

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ciencias y Humanidades



Comparación de perfiles metagenómicos de
fermentación de cacao de datos publicados en
servidores públicos, con información de distintos
orígenes de países productores de cacao

Trabajo de graduación en modalidad de Tesis
presentado por Ximena Alejandra Porras
Moscoso
para optar al grado académico de Licenciado en
Bioquímica y Microbiología

Guatemala,
2024

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ciencias y Humanidades

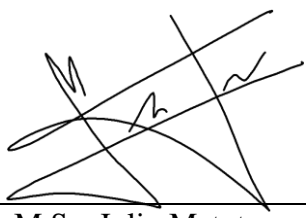


Comparación de perfiles metagenómicos de
fermentación de cacao de datos publicados en
servidores públicos, con información de distintos
orígenes de países productores de cacao

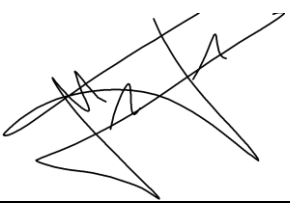
Trabajo de graduación en modalidad de Tesis
presentado por Ximena Alejandra Porras
Moscoso
para optar al grado académico de Licenciado en
Bioquímica y Microbiología

Guatemala,
2024

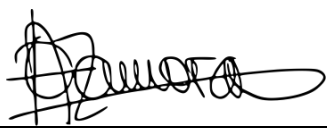
Vo.Bo.:

(f) 
M.Sc. Julio Matute

Tribunal examinador:

(f) 
M.Sc. Julio Matute

(f) 
Ph.D. Krisztina Ríos-Fulop

(f) 
Ph.D. Alejandra Zamora

Fecha de aprobación: Guatemala, 25 de enero 2024

Prefacio

A la Universidad del Valle de Guatemala, por permitirme formar parte de esta casa de estudios.

A mi asesor principal, M.Sc. Julio Matute, por su paciencia conmigo desde el día uno.

A Dra. Krisztina Ríos-Fulop y Dra. Alejandra Zamora, por creer en mí desde que me conocieron, y por preocuparse por nosotros más allá de lo académico.

A M.Sc. Miguel Morales y Lic. Marie Cosenza, por motivarme a dar siempre lo mejor de mí, especialmente como auxiliar y persona.

A cada uno de los profesores que me enseñaron, formaron, motivaron e inspiraron de alguna u otra forma.

A mis papás, Christian Porras y Lorena Moscoso, por ser mis pilares, creer en mí y motivarme, apoyarme y sostenerme siempre. A cada una de las personas de mi familia, que me ayudaron a culminar con esta etapa.

A mis compañeros de carrera, mis pollitos, con especial mención a Andrés Say, Melissa Caballeros, Silvia Vela, Valeria Overdick y Andrea de la Vega, por no dejarme sola nunca y apoyarme en las buenas, pero, sobre todo, en las malas.

Índice

Prefacio	5
Índice	6
Listado de figuras.....	8
Listado de cuadros	9
Resumen	X
Abstract.....	XI
I. Introducción	1
II. Objetivos	2
i. Objetivo general.....	2
ii. Objetivos específicos	2
III. Justificación.....	3
IV. Marco teórico	4
i. Fermentación en alimentos	4
ii. Fermentación de cacao	7
Fuente:	8
iii. Análisis de microbiomas.....	8
iv. Genes <i>16S</i> e <i>ITS</i>	9
v. Metabolómica predictiva	9
vi. Herramientas bioinformáticas	10
vii. Análisis estadístico	13
V. Marco metodológico	15
i. Variables.....	15
<i>Nota:</i> elaboración propia.....	15
ii. Hipótesis	15
VI. Materiales y métodos	17
i. Obtención de datos.....	17
ii. Análisis de datos de secuencia de amplicones de ARNr <i>16S</i>	17
iii. Análisis de datos	17

iv.	Flujo de trabajo – análisis metagenómico ARNr 16S	18
	VII. Resultados	18
i.	Abundancia relativa	18
ii.	Riqueza	19
iii.	Diversidad alfa	20
iv.	Diversidad beta	21
v.	Prueba estadística no paramétrica	23
vi.	Flujo de trabajo – genoma completo	23
	VIII. Análisis de resultados	25
	IX. Conclusiones	32
	X. Recomendaciones	33
	XI. Referencias	34

Listado de figuras

Figura 1: Algoritmo para análisis metagenómico del proceso fermentativo del cacao. (Elaboración propia)	18
Figura 2: Abundancia relativa por sitio de estudio. (Elaboración propia)	19
Figura 3: Gráficos de diversidad beta. (Elaboración propia)	22
Figura 4: Flujo de trabajo propuesto para análisis de microbiota del proceso fermentativo del cacao para datos de genoma completo.	24

Listado de cuadros

Cuadro 1: Precios de grano de cacao comercial proveniente de cada país analizado y características organolépticas asociadas	8
Cuadro 2: conceptualización, clasificación y operacionalización de las variables de estudio	15
Cuadro 3: Riqueza de especies por sitio de estudio.....	20
Cuadro 4: Índices de diversidad alfa.	21
Cuadro 5: Comparación de diversidad alfa y beta en muestras de fermentación de cacao	23

Resumen

El objetivo de este estudio fue comparar los perfiles metagenómicos del proceso fermentativo del cacao de seis sitios distintos. Los análisis bioinformáticos se realizaron utilizando seis bases de datos provenientes de México, Costa Rica, Camerún y Colombia. Se hizo uso del paquete *dada2* en *RStudio*, se realizó un análisis de limpieza de ARNr *16S* y se les asignó taxonomía a las muestras. Posteriormente, se determinó la abundancia relativa y riqueza de las especies presentes y se compararon las muestras utilizando índices de diversidad alfa y beta. Se llevó a cabo un análisis estadístico y se compararon los resultados con los precios de mercado mundial según *Selina Wamucii* (2023), cuyo estudio presentó que el cacao proveniente de México era el mejor cotizado, con un precio de 4.10 a 5.00 USD por kilogramo. Los análisis de abundancia revelaron una presencia significativa del filo Proteobacteria en todos los sitios, con excepción de México, lo cual es consistente con la significancia ecológica de la fermentación de cacao. La dominancia del filo Actinobacteriota en México sugiere beneficios antimicrobianos y un potencial biocontrol. Asimismo, el filo Cyanobacteria, observado en Camerún y Costa Rica, podría influir en la composición del producto final, debido a sus habilidades fijadoras de nitrógeno y producción de compuestos bioactivos. La alta diversidad alfa, especialmente en Colombia y Costa Rica, indican comunidades microbianas muy ricas. A pesar de la diferencia geográfica, no se encontraron diferencias significativas en las muestras. Finalmente, este estudio subraya la interacción compleja entre la diversidad microbiana, la calidad del mercado y la percepción del mercado. A pesar de las limitaciones, el mismo contribuye con perspectivas valiosas a la diversidad microbiana en los sitios de estudio, pero enfatiza la necesidad de interpretación cuidadosa de resultados. Asimismo, sugiere áreas para investigaciones futuras en este campo.

Palabras clave: *metagenómica, fermentación de cacao, diversidad microbiana*

Abstract

The aim of this study was to compare the metagenomic profiles of the cacao fermentation process from six different sites. Bioinformatics analyses were performed using six databases from Mexico, Costa Rica, Cameroon, and Colombia. The *dada2* package in *RStudio* was used, followed by a *16S* rRNA cleaning analysis and taxonomy assignment to the samples. Subsequently, the relative abundance and species richness were determined, and the samples were compared using alpha and beta diversity indices. A statistical analysis was conducted, and the results were compared with global market prices according to *Selina Wamucii* (2023), whose study indicated that cacao from Mexico was the highest priced, ranging from 4.10 to 5.00 USD per kilogram. Abundance analyses revealed a significant presence of the Proteobacteria phylum at all sites except Mexico, which is consistent with the ecological significance of cacao fermentation. The dominance of the Actinobacteriota phylum in Mexico suggests antimicrobial benefits and potential biocontrol. Additionally, the Cyanobacteria phylum, observed in Cameroon and Costa Rica, could influence the final product's composition due to its nitrogen-fixing capabilities and production of bioactive compounds. High alpha diversity, especially in Colombia and Costa Rica, indicates very rich microbial communities. Despite the geographical differences, no significant differences were found between the samples. Finally, this study highlights the complex interaction between microbial diversity, market quality, and market perception. Despite the limitations, it provides valuable insights into microbial diversity at the studied sites, but emphasizes the need for careful interpretation of the results. It also suggests areas for future research in this field.

Key words: *metagenomics, cocoa fermentation, microbial diversity*

I. Introducción

La presente tesis buscó realizar comparaciones de perfiles metagenómicos de la fermentación del cacao dentro de distintos perfiles moleculares originarios de seis sitios de estudio diferentes, obtenidos de servidores públicos. Esto se llevó a cabo haciendo uso de diferentes herramientas bioinformáticas. El propósito fue poder asociar los perfiles metagenómicos a distintos sitios de estudio y realizar comparaciones entre sí. Asimismo, se buscó generar un algoritmo para el análisis metagenómico del cacao, para poder establecer relaciones e hipótesis entre la calidad del chocolate producido y su microbiota en el futuro. Esto con el fin de sentar las bases para, potencialmente, realizar modificaciones a las microbiotas y mejorar las condiciones organolépticas del cacao, según su perfil metagenómico.

El cacao es un cultivo importante económicamente, cultivado principalmente por sus granos, los que son indispensables para la producción de chocolate. El sabor, aroma y calidad del chocolate están fuertemente influenciados por el proceso fermentativo de los granos (Afoakwa, 2010; Nielsen, 2019). Durante este proceso, ocurren transformaciones bioquímicas complejas, que son provocadas por las actividades de varios microorganismos presentes en la pulpa del cacao (Schwan y Wheals, 2004). Estas comunidades microbianas juegan un papel importante en el desarrollo de precursores de sabor deseables, así como de la degradación de compuestos no deseados (Payne, *et al*, 2010).

La comprensión de las dinámicas y composición de la microbiota del cacao durante su fermentación es de gran interés para los productores de cacao y de chocolate, así como para investigadores en ese campo (Illegheems, De Vuyst y Wecks, 2012). Avances recientes en las técnicas de biología molecular han permitido que se realice una exploración más profunda de las comunidades microbiológicas asociadas con varios ecosistemas, incluyendo los procesos de fermentación de cacao (Quintero, Santos, Lima y Bemquerer, 2013). Especialmente, el uso y abaratamiento de tecnologías de secuenciación ha revolucionado la habilidad para caracterizar la diversidad microbiana e identificarlos microorganismos clave que están envueltos en los procesos específicos (De Phillipis, *et al*, 2018).

En el presente trabajo, se buscó investigar y comparar los perfiles metagenómicos de la fermentación del cacao dentro de distintos perfiles moleculares originarios de seis estudios de sitio distintos. Asimismo, se buscó realizar una comparación de estas

secuencias y relacionarlas al chocolate que se produce a partir de esos procesos fermentativos en dichos sitios.

Estudios previos han demostrado la importancia de la diversidad microbiana y sus genes funcionales en el proceso fermentativo y desarrollo de los perfiles de sabor deseables (Papalexandratou, et al, 2011). Sin embargo, todavía falta comparar rutas metabólicas e interacciones involucradas en la fermentación del cacao de distintos lugares.

II. Objetivos

i. Objetivo general

Comparar los perfiles metagenómicos de la fermentación del cacao dentro de distintos perfiles moleculares originarios de seis sitios de estudio, provenientes de México, Costa Rica, Camerún y Colombia, los cuales fueron obtenidos de servidores públicos.

ii. Objetivos específicos

1. Generar un algoritmo para el análisis metagenómico de cacao y, con esto, lograr una comparación entre los perfiles organolépticos del chocolate producido.
2. Hacer una comparación de las secuencias encontradas de los seis lugares y relacionarlas al precio del chocolate que se produce a partir de esos procesos fermentativos.
3. Proponer un flujo de trabajo que permita realizar el mismo análisis, utilizando datos de genoma completo.

III. Justificación

El cacao es un producto altamente producido y comercializado en el mundo, con aproximadamente 4.9 millones de toneladas de cacao producidas a nivel mundial de 2021 a 2022 (International Cocoa Organization, 2022). La producción de cacao ha aumentado de forma constante en los últimos 40 años, a pesar de que es bastante volátil y está sujeto a varias fluctuaciones (Bermudez, *et al*, 2022). La producción de cacao fino de aroma en Guatemala es indispensable para la chocolatería, por su alta calidad organoléptica, posicionándose como el principal país exportador de Centroamérica (Aguirre Sarceño, 2022). Según reportes del MAGA (2014), un 96% del cacao producido pasa por procesos transformativos dentro del país, por lo que únicamente un 4% es exportado sin sufrir cambios.

