et al (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464:59–65

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO (2013) The SILVA ribosomal RNA gene  
REVIEW Yong-Xin Liu et al. © The Author(s) 2020 Protein & Cell database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–596

Quince C, Walker AW, Simpson JT, Loman NJ, Segata N (2017) Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol* 35:833  
Ren Z, Li A, Jiang J, Zhou L, Yu Z, Lu H, Xie H, Chen X, Shao L, Zhang R et al (2019) Gut microbiome analysis as a tool towards targeted non-invasive biomarkers for early hepatocellular carcinoma. *Gut* 68:1014–1023

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584  
Ross AA, Müller KM, Weese JS, Neufeld JD (2018)

Comprehensive skin microbiome analysis reveals the uniqueness of human skin and evidence for phylosymbiosis within the class mammalia. *Proc Natl Acad Sci USA* 115:E5786–E5795

Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, Costea PI, Godneva A, Kalka IN, Bar N et al (2018) Environment dominates over host genetics in shaping human gut microbiota. *Nature* 555:210

Roux S, Adriaenssens EM, Dutilh BE, Koonin EV, Kropinski AM, Krupovic M, Kuhn JH, Lavigne R, Brister JR, Varsani A et al (2019) Minimum information about an uncultivated virus genome (MIUViG). *Nat Biotechnol* 37:29–37

Saito R, Smoot ME, Ono K, Ruscheinski J, Wang P-L, Lotia S, Pico AR, Bader GD, Ideker T (2012) A travel guide to cytoscape plugins. *Nat Methods* 9:1069–1076

Salazar G, Paoli L, Alberti A, Huerta-Cepas J, Ruscheweyh H-J, Cuenca M, Field CM, Coelho LP, Cruaud C, Engelen S et al (2019) Gene expression changes and community turnover differentially shape the global ocean metatranscriptome. *Cell* 179:1068–1083.e1021

Shi W, Qi H, Sun Q, Fan G, Liu S, Wang J, Zhu B, Liu H, Zhao F, Wang X et al (2019) gcMeta: a global catalogue of metagenomics platform to support the archiving, standardization and analysis of microbiome data. *Nucleic Acids Res* 47:D637–D648

Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF (2018) Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* 3:836–843

Sinha R, Abu-Ali G, Vogtmann E, Fodor AA, Ren B, Amir A, Schwager E, Crabtree J, Ma S, Abnet CC et al (2017) Assessment of variation in microbial community amplicon sequencing by the microbiome quality control (MBQC) project consortium. *Nat Biotechnol* 35:1077–1086

Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G, Knight R, Manjurano A, Chagalucha J, Elias JE et al (2017) Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* 357:802–806

Stewart RD, Auffret MD, Warr A, Walker AW, Roehe R, Watson M (2019) Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery. *Nat Biotechnol* 37:953–961

Tange O (2018). Gnu parallel 2018 (Lulu. com). Tierney BT, Yang Z, Lubber JM, Beaudin M, Wibowo MC, Baek C, Mehlenbacher E, Patel CJ, Kostic AD (2019) The landscape of genetic content in the gut and oral human microbiome. *Cell Host Microbe* 26:283–295.e288

Tkacz A, Hortal M, Poole PS (2018) Absolute quantitation of microbiota abundance in environmental samples. *Microbiome* 6:110

Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, Tett A, Huttenhower C, Segata N (2015) MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nat Methods* 12:902–903

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI (2007) The human microbiome project. *Nature* 449:804–810

Turner TR, Ramakrishnan K, Walshaw J, Heavens D, Alston M, Swarbreck D, Osbourn A, Grant A, Poole PS (2013) Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *ISME J* 7:2248–2258

Uritskiy GV, DiRuggiero J, Taylor J (2018) MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6:158

Vandeputte D, Kathagen G, D’hoë K, Vieira-Silva S, Valles-Colomer M, Sabino J, Wang J, Tito RY, De Commer L, Darzi Y, et al (2017) Quantitative microbiome profiling links gut community variation to microbial load. *Nature* 551:507–511

Vangay P, Hillmann BM, Knights D (2019) Microbiome Learning Repo (ML Repo): A public repository of microbiome regression and classification tasks. *GigaScience* 8:giz042

Wang J, Chen L, Zhao N, Xu X, Xu Y, Zhu B (2018) Of genes and microbes: solving the intricacies in host genomes. *Protein Cell* 9:446–461

Wang J, Jia Z, Zhang B, Peng L, and Zhao F (2019) Tracing the accumulation of in vivo human oral microbiota elucidates microbial community dynamics at the gateway to the GI tract. *Gut*, gutjnl-2019–318977

Wang J, Thingholm LB, Skiecevičienė J, Rausch P, Kummén M, Hov JR, Degenhardt F, Heinsen F-A, Rühlemann MC, Szymczak S A practical guide to amplicon and metagenomic analysis of microbiome data REVIEW © The Author(s) 2020 *Protein & Cell* et al (2016) Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. *Nat Genet* 48:1396–1406

Wang J, Zheng J, Shi W, Du N, Xu X, Zhang Y, Ji P, Zhang F, Jia Z, Wang Y et al (2018) Dysbiosis of maternal and neonatal microbiota associated with gestational diabetes mellitus. *Gut* 67:1614–1625

Wang W, Yang J, Zhang J, Liu Y-X, Tian C, Qu B, Gao C, Xin P, Cheng S, Zhang W et al (2020) An Arabidopsis secondary metabolite directly targets expression of the bacterial type III secretion system to inhibit bacterial virulence. *Cell Host Microbe* 27:601–613.e607

Wang X, Wang M, Xie X, Guo S, Zhou Y, Zhang X, Yu N, and Wang E (2020b) An amplification-selection model for quantified rhizosphere microbiota assembly. *Sci Bull*

Wang Y, Song F, Zhu J, Zhang S, Yang Y, Chen T, Tang B, Dong L, Ding N, Zhang Q et al (2017) GSA: genome sequence archive\*. *Genom Proteom Bioinf* 15:14–18

Ward T, Larson J, Meulemans J, Hillmann B, Lynch J, Sidiropoulos D, Spear JR, Caporaso G, Blekhman R, Knight R et al (2017) BugBase predicts organism-level microbiome phenotypes. *bioRxiv* 133462

Wilck N, Matus MG, Kearney SM, Olesen SW, Forslund K, Bartolomaeus H, Haase S, Mähler A, Balogh A, Markó L et al (2017) Salt-responsive gut commensal modulates TH17 axis and disease. *Nature* 551:585–589

Wood DE, Lu J, and Langmead B (2019) Improved metagenomic analysis with Kraken 2. *bioRxiv* 762302

Wu Y-W, Simmons BA, Singer SW (2015) MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32:605–607

Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, Li X, Long H, Zhang J, Zhang D et al (2015) A catalog of the mouse gut metagenome. *Nat Biotechnol* 33:1103

Xu J, Zhang Y, Zhang P, Trivedi P, Riera N, Wang Y, Liu X, Fan G, Tang J, Coletta-Filho HD et al (2018) The structure and function of the global citrus rhizosphere microbiome. *Nat Commun* 9:4894

Xu Y, Zhao F (2018) Single-cell metagenomics: challenges and applications. *Protein Cell* 9:501–510

Yang J, Yu J (2018) The association of diet, gut microbiota and colorectal cancer: what we eat may imply what we get. *Protein Cell* 9:474–487

Ye SH, Siddle KJ, Park DJ, Sabeti PC (2019) Benchmarking metagenomics tools for taxonomic classification. *Cell* 178:779–794

Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G et al (2011) Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat Biotechnol* 29:415–420

Zgadzaj R, Garrido-Oter R, Jensen DB, Koprivova A, Schulze-Lefert P, Radutoiu S (2016) Root nodule symbiosis in *Lotus japonicus* drives the establishment of distinctive rhizosphere, root, and nodule bacterial communities. *Proc Natl Acad Sci USA* 113: E7996–E8005

Zhang F, Cui B, He X, Nie Y, Wu K, Fan D, Feng B, Chen D, Ren J, Deng M et al (2018) Microbiota transplantation: concept, methodology and strategy for its modernization. *Protein Cell* 9:462–473

Zhang J, Liu Y-X, Zhang N, Hu B, Jin T, Xu H, Qin Y, Yan P, Zhang X, Guo X et al (2019) NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. *Nat Biotechnol* 37:676–684

Zhang J, Zhang N, Liu Y-X, Zhang X, Hu B, Qin Y, Xu H, Wang H, Guo X, Qian J et al (2018) Root microbiota shift in rice correlates with resident time in the field and developmental stage. *Sci China Life Sci* 61:613–621

Zheng M, Zhou N, Liu S, Dang C, Liu Y-X, He S, Zhao Y, Liu W, Wang X (2019) N<sub>2</sub>O and NO emission from a biological aerated filter treating coking wastewater: main source and microbial community. *J Clean Prod* 213:365–374

Zhu W, Lomsadze A, Borodovsky M (2010) Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res* 38:e132– e132

Zou Y, Xue W, Luo G, Deng Z, Qin P, Guo R, Sun H, Xia Y, Liang S, Dai Y et al (2019) 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat Biotechnol* 37:179–185

Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA. Diversidad de la flora microbiana intestinal humana. *Ciencia*. 2005;308:1635–1638.

Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE. Sucesión de consorcios microbianos en el microbioma intestinal infantil en desarrollo. *Proc Natl Acad Sci USA*. 2011;108(Suppl 1):4578–4585.

Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. Desarrollo de la microbiota intestinal infantil humana. *PLoS Biol*. 2007;5:e177.

Dethlefsen L, Huse S, Sogin ML, Relman DA. Los efectos generalizados de un antibiótico en la microbiota intestinal humana, como lo revela la secuenciación profunda del ARNr 16S. *PLoS Biol.* 2008;6:e280.

Dethlefsen L, Relman DA. Recuperación incompleta y respuestas individualizadas de la microbiota intestinal distal humana a la perturbación antibiótica repetida. *Proc Natl Acad Sci USA.* 2011;108(Supl 1):4554–4561.

Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Variación de la comunidad bacteriana en los hábitats del cuerpo humano a través del espacio y el tiempo. *Ciencia.* 2009;326:1694–1697.

Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. El efecto de la dieta en el microbioma intestinal humano: un análisis metagenómico en ratones gnotobióticos humanizados. *Sci Transl Med.* 2009;1:6ra14.

Turnbaugh PJ, Hamady M, Yatsunencko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI. Un microbioma intestinal central en gemelos obesos y delgados. *Naturaleza.* 2009;457:480–484.

Claesson MJ, Cusack S, O'Sullivan O, Greene-Diniz R, de Weerd H, Flannery E, Marchesi JR, Falush D, Dinan T, Fitzgerald G, Stanton C, van Sinderen D, O'Connor M, Harnedy N, O'Connor K, Henry C, O'Mahony D, Fitzgerald AP, Shanahan F, Twomey C, Hill C, Ross RP, O'Toole PW. Composición, variabilidad y estabilidad temporal de la microbiota intestinal de los ancianos. *Proc Natl Acad Sci USA.* 2010;108(Suppl 1):4586–4591.

Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. Microbioma vaginal de mujeres en edad reproductiva. *Proc Natl Acad Sci USA.* 2011;108(Suppl 1):4680–4687.



Grupo de Trabajo HMP de los NIH. Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, Deal C, Baker CC, Di Francesco V, Howcroft TK, Karp RW, Lunsford RD, Wellington CR, Belachew T, Wright M, Giblin C, David H, Mills M, Salomon R, Mullins C, Akolkar B, Begg L, Davis C, Grandison L, Humble M, Khalsa J. et al. El Proyecto de Microbioma Humano de los NIH. *Genoma Res.* 2009;19:2317–2323.

Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P. et al. Un catálogo de genes microbianos intestinales humanos establecido por secuenciación metagenómica. *Naturaleza.* 2010;464:59–65.

Fierer N, Hamady M, Lauber CL, Knight R. La influencia del sexo, las manos y el lavado en la diversidad de bacterias de la superficie de la mano. *Proc Natl Acad Sci USA.* 2008;105:17994–17999.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. Patrones globales de diversidad de ARNr 16S a una profundidad de millones de secuencias por muestra. *Proc Natl Acad Sci USA.* 2011;108(Suppl 1):4516–4522.

Tannock GW, Munro K, Harmsen HJ, Welling GW, Smart J, Gopal PK. Análisis de la microflora fecal de sujetos humanos que consumen un producto probiótico que contiene *Lactobacillus rhamnosus* DR20. *Appl Environ Microbiol.* 2000;66:2578–2588.

Scanlan PD, Shanahan F, O'Mahony C, Marchesi JR. Análisis independientes del cultivo de la variación temporal de la microbiota fecal dominante y subgrupos bacterianos dirigidos en la enfermedad de Crohn. *J Clin Microbiol.* 2006;44:3980–3988.

Martinez I, Kim J, Duffy PR, Schlegel VL, Walter J. Los almidones resistentes tipos 2 y 4 tienen efectos diferenciales sobre la composición de la microbiota fecal en sujetos humanos. *PLoS Uno*. 2010;5:e15046.