La fermentación del cacao es un proceso espontáneo, dirigido por muchos microorganismos, dentro de los que se pueden mencionar levaduras, bacterias y algunos hongos filamentosos (Schwan y Wheals, 2004). Este proceso es indispensable para poder desarrollar los precursores del sabor y aroma del chocolate (Aguirre Sarceño, 2022). La fermentación debe de ser asociada al tipo de cacao disponible y dura aproximadamente siete días, aunque este número depende de muchos otros factores. La detección de los picos de ácido láctico y acético y, por consecuencia, la comparación con otras especies podría ayudar a elucidar un poco el proceso fermentativo (Nielsen, 2019). A futuro se podrían llegar a comparar estos picos con los de otras especies y buscar asociarlas a ciertos sabores o aromas del cacao, manipulando la microbiota del mismo.

Debido a que el proceso fermentativo debe de ser asociado al tipo de cacao disponible, es importante poder realizar comparaciones entre cacaos provenientes de distintos lugares. Asimismo, el poder generar un algoritmo para el análisis metagenómico del cacao, podrían sentar las bases para establecer relaciones entre la calidad del chocolate producido y su microbiota en el proceso fermentativo.

IV. Marco teórico

i. Fermentación en alimentos

La fermentación en alimentos es un proceso natural que se ha utilizado por los humanos por miles de años (Tamang, *et al*, 2016). La función principal es la preservación de comida y el mejoramiento de su sabor, así como textura y valor nutricional (Marco y Heeney, 2017). La fermentación utiliza microorganismos como bacterias, levaduras y hongos para poder descomponer compuestos complejos a compuestos más simples (Rezac, *et al*, 2018). Esto puede ser importante para mejorar su digestibilidad y crear sabores y aromas nuevos (Hughey y Johnson, 2019). Este proceso puede ocurrir de forma natural y espontánea, o puede ser inducido al agregarle microorganismos específicos a la comida (Hutkins, 2018). Los ejemplos más comunes de este proceso son el yogurt, el queso, el kimchi y el pan de masa fermentada (Schwierzy Taras, 2014).

Este proceso ocurre en tres pasos, según Ozturkoglu-Budak y Akman (2020):

1. Inicio: la comida se expone al aire y microorganismos la colonizan. Estos consumen los azúcares disponibles, produciendo ácidos orgánicos, los cuales disminuyen el pH de la comida. Al disminuir este pH, se crea un ambiente ácido que limita el crecimiento de bacterias patógenas.
2. Medio: cepas específicas de microorganismos se agregan al alimento, para controlar el proceso fermentativo y fomentar los sabores y olores deseados. Estos microorganismos pueden producir enzimas que rompen las proteínas y carbohidratos, lo cual es esencial para mejorar la textura y digestibilidad de la comida.
3. Final: el proceso se vuelve más lento, porque los azúcares disponibles son completamente consumidos. El pH se vuelve demasiado ácido, como para permitir el crecimiento o colonización de microorganismos. El producto final es normalmente estable y con mucho sabor, y puede ser guardado por periodos muy largos de tiempo.

Diversos factores pueden afectar el proceso fermentativo, como, por ejemplo, el tipo y concentración de microorganismos, la temperatura y humedad del ambiente y la composición de la comida (Steinkraus, 1997). Estos pueden afectar el producto final, no solo en términos de sabor y textura, sino de valor nutricional (Marco y Heeney, 2019). Los

alimentos fermentados tienen beneficios a la salud, esto se debe a que contienen probióticos, vitaminas y minerales (Tamang, 2015).

A continuación, se enlistan distintos alimentos comunes, los cuales son sometidos a procesos de fermentación.

a. Yogurt

El proceso de fermentación de yogurt usa bacterias ácido-lácticas para convertir lactosa en ácido láctico (Hui y Özgül Evranuz, 2018). Esto provoca que la leche se espese y vuelva más agria, además de obtener una consistencia cremosa (Hui y Özgül Evranuz, 2018). Las bacterias responsables para este proceso son principalmente cepas de *Lactobacillus bulgaricus* y *Streptococcus thermophilus*; sin embargo, existen casos documentados en los que se ven otras bacterias involucradas (Fortsythe, 2018). El proceso fermentativo inicia calentando la leche para desnaturalizar sus proteínas y crear ambientes estables para que crezcan las bacterias (Broadbent, *et al*, 2018). Luego, se enfría y se inoculan con un cultivo inicial de bacterias vivas. La mezcla se incuba a aproximadamente 42-45°C por varias horas, hasta que alcance el nivel de acidez esperada (Broadbent, *et al*, 2018). Cuando el proceso esté completo, se enfría el producto y se endulza (Tamang, *et al*, 2016). La fermentación del yogurt aumenta la digestibilidad de la leche (para personas intolerantes a la lactosa), además de que aumenta la biodisponibilidad de algunos nutrientes, como fósforo y calcio (Tamang, *et al*, 2016).

b. Cerveza

La fermentación de la cerveza utiliza a levaduras, que convierten los azúcares presentes en mosto, a alcohol y dióxido de carbono (Bamforth, 2019). El proceso normalmente se lleva a cabo en un fermentador y puede tomar de algunos días hasta varias semanas, dependiendo del estilo de cerveza que se busca producir y la temperatura de fermentación (Priest y Stewart, 2006). Hay dos tipos de fermentación de cerveza: de alta y de baja fermentación (Steiner y Hieronymus, 2014). En las cepas de levadura de alta fermentación (llamadas levaduras *ale*), así como *Saccharomyces cerevisiae*, se fermenta a temperaturas más altas y tienden a producir cervezas con aromas y sabores más frutales y picantes (Steiner y Hieronymus, 2014). En las cepas de levadura de baja fermentación (llamadas levaduras *lager*), la fermentación ocurre a temperaturas más bajas y producen cervezas con sabores y aromas más limpios (Steiner y Hieronymus, 2014). El contenido de alcohol de las cervezas típicamente está en un 3-12% (volumen/volumen), dependiendo del estilo y condiciones de fermentación (Lewis y Young, 2002). Después de esta fermentación, la cerveza normalmente se añeja, para permitir que los sabores y aromas maduren juntos (Steiner y Hieronymus, 2014).

c. Kimchi

El kimchi es un plato de coreano basado en vegetales fermentados (Jeong y Chun, 2007). Tiene un sabor único y un alto valor nutricional (Jeong y Chun, 2007). Se utiliza repollo, rábano, ajo, jengibre y polvo de pimienta roja, con sal y especias para sazonar (Jeong y Chun, 2007). Se deja fermentar a temperatura ambiente por periodos variables, de entre algunos días a varias semanas (Jeong y Chun, 2007). Durante este proceso, el pH baja, la acidez aumenta y se desarrollan los sabores (Kim y Chun, 2005). Se han encontrado bacterias ácido-lácticas y levaduras, principalmente, pero también se han encontrado bacterias de los géneros *Weissella* sp., *Pediococcus* sp. y *Bacillus* sp. (Lee, et al, 2019; Han et al, 2020) Debido a la gran variedad de microorganismos, se ha considerado que el kimchi es un platillo con muchos probióticos, lo que ayuda a la digestión (Park, Jeong, Lee y Daily III, 2014). Además, presenta propiedades antioxidantes y antiinflamatorias (Park, et al, 2014).

d. Quesos

Los quesos son productos lácteos importantes, que también pasan por un proceso fermentativo, en el que se desarrollan sus sabores y texturas, así como las propiedades necesarias para su preservación (Fox y McSweeney, 2017). Para la fermentación de queso, se utilizan levaduras, bacterias ácido-lácticas y mohos (McSweeney y Sousa, 2000). El tipo de queso establece qué tipo de microorganismos están presentes en el proceso (Adams y Moss, 2007). Las bacterias clave son especies de *Lactococcus*, *Streptococcus* y *Lactobacillus* (Smit, et al, 2005). También se encuentran especies de levaduras de los géneros *Debaryomyces*, *Kluyveromyces* y *Candida* (Smit, et al, 2005). Finalmente, los tipos de moho pueden ser especies de los géneros *Penicillium* y *Geotrichum* (Smit, et al, 2005). Los quesos se pueden clasificar de acuerdo con su proceso de fermentación y condiciones de maduración (Fox y McSweeney, 2017).

e. Sauerkraut

El sauerkraut es un alimento de repollo fermentado, que ha sido consumido por siglos (Cherif, et al, 2019). El periodo de fermentación a temperatura ambiente es variado, pero suele ser de varias semanas. Se busca desarrollar el sabor, textura y aroma (Halász, Baráth y Holzapfel, 1999). Dentro de la fermentación, se pueden encontrar especies dominantes de bacterias ácido-lácticas, como *Lactobacillus plantarum*, *Leuconostoc mesenteroides* y *Pediococcus* sp. (Wouters, et al, 2016). El sauerkraut es solicitado debido a sus efectos probióticos y su actividad antioxidante (Cherif, et al, 2019).

f. Kombucha

La kombucha es una bebida, que se obtiene a través de la fermentación de té endulzado con un cultivo simbiótico de bacterias y levaduras (SCOBY, por sus siglas en inglés) (Jayabalan, et al, 2014). Dentro de los microorganismos involucrados, se encuentran

especies de bacterias de los géneros *Acetobacter*, *Gluconacetobacter* y *Lactobacillus*; mientras que las levaduras son especies de los géneros *Saccharomyces*, *Zygosaccharomyces* y *Brettanomyces* (Greenwalt, *et al*, 2000). La komucha ha sido asociada a muchos beneficios potenciales, especialmente en cuanto a las propiedades antioxidantes y la presencia de probióticos (Greenwalt, Steinkraus y Ledford, 2000).

ii. Fermentación de cacao

La fermentación del cacao es un paso indispensable para la producción de chocolate de la más alta calidad, ya que permite el desarrollo del aroma y sabor específicos del cacao fino de aroma (Schwan y Wheals, 2004). Durante esta fermentación, los granos de cacao atraviesan cambios químicos y microbiológicos (Perea Villamil, *et al* 1998). Estos resultan en el rompimiento de compuestos, que liberan precursores de sabor (Mota-Gutierrez, *et al*, 2019). El proceso fermentativo típicamente dura entre dos a siete días, dependiendo de la región y del tipo de cacao (Meersman, *et al* 2013).

Los microorganismos implicados juegan un papel indispensable en la fermentación de este compuesto, ya que son responsables de iniciarla y dirigirla (De Vuyst y Weckx, 2016). Los más importantes son levaduras y bacterias ácido-lácticas, los cuales están presentes en el medio ambiente y colonizan los granos de cacao (Mota-Gutierrez, *et al*, 2018). Los microorganismos consumen los azúcares disponibles y, a cambio, producen ácidos orgánicos, alcoholes y otros compuestos que contribuyen al sabor y aroma del cacao final (Illegheems, *et al*, 2012)

La microbiota de los granos de cacao depende muchísimo de varios factores, especialmente de la región, el clima y las prácticas fermentativas utilizadas (Jung, *et al*, 2011). Por ejemplo, un estudio publicado por Illegheems, *et al* (2012) comparó la diversidad microbiana en el proceso fermentativo del cacao en distintas regiones geográficas y condiciones de procesamiento. Se encontraron variaciones en la composición microbiana en los granos de cacao de distintas regiones climáticas y prácticas fermentativas (Illegheems, *et al*, 2012). Asimismo, un estudio por Afoakwa, *et al* (2014) analizó el impacto de los métodos de procesamiento luego de la cosecha en la fermentación del cacao y la calidad del chocolate producido. Las variaciones en las poblaciones microbianas y cinéticas de la fermentación demostraron diferencias en los atributos sensoriales y características de calidad del chocolate (Afoakwa, *et al*, 2014). Para mantener la calidad y consistencia del producto final, es importante comprender la ecología de la microbiota del cacao, así como de su fermentación, y cómo estos procesos afectan el producto final (Pereira, *et al*, 2016).

Guatemala cuenta con aproximadamente 250,000 pequeños agricultores, que cultivan cacao fino de aroma ((USDA Foreign Agricultural Service, 2019). A pesar de su rica historia y su sabor único y cotizado mundialmente, Guatemala solo representa un 0.3% de la producción mundial del cacao, con a penas unas 1,500 toneladas anuales (Ministerio de Agricultura, Ganadería y Alimentación, 2014). En Guatemala, el cacao es normalmente fermentado en cajas de madera forradas con hojas de plátano, por entre 4 a 8 días,

mezclando los granos con la mano regularmente (USDA Foreign Agricultural Service, 2019).

A continuación, en el **Cuadro 1** se presentan los precios de grano de cacao comercial proveniente de cada país analizado. Asimismo, se pueden observar las características de sabor, aroma, textura y perfil de sabor representativas de cada uno de ellos.

Cuadro 1: Precios de grano de cacao comercial proveniente de cada país analizado y características organolépticas asociadas

País	Precio (USD/kg)	Características organolépticas			
		Sabor	Aroma	Textura	Perfil de sabor
Camerún	3.95 - 4.94	Notas afrutadas y florales	Toques cítricos y a especias	Suave, aterciopelada	Complejo y diverso
Colombia	1.97 - 4.81	Notas frutales, con toques cítricos y de bayas	Ricos y complejos, matices florales	Suave	Desde afrutados hasta terrosos
Costa Rica	3.61 - 5.77	Notas florales y frutales, con toques cítricos	Aromas frescos y limpios	Suave, sedoso	Variado
México	4.21 - 5.00	Notas terrosas, especiadas, toques de frutas secas	Aromas ricos, matices de canela y vainilla	Rugoso (en comparación)	Único, tradicional

Fuente: *Wamucii, S. (2023)*

iii. Análisis de microbiomas

El análisis de microbiomas se refiere al estudio de comunidades de microorganismos que viven en un ambiente específico, como, por ejemplo, el proceso fermentativo de cacao (Lozupone y Knight, 2005). El analizar la composición y función de la microbiota, puede significar conocimientos importantes sobre cómo funciona y cómo puede ser manipulado (Caporaso, *et al*, 2011).

Existen diversos métodos para su análisis, cada uno con sus fortalezas y limitaciones (Knight, *et al*, 2012). Dentro de los más comunes están los siguientes:

a. Métodos basados en cultivo

Estos involucran crecer microorganismos de una muestra en un medio de agar rico en nutrientes (Knight, *et al*, 2012). A pesar de que estos métodos proveen una visión más detallada de los organismos presentes, los resultados podrían estar sesgados hacia ciertas bacterias y se pierden las especies problemáticas para cultivar en laboratorio (Knight, *et al*, 2012).

b. Métodos moleculares

Se refiere a utilizar métodos moleculares para analizar ADN o ARN del microorganismo presente en la muestra (Knight, *et al*, 2012). Los más comunes son PCR (reacción de la polimerasa en cadena) y secuenciación (Knight, *et al*, 2012). Estos métodos

son más sensibles y pueden detectar un amplio espectro de especies. Sin embargo, son más caros y necesitan más especialización por parte de los laboratoristas (Knight, *et al*, 2012).

c. **Metagenómica**

Esto involucra la secuenciación de todas las moléculas de ADN presentes en la muestra y se pueden utilizar para la identificación de especies que son problemáticas para cultivar en el laboratorio (Agyirifo, *et al*, 2019). Además, la caracterización de las comunidades es el primer paso para realizar un estudio de la presencia de rutas metabólicas específicas o genes de resistencia a antibióticos (Agyirifo, Wamalwa y Otwe, 2019).

iv. **Genes 16S e ITS**

El gen ARN ribosomal *16S* y la región *ITS* (espaciador transcrito interno) son comúnmente usadas en la ecología microbiana para identificar y caracterizar comunidades microbianas (Bukin, *et al*, 2019). Estas son regiones altamente conservadas; sin embargo, varían lo suficiente entre especies para permitir su diferenciación (Durazzi, *et al*, 2021).

El gen *16S* ARNr es un componente del ribosoma de procariotas y está presente en todas las bacterias y arqueas (Johnson, *et al*, 2019). El gen está compuesto de nueve regiones hipervariables, que cambian en el largo y secuencia entre especies (Church, *et al*, 2020). El alto grado de variabilidad en la secuencia de estas regiones hace que el gen *16S* ARNr sea una herramienta indispensable para la caracterización de la diversidad microbiana, así como para la identificación de bacterias y arqueas (Muhamad Rizal, *et al*, 2020). La región *ITS* está localizada entre las subunidades pequeña y grande de genes ARNr eucarióticos (Gardes y Bruns, 1993). Esta consiste en dos subregiones, *ITS1* e *ITS2*, las cuales varían en largo y su secuencia entre diferentes especies de plantas y hongos (Hall, 1999). El *16S* y la *ITS* son secuenciadas comúnmente utilizadas para generar bases de datos largas de perfiles de la comunidad microbiana (Al-Mahrous y Shah, 2016). Las secuencias resultantes pueden ser analizadas con herramientas de bioinformática para asignar identidades taxonómicas a las secuencias (Buck, Götz, Romanowski y Mühlenweg, 2015). De esta forma, se pueden comparar estructuras de comunidades microbianas entre muestras distintas (Li, Soeckert y Roos, 2003).

A pesar de su amplio uso, existen algunas limitaciones para *16S* y la *ITS* (Lou, *et al*, 2016), debido a que el *16S* solo está presente en bacterias y arquea, mientras que la *ITS* solo está presente en eucariotas (Buck, *et al*, 2015). Asimismo, la precisión de la asignación taxonómica puede verse afectada por la integridad de la base de datos y los errores en la identificación de la secuencia (Buck, *et al*, 2015).

v. **Metabolómica predictiva**

La metabolómica predictiva es un campo relativamente nuevo que utiliza perfiles de metabolitos para predecir las respuestas biológicas de organismos ante distintos factores ambientales y experimentales (Wishart, 2016). La metabolómica es el estudio

de pequeñas moléculas de metabolitos (típicamente, con pesos moleculares menores a 1,000 Da) en sistemas biológicos (Benton, *et al*, 2008). Es una herramienta indispensable para comprender procesos bioquímicos que ocurren en organismos vivos (Creek, *et al*, 2014).

En la metabolómica predictiva, los perfiles de metabolitos son generados en diferentes condiciones experimentales o diferentes tratamientos, utilizando técnicas analíticas como espectrometría de masas o espectroscopía de resonancia magnética nuclear (Gromski, *et al*, 2015). Estos perfiles son luego utilizados para predecir las respuestas biológicas de los organismos (Guan, Zhang y Yan, 2017).

El poder predictivo de los modelos metabolómicos depende de la calidad e integridad de los datos del perfil metabolómico (Kim, *et al*, 2016). Los algoritmos de *machine learning* son utilizados para crear modelos predictivos en los perfiles de metabolitos, lo cual puede identificar patrones en los datos que están asociados a respuestas biológicas particulares (Lenz, *et al*, 2004).

Sin embargo, existen algunos retos asociados a la metabolómica predictiva (Noga, *et al*, 2016). Uno de los mayores retos es la necesidad de métodos analíticos estandarizados, así como de procedimientos de procesamiento de datos, para asegurar la consistencia y reproducibilidad de resultados a través de diferentes estudios (Righetti, *et al*, 2016). Otro reto es la necesidad de la creación de bases de datos grandes y diversas, para poder crear modelos predictivos acertados (Righetti, *et al*, 2016).

vi. Herramientas bioinformáticas

a. DADA2

El paquete *DADA2* (Callahan, *et al*, 2016) permite hacer inferencias exactas de la secuencia variante de amplicón (ASV, por sus siglas en inglés) de datos de secuenciación de amplicones de alto rendimiento, sustituyendo el método de agrupación de unidades taxonómicas operacionales (OTU, por sus siglas en inglés), que es menos preciso (Callahan, *et al*, 2016). El flujo de trabajo de *dada2* toma como entrada archivos de tipo *fastq* demultiplexados y genera variantes de secuencia y su abundancia en la muestra, luego de eliminar errores de sustitución y de quimera (secuencias que no son productos reales de la amplificación del gen *16S*) (Callahan, *et al*, 2016). La clasificación taxonómica está disponible a través de una implementación nativa del clasificador bayesiano ingenuo RDP (clasificador probabilístico fundamentado en el teorema de Bayes para asignación taxonómica) y la asignación a nivel de especie a fragmentos del gen *16S* ARNr por coincidencia exacta (Callahan, *et al*, 2016).

b. Trimmomatic

Se define como una herramienta flexible para recortar datos de extremo emparejado y extremo único (Bolger, *et al*, 2014). La selección de los pasos de recorte y sus parámetros asociados se presentan en la línea de comando (Bolger, *et al*, 2014). Puede utilizar datos comprimidos o descomprimidos de FASTQ (Bolger, *et al*, 2014).

c. **FastQC**

Es una herramienta bioinformática que provee una forma simple de hacer chequeos de control de calidad en flujos de trabajo de secuencias de alta definición (Andrews, 2010). Ofrece un conjunto modular de análisis que pueden ser utilizados para hacerse una idea rápida de si los datos presentan algún problema que deba de ser tomado en cuenta antes de seguir analizándolos (Andrews, 2010). Dentro de sus principales funciones se encuentra la importación de datos de archivos FastQ, el proveer un vistazo de en cuáles áreas debería de encontrarse problemas, un resumen de gráficas y cuadros para una evaluación de los datos y la operación fuera de línea para permitir la generación de reportes automática sin la necesidad de correr la aplicación interactiva (Andrews, 2010).

d. **MetaBAT**

Herramienta de *software* automatizada de agrupación de metagenomas para la reconstrucción de genomas individuales de comunidades microbianas, para posteriores análisis de especies microbianas no cultivadas (Kang, *et al*, 2015). *MetaBAT* se ha convertido en una de las herramientas de agrupamiento más populares debido a su eficiencia computacional y facilidad de uso, especialmente en experimentos de agrupamiento con un alto número de muestras y un gran ensamblaje (Kang, *et al*, 2015). La herramienta requiere que los usuarios elijan parámetros de sensibilidad y especificidad. Si estos no son elegidos correctamente, la precisión del agrupamiento puede verse afectada, especialmente en ensamblajes de baja calidad (Kang, *et al*, 2015).

e. **Kraken**

Es un programa muy rápido y con alta precisión para asignar etiquetas taxonómicas a secuencias de ADN metagenómico, que permite alcanzar clasificaciones con exactitud comparable a cualquier programa de BLAST (Wood y Salzberg, 2014). Esta herramienta utiliza una base de datos de referencia que contiene información sobre secuencias genéticas de una amplia gama de organismos y es particularmente útil en programas de metagenómica, ya que permite analizar comunidades microbianas y clasificarlas en función de los organismos presentes (Wood y Salzberg, 2014).

f. **MetaPhlAn**

MetaPhlAn (*Metagenomic Phylogenetic Analysis*) [Análisis filogenético metagenómico] es una herramienta computacional para el perfilamiento de la composición de comunidades microbianas de datos de secuenciación metagenómico a un nivel de especie, proporcionando información valiosa (Truong, *et al*, 2017). Permite identificar la abundancia relativa y presencia de diferentes microorganismos en las muestras analizadas. MetaPhlAn se apoya en aproximadamente 5.1 millones de genes marcadores específicos paraclados de aproximadamente un millón de genomas microbianos (Truong, *et al*, 2017). Esto permite una asignación taxonómica no ambigua, una estimación precisa de la abundancia relativa de cada organismo y la identificación y seguimiento de cepas (Truong, *et al*, 2017).

g. PICRUS2

PICRUS2 (*Plylogenetic Investigation of Communities by Reconstruction of Unobserved States*) [Investigación Filogenética de Comunidades por Reconstrucción de Estados Sin Observar] es un paquete de *software* bioinformático diseñado para predecir abundancias funcionales basadas en un gen marcador (por ejemplo, 16S ARNr). Esta versión contiene una base de datos mejorada de familias de genes y genomas de referencia. Perfiles funcionales no pueden ser identificados directamente utilizando los datos de secuencia del gen 16S ARNr debido a variaciones en las cepas, por lo que se han creado muchos métodos para predecir la función de la comunidad microbiana a partir de solo perfiles taxonómicos (secuencias de amplicones). La herramienta de PICRUS2 incluye pasos para optimizar la predicción de genoma, incluyendo colocar secuencias en filogenia de referencia, en lugar de confiar en predicciones limitadas a referencias de OTUs (Douglas, *et al*, 2020).

h. HUMAnN

Esta herramienta es importante para crear perfiles de forma eficiente y precisa de la abundancia de rutas metabólicas y otras funciones moleculares a partir de datos de metagenómica o metatranscriptómica (Beghini, *et al*, 2021). También ayuda a crear un perfil de la presencia o ausencia de rutas metabólicas (Beghini, *et al*, 2021). A este proceso se le conoce como perfilamiento funcional y busca describir el potencial metabólico de una comunidad microbiana y sus miembros (Beghini, *et al*, 2021).

i. KEGG

KEGG (*Kyoto Encyclopedia of Genes and Genome*) [Enciclopedia de genes y genomas de Kyoto] es un recurso de bases de datos para el entendimiento de funciones de alto nivel y utilidades de los sistemas biológicos, como la célula, el organismo y el ecosistema, de información a nivel molecular, grupos de datos moleculares especialmente de gran tamaño, generados por secuenciación genómica y otras tecnologías experimentales de alto rendimiento (Kanehisa, *et al*, 2021). KEGG se ha apoyado en la inteligencia humana para el desarrollo de modelos de sistemas biológicos, especialmente en la forma de mapas de rutas KEGG que son creados manualmente al capturar conocimiento de literatura publicada (Kanehisa, *et al*, 2021). Luego, los modelos KEGG pueden ser utilizados en el análisis de macrodatos biológicos, por ejemplo, para el descubrimiento de funciones sistémicas de un organismo ocultas en su secuencia genómica mediante el sencillo procedimiento de mapeo KEGG (Kanehisa, *et al*, 2021). Este mapeo asigna secuencias genéticas o proteínas a las vías metabólicas y funciones biológicas representadas en la misma base de datos (Kanehisa, *et al*, 2021). El mapeo incluye vías metabólicas y análisis funcional de datos genómicos o proteómicos (Kanehisa, *et al*, 2021).