Hamady M, Lozupone C, Knight R. Fast UniFrac: facilitando análisis filogenéticos de alto rendimiento de comunidades microbianas, incluido el análisis de pirosecuenciación y datos de PhyloChip. *ISME J*. 2009;4:17–27.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. Greengenes, una base de datos de genes 16S rRNA verificada por quimeras y banco de trabajo compatible con ARB. *Appl Environ Microbiol*. 2006;72:5069–5072. doi: 10.1128/AEM.03006-05. -

Edgar RC. Búsqueda y agrupación en grupos de órdenes de magnitud más rápido que BLAST. *Bioinformática*. 2010;26:2460–2461.

Knights D, Kuczynski J, Koren O, Ley RE, Field D, Knight R, Desantis TZ, Kelley ST. La clasificación supervisada de la microbiota mitiga los errores de etiquetado erróneo. *ISME J*. 2010;5:570–573.

# **Anexo**

## Anexo 1.

### Resumen de los procedimientos analíticos para la caracterización del suelo

En el presente anexo se ofrece un resumen de los procedimientos de análisis recomendados para ser utilizados para la caracterización de los suelos en la Base Referencial Mundial del Recurso Suelo. Las descripciones completas se pueden encontrar en Procedimientos para el análisis del suelo (Van Reeuwijk, 2002) y Manual de métodos de laboratorio para análisis de suelos del USDA (Burt, 2004).

#### 1. PREPARACIÓN DE LA MUESTRA

Las muestras se secan al aire o, alternativamente, en estufa a un máximo de 40°C. La fracción tierra fina se obtiene por tamizado de la muestra seca con un tamiz de 2mm. Los terrones que no pasan a través del tamiz, son aplastados (no molidos) y se tamiza de nuevo. Grava, fragmentos de roca, etc. que no pasan por el tamiz se tratan por separado. En casos especiales cuando el secado al aire causa cambios inaceptables e irreversibles en ciertas propiedades del suelo (por ejemplo, en turba y suelos con propiedades ándicas) las muestras se mantienen y se tratan en el estado de humedad de campo.

#### 2. CONTENIDO DE HUMEDAD

El cálculo de los resultados de análisis de suelos se realiza sobre base de masa de suelo secado en estufa (105 °C).

#### 3. ANALISIS DE TAMAÑO DE PARTICULAS

La parte mineral del suelo se separa en diversos tamaños de fracciones y se determina la proporción de estas fracciones. La determinación comprende todo el material, es decir, incluyendo grava y material más grueso, pero el procedimiento en si mismo se aplica a la fracción tierra fina (<2mm) solamente. El pretratamiento de la muestra está dirigido a la dispersión completa de partículas primarias. Por tanto, los materiales cementantes (por lo general de origen secundario) tales como materia orgánica y carbonato de calcio pueden necesitar ser eliminados. En algunos casos, también es necesario aplicar desferrificación. Sin embargo, dependiendo del objetivo del estudio, puede ser fundamentalmente incorrecto eliminar los materiales de cementación. En consecuencia, todos los pretratamientos deben considerarse opcionales. Sin embargo, para fines de caracterización de suelos la eliminación de la materia orgánica por H<sub>2</sub>O<sub>2</sub> y de carbonatos por HCl se lleva a cabo de forma rutinaria. Después de este

pretratamiento, la muestra se agita con un agente dispersante y la arena se separa de la arcilla y limo con un tamiz de 63 $\mu$ m. La arena es

-Resumen de los procedimientos analíticos para la caracterización del suelo fraccionada mediante tamizado en seco; las fracciones arcilla y limo son determinadas por el método de la pipeta o, alternativamente, por el método del hidrómetro.

#### 4. ARCILLA DISPERSABLE EN AGUA

Este es el contenido de arcilla que se encuentra cuando la muestra se dispersa con agua sin ningún pretratamiento para eliminar compuestos cementantes y sin el uso de un agente dispersante. La proporción de arcilla natural a arcilla total se puede utilizar como un indicador de estabilidad de la estructura.

#### 5. RETENCIÓN DE AGUA DEL SUELO

El contenido de agua se determina en muestras de suelo que han sido equilibradas con agua a varios valores de succión (tensión). Para valores bajos de succión, las muestras no perturbadas se equilibran en un baño de limo y caolín; para valores altos de succión, muestras perturbadas son equilibradas en placas extractoras de presión. La densidad aparente se calcula a partir de la masa de las muestras no perturbadas.