j. Cytoscape

Es una plataforma de *software* de libre acceso para visualizar las redes de interacciones

moleculares y rutas metabólicas, con el fin de integrarlas con anotaciones, perfiles de expresión génica y otros datos (Shannon, *et al*, 2003). A pesar de que *Cytoscape* se diseñó originalmente para investigación biológica, ahora es una plataforma general para análisis de redes y su visualización (Shannon, *et al*, 2003)

vii. Análisis estadístico

A continuación, se enlistan distintos índices estadísticos analizados en el presente estudio.

a. Abundancia relativa

La abundancia relativa, en el contexto ecológico, se refiere a la proporción de individuos de una especie en particular, contemplada en una comunidad, del número total de individuos de todas las especies presentes en la misma. Es una medida de dominancia de una especie en un ecosistema (Zachariasen, *et al*, 2023). Este valor puede ser influenciado por varios factores, como lo son la competencia, predación y condiciones ambientales (Zachariasen, *et al*, 2023). Se usa, comúnmente, para estudiar la estructura y dinámica de comunidades ecológicas y comprender su estabilidad (Zachariasen, *et al*, 2023).

b. Riqueza

La riqueza, se refiere al número de especies presente en una comunidad; es una medida de biodiversidad y permite proveer un vistazo a una comunidad ecológica (Hirko, *et al*, 2023). Este concepto es importante para comprender la estructura y función de ecosistemas (Hirko, *et al*, 2023).

c. Diversidad alfa

Medida de biodiversidad que describe la diversidad de especies dentro de un ecosistema (Thukral, 2017). Es calculado al contar el número de especies y tomando en cuenta su abundancia relativa (Thukral, 2017). Se utiliza en conjunto con otras medidas, como diversidad beta, para probar una visión más completa de una comunidad ecológica (Thukral, 2017).

d. Índice de Shannon

El índice de Shannon ha sido un índice de diversidad bastante popular en la literatura ecológica (Magurran, 2004). Este índice mide la biodiversidad que predice la identidad de un individuo elegido al azar en una comunidad, que contiene varias especies (Shannon, 1948). Toma en consideración tanto la riqueza de la especie como la abundancia relativa (Shannon, 1948). Por lo tanto, el índice provee un valor único que refleja la diversidad y distribución de las especies en una comunidad (Shannon, 1948).

e. Índice de Chao1

Este índice es una medida de biodiversidad (Kim, *et al*, 2017). Se utiliza para estimar el número de especies en una comunidad, al considerar tanto las especies observadas como la riqueza estimada (Kim, *et al*, 2017). Es una herramienta muy útil para estudiar la diversidad alfa y es particularmente utilizada para estimar la riqueza de especies en comunidades con abundancias desiguales (Kim, *et al*, 2017).

f. Diversidad beta

Normalmente utilizada en conjunto con la diversidad alfa, es una medida de biodiversidad, que describe las diferencias en la composición de especies entre dos o más ecosistemas (Socolar, *et al*, 2016). Normalmente, se usa junto a la diversidad alfa (Socolar, *et al*, 2016). Comprender esta diversidad es crucial, ya que permite identificar áreas de alta biodiversidad y áreas en donde los cambios ambientales pueden tener impactos significativos en la distribución de especies y estructuras de la comunidad (Socolar, *et al*, 2016)

g. Índice de Bray-Curtis

Es una medida de la disimilitud de composición entre dos sitios o muestras distintas (Vega, *et al*, 2020). Toma en consideración la abundancia de distintas especies en las muestras, más que su presencia o ausencia (Vega, *et al*, 2020). Es comúnmente utilizado en ecología para comparar la similitud de composición de especies entre distintas muestras y es particularmente útil cuando se trata de datos que contienen muchos ceros (Vega, *et al*, 2020). Considera tanto la presencia y abundancia de elementos, basado en las diferencias en abundancia absoluta (Vega, *et al*, 2020).

4.7.5 Índice de Jaccard

También es una medida de disimilitud entre dos grupos de datos, pero se calcula como la tasa del número de elementos que no son compartidos por los conjuntos del número total de elementos en ambos sitios o muestras (Fletcher e Islam, 2018). Hace énfasis en la presencia o ausencia de elementos y es una medida de la proporción de elementos compartidos (Fletcher e Islam, 2018).

V. Marco metodológico

i. Variables

A continuación, se presenta la conceptualización, clasificación y operacionalización de las variables del presente estudio.

Cuadro 2: conceptualización, clasificación y operacionalización de las variables de estudio

	Clasificación	Definición conceptual	Definición operacional
Microbiota del cacao	Cualitativa	Comunidad de microorganismos, como bacterias, hongos y levaduras, que se encuentran en la superficie de los granos de cacao y su entorno, incluyendo el suelo, el agua y las plantas circundantes.	Comunidad microbiana que se identifica a través de técnicas de secuenciación.

Nota: elaboración propia

ii. Hipótesis

H_i: los perfiles metagenómicos de la fermentación del cacao dentro de distintos perfiles moleculares originarios de los seis lugares a estudiar tendrán diferencias significativas entre sí.

H₀: los perfiles metagenómicos de la fermentación del cacao dentro de distintos perfiles moleculares originarios de los seis lugares a estudiar no tendrán diferencias significativas entre sí.

H₁: la generación de un algoritmo para el análisis metagenómico de cacao demostrará si hay diferencias significativas en los perfiles organolépticos del chocolate producido en los seis sitios de estudio.

H₀: la generación de un algoritmo para el análisis metagenómico de cacao no demostrará si hay diferencias significativas en los perfiles organolépticos del chocolate producido en los seis sitios de estudio.

H₂: los resultados de la comparación de las secuencias de los seis diferentes sitios de estudio mostrarán diferencias significativas, que podrán ser relacionadas al precio del chocolate producido en los seis sitios.

H0: los resultados de la comparación de las secuencias de los seis diferentes sitios de estudio no mostrarán diferencias significativas, que no podrán ser relacionadas al precio del chocolate producido en los seis sitios.

VI. Materiales y métodos

i. Obtención de datos

Los datos fueron obtenidos de la página de *EMBL's European Bioinformatics Institute*, con las palabras clave “*cocoa fermentation*”. Se seleccionaron aquellos datos que contaran con un mínimo de 15 muestras. Las bases de datos utilizadas contenían distintos datos de secuenciación. Colombia contó con datos de tres sitios distintos, Valles Interandinos Secos (15 pares de datos), Montaña Santandereana (44 pares de datos) y Cafetalera Marginal Baja (38 pares de datos). Por otro lado, Costa Rica contó con 40 pares de datos, mientras que México contó con 20 pares de datos. Finalmente, la base de datos proveniente de Camerún contenía 39 pares de datos. No se contaba con metadata adjunta para ningún conjunto de datos.

ii. Análisis de datos de secuencia de amplicones de ARNr 16S

Se instaló y cargó la librería *DADA2* (Callahan, *et al*, 2016) a *RStudio*. Se cargaron los archivos en formato FASTQ que contenían las lecturas de secuenciación. Posteriormente, se realizó un filtrado de calidad y recorte de las secuencias, para luego realizar una corrección de error y verificar la inferencia de las variantes en secuencias. Se alinearon las lecturas *forward* y *reverse* para obtener secuencias de tamaño completo y se generó una tabla de variantes únicas y sus abundancias. Con ayuda de la base de datos SILVA (2020) para ARNr 16S, se realizó una asignación taxonómica a los variantes de secuencia de amplicones (ASV). Esto se repitió para cada una de las seis localidades.

iii. Análisis de datos

a. Diversidad y riqueza

Se cargaron los paquetes *dplyr* versión 1.1.3 (Wickham, *et al*, 2021), *ggplot2* versión 3.4.4 (Wickham, 2016) y *tidyr* versión 1.3.0 (Wickham, *et al*, 2021). Se leyeron los archivos que contenían la información de frecuencias de ASVs y los archivos que contenían la información taxonómica a nivel de filo. Posteriormente, se generaron gráficos de abundancia de filo por sitio de estudio y riqueza de filios por sitio.

b. Comparación de muestras

Se cargaron las librerías *permute* versión 0.9.7 (Simpson, 2022), *lattice* versión 0.21.08 (Sarkar, 2008), *vegan* versión 2.6.4 (Oskanen, *et al*, 2022) y *heatmap* versión 1.0.12 (Kolde, 2019). Se calcularon los índices de diversidad alfa Shannon y Chao1, y de diversidad beta, distancia de Bray-Curtis y Jaccard. Se realizaron gráficos para presentar los resultados.

c. Análisis estadístico

Se llevó a cabo una prueba de Kruskal-Wallis para verificar si había diferencia significativa entre los diversos sitios, utilizando los índices antes mencionados.

d. Verificación de precios de cacao

Se utilizaron reportes de Selina Wamucii (2023) para verificar los precios del cacao proveniente de los distintos países. Esta es una plataforma para comprar y vender productos agrícolas, que mantiene precios actualizados de los productos en el mercado actual.

iv. Flujo de trabajo – análisis metagenómico ARNr 16S

A continuación, se presenta un diagrama del algoritmo desarrollado (**Figura 1**) para analizar los datos metagenómicos encontrados y susubsiguiente comparación con precios publicados.

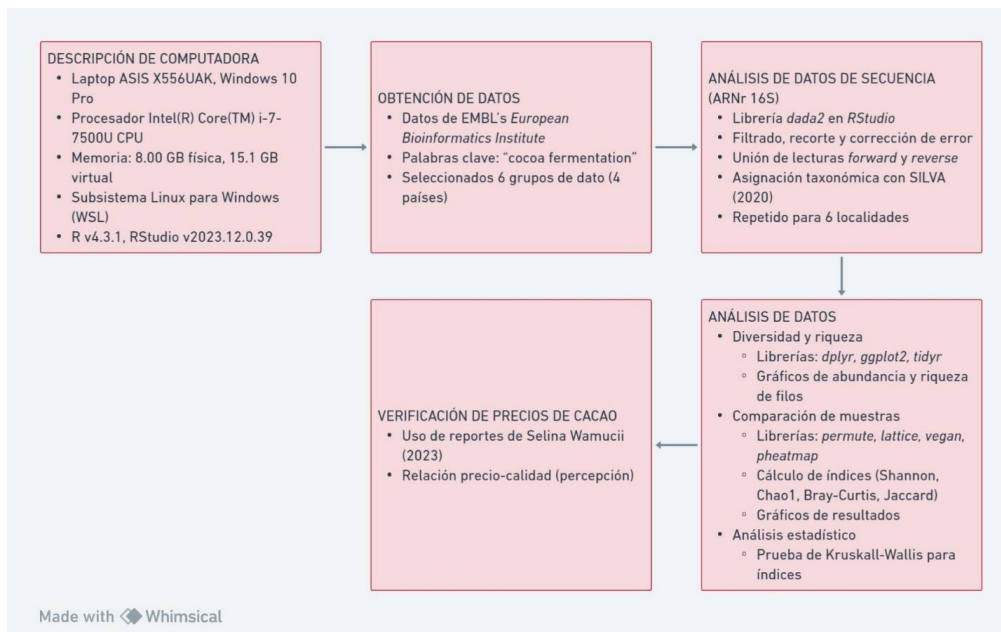


Figura 1: Algoritmo para análisis metagenómico del proceso fermentativo del cacao. (Elaboración propia)

VII. Resultados

i. Abundancia relativa

Con el fin de proporcionar una visión general de la composición taxonómica de cada sitio de estudio, se realizó un gráfico (**Figura 2**) que muestra la proporción de cada filo en relación con el total de filós identificados en cada sitio. Las barras representan la

abundancia relativa de cada filo.

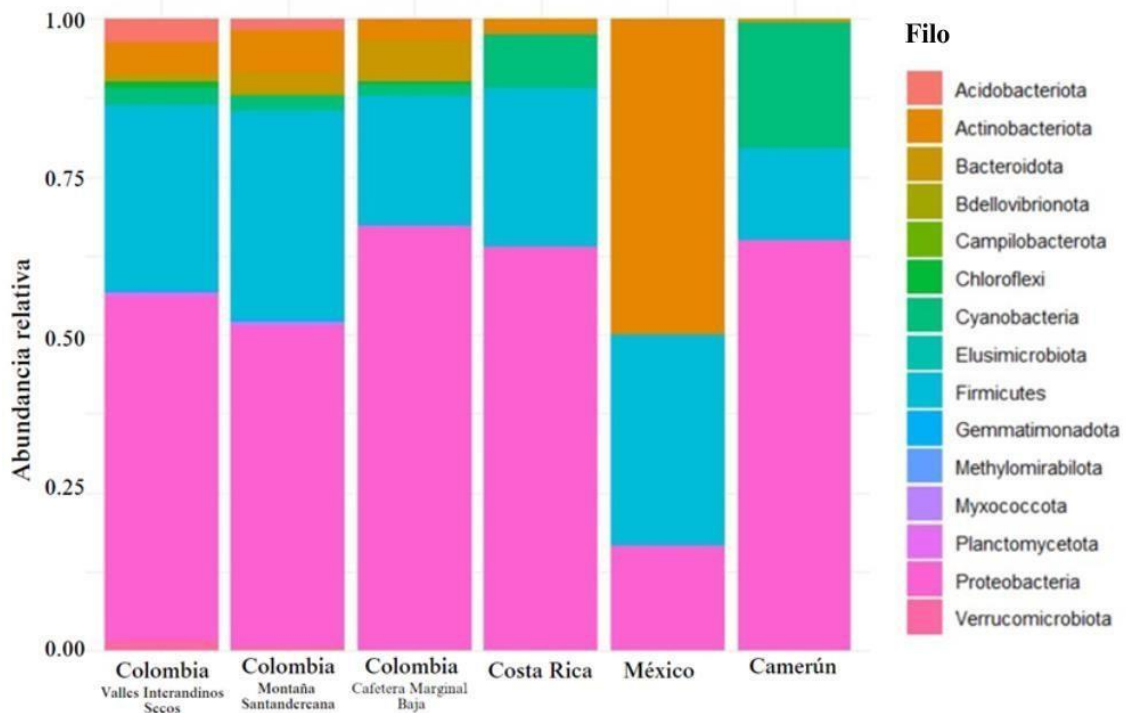


Figura 2: Abundancia relativa por sitio de estudio. (Elaboración propia)

Representación de un gráfico de barras, donde cada barra representa la abundancia relativa de especies en un rango de 0 a 1, para distintos sitios de estudio. Se coordina por colores, visualizada en la leyenda del lado derecho, facilitando la identificación rápida de las contribuciones relativas de cada sitio a la diversidad biológica. Cada color específico corresponde a un filo único.

Se puede observar que el filo predominante es Proteobacteria en todos los sitios, exceptuando México, en el que el filo predominante es Actinobacteriota. También se observa un gran porcentaje de Firmicutes en todos los sitios de estudio. Finalmente, se puede resaltar la alta presencia del filo Cyanobacteria en las muestras provenientes de Camerún.

ii. Riqueza

Con el objetivo de obtener información sobre la diversidad taxonómica de las muestras por cada sitio de estudio, se realizó un cuadro tabulando la riqueza de especies, es decir, la cantidad total de especies únicas identificadas en cada sitio.

Cuadro 3: *Riqueza de especies por sitio de estudio.*

Sitio	Riqueza de especies
Colombia (Valles Interandinos Secos)	225
Colombia (Montaña Santandereana)	604
Colombia (Cafetera Marginal Baja)	482
Costa Rica	536
México	52
Camerún	327

Nota: elaboración propia

Se puede observar que en las muestras provenientes de la Montaña Santandereana de Colombia se lograron identificar más de 600 especies, siendo la muestra con mayor riqueza de todo el estudio. Por otro lado, las muestras provenientes de México presentaron a penas 52 especies asignadas.

iii. Diversidad alfa

La diversidad alfa permite determinar la diversidad dentro de una comunidad o un hábitat en específico, siendo así una medida de variabilidad de especies dentro de un solo lugar. Es por esto que un alto valor de diversidad alfa es un indicador de que existe presencia de muchas especies distintas en un lugar en específico. El índice de Shannon permite medir la diversidad de especies teniendo en cuenta la riqueza y la uniformidad en la abundancia de las mismas. Un valor más alto es indicador de una mayor diversidad en términos de número de especies y equidad en la distribución de los mismos. Por otro lado, el índice de Chao1 es un estimador de riqueza total de filós, que ofrece información sobre la cantidad de filós presentes, incluidos aquellos que podrían no haber sido observados o detectados. Esto quiere decir que un valor más alto de este índice sugiere que la riqueza de especies real podría ser mayor a la que se pudo observar.

En el **Cuadro 4** se tabulan los sitios de estudio, la riqueza de filós y los índices de Shannon y Chao1. Se puede observar que las muestras provenientes de la Montaña Santandereana de Colombia, como las provenientes de Costa Rica presentaron los índices más altos.

Cuadro 4: Índices de diversidad alfa.

Sitio	Riqueza de filos	Índice de Shannon	Índice de Chao1
Colombia (Valles Interandinos Secos)	11	4.902	182
Colombia (Montaña Santandereana)	15	5.866	493
Colombia (Cafetera Marginal Baja)	10	5.502	369
Costa Rica	5	6.008	528
México	5	3.698	51
Camerún	5	4.869	151

Nota: elaboración propia

iv. Diversidad beta

La diversidad beta engloba la variabilidad de especies entre diferentes comunidades microbianas, evaluando cómo estas cambian o se diferencian entre lugares. En el caso de una matriz de Bray- Curtis, se observa la representación de las disimilitudes en la composición y abundancia relativa de elementos entre conjuntos, tomando en cuenta las diferencias en la abundancia de aquellos compartidos. Así, valores más bajos indican mayor similitud aparente en la composición de especies. Por otro lado, con la matriz de distancias de Jaccard se puede ilustrar la proporción de elementos que son comunes entre dos conjuntos, en relación con el total de elementos presentes en ambos. En este caso, los conjuntos idénticos estarían 0, mientras que los que no comparten ninguna especie estarían representados por un 1.

A continuación, se presenta la **Figura 3**, en la cual se puede observar la diversidad beta. En la **Figura 3A** se observa una matriz gráfica de los índices de Bray-Curtis, la cual ilustra las diferencias en la composición de especies entre los sitios mediante el Índice de Bray-Curtis, en el que los valores más bajos indican mayor similitud en la composición de especies. Por otro lado, la **Figura 3B** grafica un mapa de calor del índice de Jaccard. Este muestra la similitud entre pares de sitios, en el que valores más altos indican una mayor similitud en la composición de especies entre los sitios.

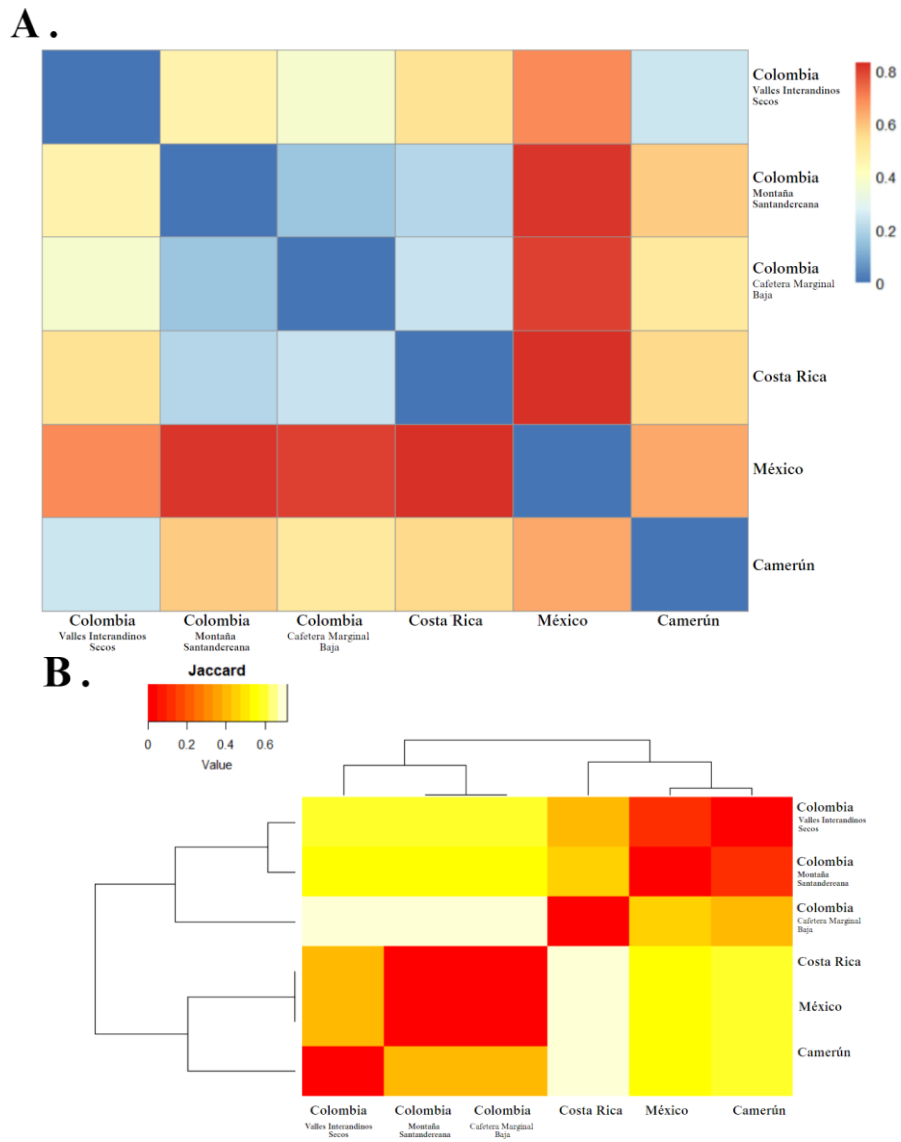


Figura 3: Gráficos de diversidad beta. (Elaboración propia)

En la **Figura 3A** se puede observar que las muestras 2 y 3 (Montaña Santandereana y Cafetera MarginalBaja, ambos en Colombia) presentaron conjuntos bastante similares en composición y abundancia, mientras que el sitio 5 (México) presentó la mayor disimilitud con los demás sitios, especialmente con las muestras provenientes de la Montaña Santandereana colombiana y aquellas de Costa Rica. Por otro lado, se pudo observar en la **Figura 3B** que las muestras provenientes de la Montaña Santandereana de Colombia son las que presentan mayor disimilitud aparente con aquellas provenientes de Costa Rica, México y Camerún.

v. Prueba estadística no paramétrica

El **Cuadro 5** tabula la comparación de diversidad alfa y beta en las muestras entre sitios. Se observan los índices Shannon, Chao1, Jaccard y Bray-Curtis, el estadístico de χ^2 obtenido, los grados de libertad y el valor p obtenido por la prueba de Kruskal-Wallis.

Cuadro 5: Comparación de diversidad alfa y beta en muestras de fermentación de cacao

	Índice	Estadístico χ^2	Grados de libertad	Valor p
Diversidad alfa	Shannon	5	5	0.4159
	Chao1	5	5	0.4159
Diversidad beta	Jaccard	9.0383	4	0.0602
	Bray-Curtis	2.6292	4	0.6217

No se observó ninguna diferencia estadísticamente significativa entre las muestras, ya que los valores p son altos (mayores a 0.05). El único índice que sugeriría una posible diferencia significativa sería el de Jaccard, ya que presentó un valor p muy cercano a 0.05 (0.0602).

vi. Flujo de trabajo – genoma completo

A continuación, en la **Figura 4** se presenta diagrama del algoritmo propuesto para análisis de microbiota con el uso de metagenoma completo.



Figura 4: Flujo de trabajo propuesto para análisis de microbiota del proceso fermentativo del cacao para datos de genoma completo.

VIII. Análisis de resultados

Con el objetivo de comparar los perfiles metagenómicos de la fermentación del cacao originarios de seis sitios de estudio distintos, obtenidos de servidores públicos, se realizó un análisis bioinformático. Para esto, se empleó una laptop marca ASUS con Windows 10 Pro y Linu a través de WSL. Los datos fueron obtenidos del EMBL's *European Bioinformatics Institute* e incluyeron seis grupos de datos de México, Costa Rica, Camerún y Colombia. Con ayuda de la librería *dada2* en *RStudio*, se realizó un análisis de secuencias de ARNr 16S, incluyendo asignación taxonómica. Posteriormente, se analizó la diversidad y riqueza utilizando paquetes como *dplyr* y *ggplo2*, y se realizaron comparaciones de muestras mediante índices de diversidad alfa y beta con pruebas estadísticas de Kruskal-Wallis. Finalmente, se verificaron los precios del cacao mediante reportes de Selina Wamucii para inferir el posible impacto de la microbiota en la calidad del chocolate producido.