#### 6. DENSIDAD APARENTE

La densidad aparente del suelo es la masa por la unidad de volumen de suelo. Como la densidad aparente cambia con el contenido de agua, debe especificarse el estado de humedad de la muestra. Pueden utilizarse dos procedimientos diferentes:

» **Muestras no perturbadas.** Un cilindro de metal de volumen conocido se presiona en el suelo. Se registra la masa de la muestra húmeda. Esto puede ser en el estado de humedad de campo o en el estado después de equilibrar la muestra a una tensión de agua específica. La muestra es secada en estufa y pesada de nuevo. La densidad aparente es el cociente entre la masa seca y el volumen en el contenido de agua determinado y/o la tensión de agua especificada.

» **Terrones revestidos.** Terrones recogidos en el campo son recubiertos con laca plástica (por ejemplo, Saran disueltos en metil etil cetona) para permitir la determinación de la masa sumergida en agua. Esto da el volumen al terrón. Se registra la masa de muestra húmeda. Esto puede ser en el estado de humedad de campo o en el estado después de equilibrar el terrón con un valor específico de succión de agua. La muestra es secada en estufa y pesada de nuevo. La

densidad aparente es el cociente entre la masa seca y el volumen a la succión de agua especificada.

Nota: la determinación de la densidad aparente es muy sensible a errores, particularmente causados por elementos no representativos de las muestras (piedras, grietas, raíces, etc). Por lo tanto, las determinaciones deben hacerse siempre por triplicado.

#### 7. COEFICIENTE DE EXTENSIBILIDAD LINEAL (COLE)

El COLE da una indicación de la capacidad reversible de expansión-retracción de un suelo. Se calcula a partir de la densidad aparente en seco y la densidad aparente a 33 kPa de succión de agua. El valor COLE se expresa en centímetros por centímetro o como un valor porcentual.196

Base referencial mundial del recurso suelo 2014

#### 8. PH

El pH del suelo se mide potenciométricamente en la suspensión sobrenadante de una mezcla suelo:líquido. Si no se indica lo contrario, suelo:líquido están en una relación 1:5 (volumen: volumen) (de acuerdo con las normas ISO). El líquido es agua destilada (pHagua) o una solución 1M de KCl (pHKCl). Sin embargo, en algunas definiciones, se utiliza una relación de suelo: agua 1:1.

#### 9. CARBONO ORGÁNICO

Se sigue el procedimiento Walkley-Black. Esto implica una combustión en húmedo de materia orgánica con una mezcla de dicromato de potasio y ácido sulfúrico a aproximadamente 125°C. El dicromato residual se valora con sulfato ferroso. Para compensar por la destrucción incompleta, se aplica un factor de corrección empírico de 1.3 en el cálculo del resultado.

Nota: Otros procedimientos, incluyendo analizadores de carbono (por ejemplo, combustión seca) pueden también ser utilizados. En estos casos, se recomienda una prueba cualitativa de carbonatos por efervescencia con HCl y, si están presentes, se requiere una corrección para C inorgánico (ver carbonatos más abajo).

#### 10. CARBONATOS

Se utiliza el método de valoración rápida de Piper (también llamado método de neutralización ácida). La muestra se trata con HCl diluido y el ácido residual se valora. Los resultados están referidos como carbonato de calcio equivalente ya que la disolución no es selectiva para calcita y otros carbonatos tales como dolomita también se disuelven hasta cierto punto.

Nota: También pueden utilizarse otros procedimientos tales como el método volumétrico Scheibler o el calcímetro Bernard.

#### 11. YESO

El yeso se disuelve por agitación de la muestra con agua. A continuación, se precipita selectivamente del extracto mediante la adición de acetona. Este precipitado se redisuelve en agua y se determina la concentración de Ca como una medida para el yeso.

#### 12. CAPACIDAD DE INTERCAMBIO CATIONICO (CIC) Y BASES INTERCAMBIABLES

Se utiliza el método de acetato de amonio pH 7. La muestra se percola con acetato de amonio (pH 7) y las bases se miden en el percolado. La muestra es posteriormente percolada con acetato de sodio (pH 7), se retira el exceso de sal y el sodio adsorbido se intercambia por percolación con acetato de amonio (pH 7). El Na en este percolado es una medida de la CIC. Alternativamente, después de la percolación con acetato de amonio, la muestra puede ser lavada para quitarle el exceso de sal, se destila la muestra completa y se determina el amonio liberado. La percolación en tubos también puede ser reemplazada por agitación en frascos. Cada extracción debe repetirse tres veces y se deben combinar los tres extractos para el análisis

Nota 1: Pueden utilizarse otros procedimientos para la CIC siempre que se realice la determinación a pH 7.

Nota 2: En casos especiales donde la CIC no es un criterio de diagnóstico, por ejemplo, en suelos salinos y alcalinos, puede determinarse la CIC a pH 8.2.

Nota 3: La saturación de bases de suelos salinos, calcáricos y gipsíricos puede considerarse al 100%.

Nota 4: Cuando hay involucradas arcillas de baja- actividad, la CIC de la materia orgánica debe deducirse. Esto puede hacerse por el método gráfico (FAO, 1966), o por el análisis de la CIC de la materia orgánica o de los coloides minerales por separado.

#### 13. ACIDEZ DE INTERCAMBIO Y ALUMINIO INTERCAMBIABLE

Acidez de intercambio (H+Al) y aluminio intercambiable son liberados por intercambio con una solución no tamponada de KCl 1M. La acidez de intercambio también puede designarse como acidez real (en comparación con acidez potencial o extraíble).

#### 14. SÍLICE, MANGANESO, ALUMINIO Y HIERRO EXTRACTABLES

Estos análisis comprenden:

» Fedith, Aldith, Mndith: Compuestos de Fe, Al y Mn libres en el suelo extraídos por una solución de ditionito-citrato-bicarbonato (pueden usarse ambos procedimientos, el Mehra y Jackson y el Holmgren).

» Feox, Alox y Siox: Compuestos activos de bajo grado de ordenamiento o amorfosde Fe, Al y Si extraídos por una solución de oxalato ácido de amonio (pH 3). (Blakemore et al., 1987).

» Fepy, Alpy: Fe y Al orgánicamente ligados extraídos por una solución de pirofosfato.

## 15. SALINIDAD

Las características asociadas a la salinidad de los suelos se determinan en el extracto de saturación. Las características incluyen: pH, conductividad eléctrica (ECe), relación de adsorción de sodio (SAR), cationes y aniones de las sales disueltas. Estos incluyen Ca, Mg, Na, K, carbonato y bicarbonato, cloruro, nitrato y sulfato. El SAR y el porcentaje de sodio intercambiable (ESP) pueden estimarse a partir de las concentraciones de cationes disueltos.

## 16. FOSFATO Y RETENCIÓN DE FOSFATO

Estos análisis comprenden:

» Método Olsen: Extracción con una solución de  $\text{NaHCO}_3$  0,5 M a pH 8.5 (Olsen et al 1954).

» Método de ácido cítrico: extracción con una solución al 1% de ácido cítrico (Blanck, 1931; Van Reeuwijk, 2002).

» Método Mehlich-1: extracción con una solución de  $\text{HCl}$  0.05 M y  $\text{H}_2\text{SO}_4$  0.025 M (Mehlich, 1953).

» Para la retención de fosfatos, se utiliza el procedimiento Blakemore. La muestra se equilibra con una solución de fosfato a pH 4.6 y se determina la proporción de fosfatos extraída de la solución (Blakemore et al., 1987).

## 17. DENSIDAD ÓPTICA DEL EXTRACTO DE OXALATO (ODOE)

La muestra se filtra o se agita con una solución de oxalato ácido de amonio (pH 3). Se mide la densidad óptica del extracto a una longitud de onda de 430 nm.

## 18. INDICE MELÁNICO

Se agita la muestra con una solución de  $\text{NaOH}$  0.5 M y se mide la absorbencia del extracto a 450 y 520 nm, respectivamente. El índice melánico se obtiene dividiendo la absorbencia a 450 nm por la absorbencia a 520 nm.

## 19. ANÁLISIS MINERALÓGICO DE LA FRACCIÓN ARENA



Después de eliminar los materiales de revestimiento y cementantes, la arena se separa de la arcilla y limo mediante tamizado en húmedo. Se separa la fracción 63-420  $\mu\text{m}$  de la arena por tamizado en seco. Esta fracción se divide en fracción pesada y fracción ligera con la ayuda de un líquido de alta densidad: una solución de politungstato de sodio con una densidad específica de 2,85  $\text{kg dm}^{-3}$ . De la fracción pesada, se hace una lámina para el microscopio; la fracción ligera se tiñe selectivamente para a identificación microscópica de cuarzo y feldespatos. El vidrio volcánico puede ser reconocido generalmente como granos isotrópicos con vesículas.