Dentro del análisis de abundancia se determinó que el filo Proteobacteria tuvo gran presencia en todos los sitios de estudio, con excepción de México (Figura 1). Esto era de esperarse, ya que el filo Proteobacteria tiene una alta significancia ecológica en el proceso fermentativo del cacao (Illeghes, et al, 2012). El filo engloba un grupo muy diverso de bacterias, que presentan distintas capacidades metabólicas (Liu, et al, 2021) y, en el proceso de fermentación de cacao, miembros del filo Proteobacteria actúan como fermentadores primarios, lo que los lleva a jugar un papel indispensable en la descomposición de compuestos orgánicos complejos presentes en la pulpa del cacao (Illeghes, et al, 2012). Algunas bacterias del filo, como las bacterias ácido-acéticas (AAB, por sus siglas en inglés) dentro de la clase Alfaproteobacteria, contribuyen a la conversión de azúcares a ácidos orgánicos, como ácido acético (Gerard, 2016). Esta descomposición influye el pH y la composición química general de la pulpa (De Vuys y Weckx, 2015). La acidificación es indispensable para poder crear las condiciones necesarias para las actividades de otros microorganismos (De Vuys y Weckx, 2015). Los ácidos orgánicos producidos también contribuyen al desarrollo de los sabores característicos en los granos de cacao, lo que influye en las características organolépticas del producto final, que, este caso, es el chocolate (De Vuys y Weckx, 2015).

Asimismo, es importante mencionar que las bacterias pertenecientes al filo Proteobacteria muchas veces se encuentran dentro de los colonizadores tempranos del proceso de fermentación (Illeghes, et al, 2012). Sus actividades metabólicas crean un ambiente propenso al crecimiento de levaduras y bacterias ácido-lácticas, lo que lleva a una dinámica exitosa (Illeghes, et al, 2012). Además, estas bacterias se involucran en interacciones complejas con otros grupos microbianos, lo que influye en la estructura compleja de la comunidad (Illeghes, et al, 2012). Estas interacciones contribuyen a la diversidad y riqueza observada en el ecosistema (Illeghes, et al, 2012). Las proteobacterias

presentan versatilidad metabólica y son muy adaptables a las condiciones ambientales tan cambiantes (Rizzatti, *et al*, 2017). Esta adaptabilidad les permite sobrevivir en el ambiente fermentativo complejo, prosperando en diversas temperaturas, pHs y disponibilidad de nutrientes (Rizzatti, *et al*, 2017). Es posible que, por eso, el proceso fermentativo del cacao sea tan resiliente, influyendo en la consistencia y calidad del chocolate de diversas regiones geográficas.

En el caso de México, se observó una gran presencia del filo Actinobacteriota. Estas son conocidas por su habilidad de producir un amplio espectro de metabolitos secundarios, dentro de los que se encuentran compuestos antimicrobianos, los que podrían tener un potencial de biocontrol, influyendo en la abundancia y actividades de algunos patógenos (Silva, *et al*, 2022). Sirven, también, para la prevención de enfermedades, al limitar e impedir el crecimiento de microorganismos indeseables, lo que asegura la salud del cacao final (Silva, *et al*, 2022). Asimismo, esta diversidad metabólica podría jugar un papel importante en el moldeamiento de las interacciones microbianas y las dinámicas de la comunidad durante el proceso fermentativo (Viesser, *et al*, 2021). Asimismo, los miembros de este filo son descomponedores efectivos de sustratos orgánicos complejos, como celulosa y hemicelulosa, contribuyendo a la descomposición de compuestos presentes de la pared celular vegetal presentes en la pulpa de cacao (De Vuyst y Weckx, 2015). La descomposición de algunos sustratos resulta en la liberación de nutrientes y otros compuestos orgánicos más simples, contribuyendo a la disponibilidad de nutrientes para otros microorganismos involucrados en el proceso fermentativo (De Vuyst y Weckx, 2015).

Por otro lado, las actinobacterias pueden contribuir a la producción de metabolitos que influyen en las características organolépticas del cacao, especialmente en el sabor y aroma (De Vuyst y Weckx, 2015). Algunos compuestos volátiles producidos por este filo de bacterias pueden contribuir a las características sensoriales del chocolate final (De Vuyst y Weckx, 2015). Asimismo, las interacciones que estas bacterias tienen con levaduras y bacterias ácido-lácticas pueden influir la producción de compuestos aromáticos a través de comportamientos de cooperación (De Vuyst y Weckx, 2015). Es importante mencionar que las actinobacterias son conocidas por su capacidad de formar *biofilm*, lo que, se hipotetiza, tiene implicaciones en la adhesión microbiana y la colonización de los granos de cacao, contribuyendo a la estabilidad de comunidades microbianas y protegiendo a los microorganismos de estresores ambientales (Sauer, *et al*, 2022).

El filo Firmicutes también tuvo relevancia dentro de los seis sitios estudiados, estudios anteriores han demostrado que este forma parte grupos microbianos dominantes involucrados en el proceso fermentativo de cacao (Hamdouche, *et al*, 2019). Esta codominancia se puede atribuir a que la pulpa del cacao es una fuente bastante prominente de glucosa, sacarosa y glicerol, proveyendo un ambiente ideal para el crecimiento microbiano y actividades metabólicas (Kresnowati, *et al*, 2015). Además, los firmicutes también están involucrados en actividades metabólicas durante este proceso, lo que incluye la producción de ácido láctico y acético (Sorrenti, *et al*, 2020). Algunas bacterias de este filo contribuyen a la conversión de ácido cítrico a ácido láctico y a la reducción de fructosa a manitol (Zong, *et al*, 2023). La producción de ácido láctico contribuye a la acidificación de la pulpa del cacao, lo que es esencial para la creación de condiciones necesarias para las actividades microbianas (Kresnowati, *et al*, 2015).

A pesar de que no existen estudios que demuestren la importancia de las bacterias del género *Cyanobacteria* sp. en el proceso fermentativo, se pudo observar su abundancia en Camerún y un poco en Costa Rica. Como se mencionó anteriormente, el proceso de fermentación de cacao cuenta con una diversidad de microorganismos, por lo que no es extraño que se encuentren bacterias de este filo en las muestras analizadas. Estas son conocidas por su capacidad de fijación de nitrógeno y su producción de compuestos bioactivos (Herrero y Muro-Pastor, 2001), por lo que su presencia en la fermentación del cacao podría llegar a influir en la composición química del producto final y, por ende, de las características organolépticas.

En cuanto a la riqueza de especies (*Cuadro 2*), se observó que las muestras provenientes de la Montaña Santandereana de Colombia presentaron la mayor cantidad de especies, mientras que las de México fueron las que presentaron menos de cien especies asignadas. Esta riqueza microbiológica puede variar debido a diversos factores, como por ejemplo las condiciones climáticas, la calidad del suelo y la presencia de algunos microorganismos en el ambiente, así como los métodos de muestreo utilizados y la ubicación geográfica (López Suárez, *et al*, 2021). A pesar de que tanto en Colombia como en México se cultiva el cacao fino de aroma (Sandoval, *et al*, 2020), existen diferencias en cuanto a ecosistemas y climas predominantes en ambos países. En Colombia predomina el clima tropical y las regiones montañosas (World Bank Group, 2023), mientras que en México también se encuentran regiones desérticas y áreas costeras (Pacheco, 2022). Esta diversidad climática está influenciada por factores como la altitud, proximidad al ecuador y características topográficas (Pacheco, 2022). Es importante resaltar que un análisis de *16S* ignora por completo las poblaciones fúngicas. Estas tienen un rol crítico en la fermentación, ya que las levaduras producen etanol, que funciona como un agente bacteriostático (Schwan, *et al*, 2023). Debido a que no se contaban con datos del ARNr *18S*, se ve bastante limitada la población analizada. Por ejemplo, un estudio probó que en una muestra proveniente de Colombia se encontraron ocho especies distintas de levaduras (Tovar, *et al*, 2020). Además, se ha comprobado que las levaduras son indispensables precursores de sabor, además de iniciadores de cultivo (Gutiérrez-Ríos, *et al*, 2022). Esto se debe a que fermentan la glucosa de la pulpa del cacao a etanol y producen compuestos de sabores como aldehídos, ácidos orgánicos y ésteres, responsables de notas florales y frutales (Gutiérrez-Ríos, *et al*, 2022).

La diversidad alfa alta (*Figura 3*) observada especialmente en las muestras de Costa Rica y de la Montaña Santandereana de Colombia sugiere una rica comunidad microbiana en esas regiones, a pesar de no haber encontrado diferencias estadísticamente significativas ($p = 0.4159$). La diversidad en las poblaciones contribuye a las dinámicas de fermentación, en las que se hipotetiza que las especies interactúan para influir en el desarrollo del sabor, aroma y calidad de los granos de cacao fermentados, que llevan a un producto final ejemplar (Agyirifo, *et al*, 2019). Una gran diversidad alfa también podría significar que, dentro de la comunidad microbiana, hay especies con funciones metabólicas distintivas (Agyirifo, *et al*, 2019). Esta diversidad funcional podría potenciar la eficiencia del proceso fermentativo, lo que podría llevar a un producto final con características favorables (Agyirifo, *et al*, 2019). Dentro de esa misma línea, especies distintas de microorganismos podrían producir un sinnúmero de metabolitos durante la fermentación (Agyirifo, *et al*, 2019). Esto podría traducirse a la formación de compuestos únicos de sabor y olor, influyendo en la calidad final del

chocolate (Agyirifo, *et al*, 2019).

Se observó en la *Figura 3* que las muestras provenientes de la Montaña Santandereana y Cafetera Marginal Baja presentaron conjuntos bastante similares en composición y abundancia. Esto era de esperarse, debido a que ambos sitios de estudio se encuentran en Colombia; sin embargo, los sitios no tienen nada más en común. La Montaña Santandereana de Colombia está ubicada en el departamento de Santander, con altitudes muy variadas, desde 1,000 a 3,000 metros sobre el nivel del mar (Pacheco-Montealegre, *et al*, 2019). Santander se encuentra en la región central de Colombia y la variación contribuye a la presencia de distintos tipos de vegetación y muchos climas diversos, dentro de los que se pueden mencionar bosques de niebla y tropicales (Pacheco-Montealegre, *et al*, 2019). Por otro lado, la Cafetera Marginal Baja en Colombia se encuentra a altitudes más bajas, desde 600 a 1,200 metros sobre el nivel del mar, encontrándose en las tierras bajas o estribaciones de las cordilleras andinas colombianas (Pacheco-Montealegre, *et al*, 2019). El clima en esa región presenta tendencia a ser más cálido y la vegetación puede incluir bosques tropicales (Pacheco-Montealegre, *et al*, 2019). Una limitación grande para poder llegar a conclusiones es que no se tiene información sobre las técnicas de fermentación usadas, ya que esto podría haber jugado un papel importante en las características del proceso fermentativo.

Es importante mencionar que el presente estudio reveló una observación interesante: a pesar de que las muestras provenían de cuatro distintos países, no se presentaron diferencias estadísticamente significativas en las medidas de diversidad alfa y beta. Esto sugiere que, a un nivel general de diversidad y estructura de las comunidades microbianas, el proceso de fermentación de cacao muestra similitudes notables a lo largo de distintos países. La ausencia de diferencias estadísticamente significativas en los índices de diversidad alfa entre muestras indica un patrón consistente en la riqueza y distribución homogénea en las comunidades microbianas involucradas en el proceso fermentativo. Al ser una medida de diversidad dentro de una muestra específica, la diversidad alfa engloba factores como riqueza y diversidad de especies (Andermann, *et al*, 2022). El hecho que no se encontraran diferencias estadísticamente significativas podría sugerir un nivel de uniformidad en y distribución equitativa en estos aspectos a lo largo de los distintos orígenes de las muestras (Whittaker, *et al*, 2001).

Por otro lado, la falta de diferencias estadísticamente significativas ($p = 0.0602$ para el índice de Jaccard y $p = 0.6217$ para el índice de Bray-Curtis) en la diversidad beta refuerza la homogeneidad observada en la comunidad microbiana. Se observó que el índice de Jaccard estuvo muy cerca de ser considerado estadísticamente significativo, lo que podría plantear preguntas sobre la estructura específica de la comunidad. Esto, porque el índice de Jaccard mide la presencia o ausencia de especies compartidas entre muestras (Ederra y Miguel-Velasco, 1992). A pesar de que, en este caso, no se encontró estadísticamente significativo, el hecho de que se encuentre en el límite podría sugerir variaciones sutiles entre los miembros de la comunidad.

Se pueden mencionar implicaciones potenciales para la estructura comunitaria microbiana. El hecho de que el índice de Jaccard se encontrara cerca del límite de significancia estadística, podría ser indicador de variaciones sutiles en la identidad de las especies microbianas presentes en las distintas muestras del proceso fermentativo. A pesar que la diversidad y la riqueza generales siguen siendo bastante similares, las especies que constituyen las comunidades podrían ser ligeramente distintas. Esto se debe a que pueden ser influidas por factores del medio ambiente locales, características propias del grano de cacao o prácticas fermentativas de cada región. Aunque no se probó la significancia estadística en las variaciones, estas podrían tener implicaciones para los atributos sensoriales y de calidad de los productos finales del cacao, como el chocolate, debido a que las diferentes especies microbianas pueden aportar cualidades organolépticas distintas durante el proceso fermentativo.