## 20. DIFRACTOMETRÍA DE RAYOS X

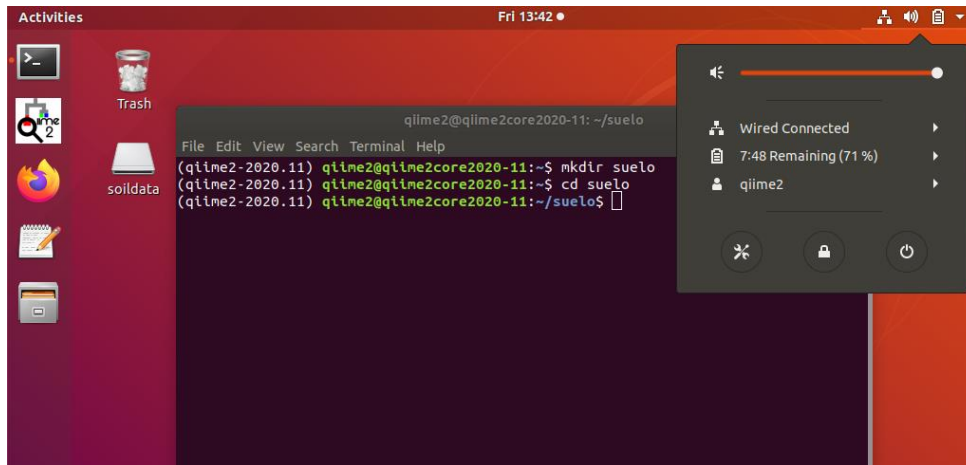
La fracción arcilla se separa de la tierra fina y se deposita de manera orientada en un portaobjetos de vidrio o placas de cerámica porosa para ser analizadas en un difractómetro de rayos x. Muestras de polvo de arcilla y otras fracciones sin orientar se analizan en el mismo aparato o con una cámara Guinier de rayos x (fotografías). El Bromoformo también puede ser usado como líquido de alta densidad, pero su uso no se recomienda debido a su vapor altamente tóxico.

## 21. SULFUROS

El S inorgánico reducido se convierte en  $\text{H}_2\text{S}$  en una solución ácida de  $\text{CrCl}_2$  caliente. El  $\text{H}_2\text{S}$  producido es atrapado cuantitativamente en una solución de acetato de Zn como  $\text{ZnS}$  sólido. El  $\text{ZnS}$  se trata entonces con  $\text{HCl}$  para liberar  $\text{H}_2\text{S}$  en la solución, el cual es rápidamente valorado con una solución de  $\text{I}_2$  hasta el punto final de color azul indicado por la reacción del  $\text{I}_2$  con almidón (Sullivan et al., 2000). Precaución: los residuos tóxicos se deben manejar cuidadosamente. (IUSS Working Group WRB, 2014)

## Anexo 2.

### Script de análisis bioinformático con QIIME2



- Import the multiplexed sequences

```
qiime tools import \  
  --type MultiplexedSingleEndBarcodeInSequence \  
  --input-path forward.fastq.gz \  
  --output-path multiplexed-seqs.qza
```

- Demultiplex the reads

```
qiime cutadapt demux-single \  
  --i-seqs multiplexed-seqs.qza \  
  --m-barcodes-file metadata.txt \  
  --m-barcodes-column Barcode \  
  --p-error-rate 0 \  
  --o-per-sample-sequences demultiplexed-seqs.qza \  
  --o-untrimmed-sequences untrimmed.qza \  
  --verbose
```

- Trim adapters from demultiplexed reads

```
qiime cutadapt trim-single \  
  --i-demultiplexed-sequences demultiplexed-seqs.qza \  
  --p-front GCTACGGGGGG \  
  --p-error-rate 0
```

```
--o-trimmed-sequences trimmed-seqs.qza \  
--verbose
```

- Summarize demultiplexed and trimmed reads

```
qiime demux summarize \  
--i-data trimmed-seqs.qza \  
--o-visualization trimmed-seqs.qzv  
qiime tools view trimmed-seqs.qzv
```

- Sequence quality control and feature table construction

```
qiime dada2 denoise-single \  
--i-demultiplexed-seqs demultiplexed-seqs.qza \  
--p-trim-left 23 \  
--p-trunc-len 154 \  
--o-representative-sequences rep-seqs-dada2.qza \  
--o-table table-dada2.qza \  
--o-denoising-stats stats-dada2.qza
```

```
mv rep-seqs-deblur.qza rep-seqs.qza  
mv table-deblur.qza table.qza
```

```
qiime feature-table summarize \  
--i-table table.qza \  
--o-visualization table.qzv \  
--m-sample-metadata-file sample-metadata.tsv  
qiime feature-table tabulate-seqs \  
--i-data rep-seqs.qza \  
--o-visualization rep-seqs.qzv
```

- Generate a tree for phylogenetic diversity analyses

```
qiime phylogeny align-to-tree-mafft-fasttree \  
--i-sequences rep-seqs.qza \  
--o-alignment aligned-rep-seqs.qza \  
--o-masked-alignment masked-aligned-rep-seqs.qza \  
--o-tree tree-rep-seqs.qza
```

```
--o-tree unrooted-tree.qza \  
--o-rooted-tree rooted-tree.qza
```

- Alpha and beta diversity analysis

```
qiime diversity core-metrics-phylogenetic \  
--i-phylogeny rooted-tree.qza \  
--i-table table.qza \  
--p-sampling-depth 151122 \  
--m-metadata-file sample-metadata.tsv \  
--output-dir core-metrics-results
```

```
qiime diversity alpha-group-significance \  
--i-alpha-diversity core-metrics-results/faith_pd_vector.qza \  
--m-metadata-file sample-metadata.tsv \  
--o-visualization core-metrics-results/faith-pd-group-significance.qzv
```

```
qiime diversity alpha-group-significance \  
--i-alpha-diversity core-metrics-results/evenness_vector.qza \  
--m-metadata-file sample-metadata.tsv \  
--o-visualization core-metrics-results/evenness-group-significance.qzv
```

Next We'll Analyze Sample Composition In The Context Of Categorical Metadata Using Permanova (First Described In [Anderson \(2001\)](#)) Using The Beta-Group-Significance Command.

```
qiime diversity beta-group-significance \  
--i-distance-matrix core-metrics-results/unweighted_unifrac_distance_matrix.qza \  
--m-metadata-file sample-metadata.tsv \  
--m-metadata-column sample-type \  
--o-visualization core-metrics-results/unweighted-unifrac-sample-type-significance.qzv --p-pairwise
```

```
qiime diversity beta-group-significance \  
--i-distance-matrix core-metrics-results/unweighted_unifrac_distance_matrix.qza \  
--o-visualization core-metrics-results/unweighted-unifrac-sample-type-significance.qzv --p-pairwise
```

```
--m-metadata-file sample-metadata.tsv \
```

```
--m-metadata-column Condition \
```

```
--o-visualization core-metrics-results/unweighted-unifrac-Condition-significance.qzv \
```

```
--p-pairwise
```

- Alpha rarefaction plotting

```
qiime diversity alpha-rarefaction \
```

```
--i-table table.qza \
```

```
--i-phylogeny rooted-tree.qza \
```

```
--p-max-depth 159317\
```

```
--m-metadata-file sample-metadata.tsv \
```

```
--o-visualization alpha-rarefaction.qzv
```

- Taxonomic analysis

```
wget \
```

```
-O "gg-13-8-99-515-806-nb-classifier.qza" \
```

```
qiime feature-classifier classify-sklearn \
```

```
--i-classifier gg-13-8-99-515-806-nb-classifier.qza \
```

```
--i-reads rep-seqs.qza \
```

```
--o-classification taxonomy.qza
```

```
qiime metadata tabulate \
```

```
--m-input-file taxonomy.qza \
```

```
--o-visualization taxonomy.qzv
```

```
qiime taxa barplot \
```

```
--i-table table.qza \
```

```
--i-taxonomy taxonomy.qza \
```

```
--m-metadata-file sample-metadata.tsv \
```

```
--o-visualization taxa-bar-plots.qz
```

Anexo 4

Clasificación de Suelo de Panamá 2010

<b>Abreviatura de suelo</b>	<b>Descripción</b>
ACct	Haplic Acrisols (Cutanic)
ACdy	Haplic Acrisols (Hyperdystric)
ACpl	Acrisol Plinthic
ALumct	Alisoles úmbricos arenicos
Anum	Andosoles úmbricos dístricos
CMch	Cambisoles Chernic ch
CMdy	Cambisoles Dystric
CMdy/eu	Cambisoles districos/eutricos gleyicos,
CMeu	Cambisoles eutricos
CMfl	Cambisoles Ferralic
CMha	Cambisoles Haplic
FLeu	Fluvisol Eutric
GLmo	Gleysol Mollic
LPdyeu	Leptosol Dystric/ Eutric
LXcr	Lixisol Chromic
NTha	Nitisol Haplic
NThudy	Nitisol Humic
NTro	Nitisol Rhodic
NTum	Nitisol Umbric
PHha	Phaeozem Haplic
RGmo	Regosol Mollic
VEum	Vertisol Umbric
VRha	Vertisol Haplic