Los perfiles metagenómicos de la fermentación de cacao pueden influir en la calidad de los granos de cacao, lo que puede afectar el precio de mercado. Una mayor diversidad microbiana o composiciones específicas de microorganismos durante el proceso podrían contribuir a una mejora en el desarrollo del sabor o calidad. Asimismo, algunos metabolitos microbianos producidos durante la fermentación podrían tener una relación directa con el impacto de las características organolépticas de los granos de cacao, lo que podría influir en el valor de mercado.

El entendimiento de algunas rutas metabólicas involucradas en el proceso de fermentación permitiría la identificación de contribuciones microbianas a los procesos bioquímicos clave. Algunos metabolitos, como ácidos orgánicos, alcoholes y otros compuestos volátiles podrían correlacionarse con características de aroma que podrían influir en los precios. Además, algunos taxones microbianos podrían estar asociados con la producción de compuestos que podrían darle al chocolate, como producto final, distintos sabores, aromas y textura (Hirko, *et al*, 2023). Compuestos como ácido acético, ácido butírico y algunos ésteres podrían darle al chocolate aromas frutales o a nueces, afectando las características organolépticas (Zareie, *et al*, 2023; Hirko, *et al*, 2023).

La producción de polifenoles y otros metabolitos secundarios producidos por ciertos géneros bacterianos podrían contribuir a la percepción de los beneficios a la salud y los atributos sensoriales de los productos derivados de cacao (Shahanas, *et al*, 2019; Sorrenti, *et al*, 2020). La relación entre la diversidad microbiana y la consistencia de las características organolépticas es crucial. El poder comprender cómo comunidades microbianas específicas contribuyen a perfiles de sabor consistentes podría permitir a los productores mantener la calidad de su producto, influyendo su posición en el mercado.

Como se puede apreciar en el *Cuadro 1*, el cacao mejor cotizado a nivel mundial según *Selina Wamucii* (2023) el de México, con aquel proveniente de Camerún en segundo lugar, bajo la premisa de que el precio es equivalente a la calidad del cacao producido. Relacionando esto con los resultados anteriores (**Figura 1**), en México se observó una gran cantidad de microorganismos pertenecientes a los filos Actinobacteriota y Firmicutes, con una pequeña abundancia de Proteobacteria. Por otro lado, en Camerún se observó una mayoría del filo Proteobacteria, con una presencia relativamente similar de Firmicutes y Cyanobacteria. Esto es consistente con los demás sitios, pero con abundancias relativas distintas. Sin embargo, es importante mencionar que las muestras de México fueron la que

presentaron la menor riqueza de especies por sitio de estudio (**Cuadro 2**), junto con las muestras de los Valles Interandinos Secos de Colombia y las de Camerún.

Por otro lado, la correlación entre el precio del cacao y los análisis metagenómicos realizados en el presente estudio, revelaron que el cacao de México es altamente valorado a nivel global. Este reconocimiento podría estar relacionado con la diversidad microbiana observada en el proceso fermentativo del cacao mexicano, en donde los filos predominantes Actinobacteriota y Firmicutes podrían desempeñar un papel importante en el desarrollo de las características organolépticas. Asimismo, en Camerún, la presencia alta de Proteobacteria sugiere otra faceta de la diversidad de microorganismos, que podría contribuir a otros perfiles distintivos de sabor. En cambio, con Colombia, se puede observar que los precios asignados a nivel internacional al cacao tienen un rango bastante grande (1.97 a 4.81 USD). Esto se puede deber a la diversidad genética presente en el país, en donde el cacao es endémico, así con la diversidad microbiana observada en los tres sitios de estudio, que presentaron la mayor cantidad de filos identificados. Finalmente, a pesar de la baja riqueza de especies en las muestras provenientes de México, comparada con los demás sitios de estudio, la calidad percibida del cacao podría ser atribuible a otros factores más específicos dentro de la composición de microorganismos en la comunidad. Estos resultados resaltan la interconexión compleja entre la diversidad microbiana, calidad de cacao y percepción de mercado.

Una de las limitaciones más resaltables del estudio es la ausencia de metadatos acompañando a las muestras microbianas. La indisponibilidad de información contextual detallada, como las condiciones ambientales de muestreo y las prácticas específicas de cultivo, así como las condiciones de fermentación, obstaculiza un poco la comprensión total de los factores que influyen la composición microbiana en el proceso de fermentación de cacao. Otra limitante resaltable es el hecho de que no se cuentan con datos de levaduras, ya que la fermentación de cacao es un sistema de fermentación conformado por levaduras y bacterias. Las levaduras son las que permiten la selección de poblaciones microbianas aptas para colonizar, ya que producen etanol que es un inhibidor de crecimiento de la mayoría de las bacterias (Koffi, *et al*, 2018). Es más, un estudio publicado por Mota-Gutierrez, *et al* (2019), menciona que algunas de las especies de levaduras más comúnmente presentes en el proceso fermentativo generan compuestos volátiles funcionales (VOCs, por sus siglas en inglés). Estos VOCs, al ser ingeridos, han sido demostrados de prevenir e inhibir el crecimiento de microorganismos e incluso tumores, así como de reducir la percepción del dolor, el estrés y la ansiedad, entre otros.

Asimismo, la omisión de datos fúngicos no permite determinar si las muestras analizadas padecieron de sobrefermentación, ya que una de las características es un exceso de etanol y la aparición de hongos filamentosos (de Araújo, *et al*, 2019; de Vuyst y Leroy, 2011). Otro de los retos del estudio estuvo relacionado con las diferencias en tamaños de muestras de los distintos sitios. La variabilidad en estos tamaños puede inducir sesgos en la representación de las comunidades microbianas. A pesar de que se realizaron esfuerzos para extraer perspectivas significativas de cada set de datos, la heterogeneidad en el tamaño de las muestras restringe la capacidad de generalizar los descubrimientos.

Por último, otra de las limitaciones nace de la ausencia de duplicados o triplicados en

las muestras dentro de cada sitio de estudio. La falta de réplicas restringe la habilidad de evaluar la reproducibilidad y consistencia de los patrones microbianos en la comunidad fermentativa. A pesar de las limitaciones, el presente estudio contribuyó con perspectivas a la diversidad microbiana de los sitios de estudio analizados. Reconocer estas limitaciones subraya la necesidad de interpretación cuidadosa de resultados y resalta áreas de mejora para futuras investigaciones en este campo de estudio, y enfatiza la necesidad de realizar más experimentos para la obtención de datos de este tipo.

IX. Conclusiones

1. Se generó una metodología para el análisis metagenómico de cacao, para poder determinar una relación entre los perfiles organolépticos del chocolate producido. Esto se logró con la correlación entre perfiles metataxonómicos y organolépticos, y los precios del cacao.
2. Se observó una presencia consistente del filo Proteobacteria en la fermentación del cacao de todos los sitios, exceptuando a México, lo que respalda la importancia ecológica de este filo en el proceso fermentativo.
3. México mostró una presencia relevante del filo Actinobacteriota, lo que podría señalar un potencial único en el perfil de aroma del cacao.
4. Se observó la presencia del filo Firmicutes en todos los sitios, sugiriendo un papel relevante en el proceso fermentativo.
5. No se encontraron diferencias estadísticamente significativas entre las diversidades alfa y beta, lo cual es un posible indicador de similitudes en la estructura de la comunidad microbiana de los distintos sitios estudiados, señalando una posible uniformidad en los procesos.
6. Se propuso un flujo de trabajo que permitiera realizar el mismo análisis, pero utilizando datos de metagenoma completo, en lugar de ASVs.

X. Recomendaciones

1. Para futuros estudios, se recomienda contar e incorporar metadatos robustos con el fin de poder realizar análisis más profundos e interpretables.
2. Se recomienda priorizar una recolección de muestras más uniforme, para poder asegurar una comparación balanceada y equitativa dentro de los distintos sitios de estudio.
3. Se recomienda incluir muestras en duplicado para poder proveer perspectivas en cuanto a la estabilidad y confiabilidad de las composiciones microbianas observadas, dándole más robustez a los análisis de diversidad microbiana.

XI. Referencias

- Afoakwa, E. O. (2010). *Chocolate science and technology* [Ciencia y tecnología del chocolate]. Wiley-Blackwell.
- Aguirre Sarceño, J.A. (2022). Análisis del Potencial de trayectorias de mejoramiento industrial para Guatemala dentro de la cadena global de cacao. [Tesis de licenciatura, Universidad del Valle de Guatemala].
- Agyirifo, D. S., Wamalwa, M., Otwe, E. P., Galyuon, I., Runo, S., Takrama, J., y Ngeranwa, J. (2019). Metagenomics analysis of cocoa bean fermentation microbiome identifying species diversity and putative functional capabilities. [Análisis metagenómico del microbioma de la fermentación del cacao en grano para identificar la diversidad de especies y las capacidades funcionales putativas] *Heliyon*, 5(7), e02170. <https://doi.org/10.1016/j.heliyon.2019.e02170>.
- Al-Mahrous, M. M., y Shah, A. H. (2016). Conserved bacterial and fungal genes: their importance in predicting pathogenicity and drug resistance. [Genes bacterianos y fúngicos conservados: su importancia en la predicción de la patogenicidad y la farmacorresistencia]. *Journal of microbiology and biotechnology research*, 1(1), 1-8. doi: <https://doi.org/10.15436/2377-0610.16.002>.
- Alonso, A. (2022) Análisis de comunidades: ANOSIM y NMDS. Rpubs, <https://rpubs.com/aafernandez1976/ANOSIMyNMDS>.
- Amann, R., Fuchs, B. M., y Behrens, S. (2001). The identification of microorganisms by fluorescence in situ hybridization. [Identificación de microorganismos mediante hibridación fluorescente in situ]. *Current Opinion in Biotechnology*, 12(3), 231-236. [https://doi.org/10.1016/S0958-1669\(00\)00208-6](https://doi.org/10.1016/S0958-1669(00)00208-6).
- Andermann, T., Antonelli, A., Barrett, R. L., et al. (2022). Estimating alpha, beta, and gamma diversity through deep learning [Estimación de la diversidad alfa, beta y gamma mediante aprendizaje profundo]. *Frontiers in plant science*, 13, 839407. <https://doi.org/10.3389/fpls.2022.839407>.
- Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online] [FastQC: Una herramienta de control de calidad para datos de secuencias de alto rendimiento [en línea]]. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Bajaj, V., Lucas, M., y Tammam, S. (2018). Challenges and strategies for culturing anaerobic bacteria [Retos y estrategias para el cultivo de bacterias anaerobias].

Journal of microbiology and biology education, 19(2), 1-8. doi: 10.1128/jmbe.v19i2.1434.

- Bamforth, C. W. (2019). Beer: Tap into the art and science of brewing (4th ed.) [Cerveza: Explore el arte y la ciencia de elaborar cerveza]. Oxford University Press.
- Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., ... y Segata, N.
- Beghini, F., McIver, F. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Maltez, A., Mireia Valles-Colomer, T., Weingart, G., Zhang, Y., Zolfo, M., Huttenhower, C., Franzosa, E. A. y Segata, N. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3 [Integración de perfiles taxonómicos, funcionales y a nivel de cepa de diversas comunidades microbianas con bioBakery 3]. *elife*, 10, e65088. <https://doi.org/10.7554/eLife.65088>.
- Benton, H. P., Wong, D. M., Trauger, S. A. y Siuzdak, G. (2008). XCMS2: processing tandem mass spectrometry data for metabolite identification and structural characterization [XCMS2: procesamiento de datos de espectrometría de masas en tándem para la identificación de metabolitos y la caracterización estructural]. *Analytical chemistry*, 80(16), 6382-6389. <https://doi.org/10.1021/ac800795f>.
- Bermudez, S., Voora, V., Larrea, C. y Luna, E. (2022). Cocoa prices and sustainability - Global market report [Precios del cacao y sostenibilidad - Informe sobre el mercado mundial]. <https://www.iisd.org/publications/report/2022-global-market-report-cocoa>.
- Bolger, A. M., Lohse, M. y Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data [Trimmomatic: Un recortador flexible para datos de secuencias de Illumina]. *Bioinformatics*, btu170. <https://doi.org/10.1093/bioinformatics/btu170>.
- Broadbent, J. R., Neeno-Eckwall, E., Stahl, B. y Tandee, K. (2018). Fermented foods: Definitions and characteristics, impact on the microbiota and potential health benefits [Alimentos fermentados: Definiciones y características, impacto en la microbiota y posibles beneficios para la salud]. *Microbial ecology of fermented foods* (pp. 1-24). Springer. <https://doi.org/10.3390/nu11081806>.
- Brooks, J. P., Edwards, D. J. y Harwood, V. J. (2014). Failure to culture select fecal indicator bacteria under laboratory conditions simulating environmental stress [Fracaso en el cultivo de bacterias indicadoras fecales seleccionadas en condiciones de laboratorio que simulan el estrés ambiental.]. *Applied and environmental microbiology*, 80(21), 6659-6667. <https://doi.org/10.1128/AEM.01992-14>.
- Buchfink, B., Xie, C. y Huson, D.H. (2014). Fast and sensitive protein alignment using DIAMOND [Alineación rápida y sensible de proteínas mediante DIAMOND]. *Nature Methods*, 12: 59-60. <https://doi.org/10.1038/nmeth.3176>.

- Buck, U., Götz, M., Romanowski, G., & Mühlenweg, A. (2015). In silico identification and in vivo verification of intergenic regions conserved between *Bacillus subtilis* and *Bacillus anthracis* [Identificación in silico y verificación in vivo de regiones intergénicas conservadas entre *Bacillus subtilis* y *Bacillus anthracis*]. *Applied and environmental microbiology*, 81(16), 5485-5496. <https://doi.org/10.1128/AEM.01185-15>.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A. y Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data [DADA2: Inferencia de muestras de alta resolución a partir de datos de amplicones de Illumina]. *Nature Methods*, 13(7), 581-583. <https://doi.org/10.1038/nmeth.3869>.
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., y Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms [Análisis de comunidades microbianas de alto rendimiento en las plataformas Illumina HiSeq y MiSeq]. *The ISME Journal*, 6(8), 1621-1624. <https://doi.org/10.1038/ismej.2012.8>.
- Caporaso, J.G., Lauber, C.L., Walters, W.A., *et al.* (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample [Patrones globales de diversidad del ARNr 16S a una profundidad de millones de secuencias por muestra]. *Proc Natl Acad Sci USA*, 108, Suppl 1:4516-22. <https://doi.org/10.1073/pnas.1000080107>.
- Cherif, S., *et al.* (2019). Effects of Fermentation Process on the Nutritional and Functional Properties of Sauerkraut: A Review [Efectos del proceso de fermentación en las propiedades nutricionales y funcionales del chucrut: Una revisión]. *Food Science of Animal Resources*, 39(6), 751-763. <https://doi.org/10.5851/kosfa.2019.e70>.
- Church, D. L., Cerutti, L., Gürtler, A., Griener, T., Zelazny, A., & Emler, S. (2020). Performance and application of 16S rRNA gene cycle sequencing for routine identification of bacteria in the clinical microbiology laboratory [Rendimiento y aplicación de la secuenciación cíclica del gen ARNr 16S para la identificación rutinaria de bacterias en el laboratorio de microbiología clínica]. *Clinical microbiology reviews*, 33(4), e00053-19. <https://doi.org/10.1128/cmr.00053-19>.
- Creek, D. J., Dunn, W. B., Fiehn, O., Griffin, J. L., Hall, R. D., Lei, Z., Mistrik, R., Neumann, S., Schymansky, E. L., Sumner, L. W., Trengove, R. y Wolfender, J. L. (2014). Metabolite identification: are you sure? And how do your peers gauge your confidence? [Identificación de metabolitos: ¿está seguro? ¿Y cómo calibran tus compañeros tu confianza?]. *Metabolomics*, 10(3), 350-353. <https://doi.org/10.1007/s11306-014-0633-x>.
- Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L. y Toohey- Kurth, K. (2020). Guidelines for Sanger sequencing and molecular assay monitoring [Directrices para la secuenciación Sanger y el seguimiento de ensayos moleculares]. *Journal of Veterinary Diagnostic Investigation*, 32(6), 767-775.

<https://doi.org/10.1177/1040638720905833>.

- De Filippis, F., Troise, A. D., Vitaglione, P. y Ercolini, D. (2018). Different temperatures select distinctive acetic acid bacteria species and promotes organic acids production during Kombucha tea fermentation [Diferentes temperaturas seleccionan especies distintivas de bacterias del ácido acético y promueven la producción de ácidos orgánicos durante la fermentación del té Kombucha]. *Food Microbiology*, 73, 11-16. <https://doi.org/10.1016/j.fm.2018.01.008>.
- De Vuyst, L. y Wecks, S. (2016). The cocoa bean fermentation process: from ecosystem analysis to starter culture development [El proceso de fermentación del cacao en grano: del análisis del ecosistema al desarrollo de cultivos iniciadores]. *Journal of Applied Microbiology*, 121(1). <https://doi.org/10.1111/jam.13045>.
- De Vuyst, L. y Weckx, S. (2015). The functional role of lactic acid bacteria in cocoa bean fermentation [Papel funcional de las bacterias lácticas en la fermentación del cacao en grano]. *Biotechnology of lactic acid bacteria: novel applications*, 248-278. <https://doi.org/10.1002/9781118868386.ch16>.
- Douglas, G.M., Maffei, V.J., Zaneveld, J.R., Yurgel, S.N., Brown, J.R., Taylor, C.M., Huttenhower, C. y Langille, M.G.I. (2020). PICRUSt2 for prediction of metagenome functions [PICRUSt2 para la predicción de las funciones del metagenoma]. *Nature Biotechnology*, 38. <https://doi.org/10.1038/s41587-020-0548-6>.
- Durazzi, F., Sala, C., Castellani, G., Manfreda, G., Remondini, D. y De Cesare, A. (2021). Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota [Comparación entre datos de ARNr 16S y secuenciación shotgun para la caracterización taxonómica de la microbiota intestinal]. *Scientific reports*, 11(1), 1-10. <https://doi.org/10.1038/s41598-021-82726-y>.
- Ederra, A. y Miguel-Velasco, A. M. (1992). Estudio de la afinidad entre hayedos y robledales de Navarra (España) mediante la aplicación del índice de Jaccard a su brioflora. *Actes del Simposi Internacional de Botanica Pius Font i Quer*, Vol 1.
- Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J. y Otto, G. (2009) Real-Time DNA Sequencing from Single Polymerase Molecules [Secuenciación del ADN en tiempo real a partir de moléculas de polimerasa única]. *Science*, 323(5910), 133- 138. <https://doi.org/10.1126/science.1162986>
- Eskes, A. B., Lass, R. A., Phillips-Mora, W. y Astorga, C. (2010). Factors influencing seed quality and cacao production through genetic improvement [Factores que influyen en la calidad de las semillas y la producción de cacao mediante la mejora genética]. In J. Flood (Ed.), *Achieving sustainable cultivation of cocoa* (pp. 197-226). Burleigh Dodds Science Publishing. <https://doi.org/10.1016/j.foodres.2016.01.012>.
- Forsythe, S. J. (2018). *The microbiology of safe food* (2nd ed.) [Microbiología de los

- alimentos seguros]. Wiley-Blackwell. <https://doi.org/10.1002/9780470999431>.
- Fox, P.F. y McSweeney, P.L.H. (Eds.). (2017). *Cheese: Chemistry, Physics, and Microbiology* (4th ed.) [Queso: Química, Física y Microbiología]. Academic Press.
- Gardes, M. y Bruns, T. D. (1993). *ITS primers with enhanced specificity for basidiomycetes—application to the identification of mycorrhizae and rusts* [Cebadores *ITS* con especificidad mejorada para basidiomicetos - aplicación a la identificación de micorrizas y royas]. *Molecular ecology*, 2(2), 113-118. <https://doi.org/10.1111/j.1365-294x.1993.tb00005.x>.
- Gaudreau, C., Gagné, M. J. y Couture, F. (2014). Comparison of culture methods for detection of *Salmonella enterica* serotype Enteritidis in poultry environmental samples [Comparación de métodos de cultivo para la detección de *Salmonella enterica* serotipo Enteritidis en muestras ambientales de aves de corral]. *Journal of food protection*, 77(5), 742-748. <https://doi.org/10.4315/0362-028X.JFP-11-297>.
- Gerard, L. M. (2016). Caracterización de bacterias del ácido acético destinadas a la producción de vinagres de frutas [Tesis para doctorado], Universitat Politècnica de València.
- Greenwalt, C.J., Steinkraus, K.H. y Ledford, R.A. (2000). Kombucha, the fermented tea: microbiology, composition, and claimed health effects [Kombucha, el té fermentado: Microbiología, composición y supuestos efectos saludables]. *Journal of Food Protection*, 63(7), 976-981. <https://doi.org/10.4315/0362-028X-63.7.976>.
- Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L. y Goodacre, R. (2015). A tutorial review: metabolomics and partial least squares- discriminant analysis—a marriage of convenience or a shotgun wedding [Revisión de un tutorial: metabolómica y análisis discriminante por mínimos cuadrados parciales: ¿un matrimonio de conveniencia o una boda escopeta?]. *Analytica chimica acta*, 879, 10-23. <https://doi.org/10.1016/j.aca.2015.02.012>.
- Guan, Q., Zhang, B. y Yan, G. (2017). Metabolomic profiling reveals correlation between corticosterone and oxidative stress in the hippocampus of adrenalectomized rats [Los perfiles metabolómicos revelan una correlación entre la corticosterona y el estrés oxidativo en el hipocampo de ratas adrenalectomizadas.]. *Neurochemical research*, 42(2), 658-666.
- Hadley Wickham (2016). *ggplot2: Elegant Graphics for Data Analysis* [*ggplot2: Gráficos Elegantes para el Análisis de Datos.*]. Springer, New York. ISBN 978-3- 319-24277-4, <https://ggplot2.tidyverse.org>
- Halász, A., Baráth, A. y Holzapfel, W. H. (1999). The influence of starter culture selection on sauerkraut fermentation. *Z Lebensm Unters Forsch*, 208. <https://doi.org/10.1007/s002170050443>.
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT [BioEdit: un editor de alineación de

secuencias biológicas y programa de análisis de fácil uso para Windows 95/98/NT]. *Nucleic acids symposium series*, 41, 95-98.

- Hamdouche, Y., Meile, J. C., Lebrun, M., Guehi, T., Boulanger, R., Teyssier, C. y Montet, D. (2019). Impact of turning, pod storage and fermentation time on microbial ecology and volatile composition of cocoa beans [Efectos del volteo, el almacenamiento de las vainas y el tiempo de fermentación en la ecología microbiana y la composición volátil del cacao en grano]. *Food Research International*, 119, 477-491. <https://doi.org/10.1016/j.foodres.2019.01.001>.
- Herrero, A., Muro-Pastor, A. M. y Flores, E. (2001). Nitrogen control in cyanobacteria [Control del nitrógeno en las cianobacterias]. *Journal of bacteriology*, 183(2), 411-425. <https://doi.org/10.1128/jb.183.2.411-425.2001>.
- Hirko, B., Mitiku, H. y Getu, A. (2023). Role of fermentation and microbes in cacao fermentation and their impact on cacao quality [Papel de la fermentación y los microbios en la fermentación del cacao y su impacto en la calidad del cacao]. *Systems Microbiology and Biomanufacturing*, 1-12. <https://doi.org/10.1007/s43393-023-00160-9>.
- Hughey, V. L. y Johnson, M. G. (2019). A review of the safety of probiotics and functional foods in the United States: what do we know about their safety? [Revisión de la seguridad de los probióticos y los alimentos funcionales en Estados Unidos: ¿qué sabemos sobre su seguridad?]. *Nutrients*, 11(2), 261. <https://doi.org/10.3390/nu11020261>.
- Hui, Y. H. y Özgül Evranuz, E. (Eds.) (2018). *Handbook of fermented food and beverage technology (2nd ed.) [Manual de tecnología de alimentos y bebidas fermentados]*. CRC Press.
- Hutkins, R. W. (2018). *Microbiology and technology of fermented foods [Microbiología y tecnología de los alimentos fermentados]*. JohnWiley & Sons.
- Illegheems, K., De Vuyst, L., Papalexandratou, Z. y Veckx, S. (2012). Phylogenetic Analysis of a Spontaneous Cocoa Bean Fermentation Metagenome Reveals New Insights into Its Bacterial and Fungal Community Diversity [El análisis filogenético de un metagenoma espontáneo de fermentación de habas de cacao revela nuevos conocimientos sobre la diversidad de su comunidad bacteriana y fúngica]. *PLoS ONE*, 7(5). <https://doi.org/10.1111/jam.13045>.
- Illegheems, K., Wecks, S. y De Vuyst, L. (2015). Applying meta-pathway analysis through metagenomics to identify the functional properties of the major bacterial communities of a single spontaneous cocoa bean fermentation process sample [Aplicación del análisis de meta-vías mediante metagenómica para identificar las propiedades funcionales de las principales comunidades bacterianas de una única muestra espontánea del proceso de fermentación del cacao en grano]. *Food Microbiology*, 50. <https://doi.org/10.1016/j.fm.2015.03.005>.
- International Cocoa Organization. (2022). November 2022 quarterly bulletin of cocoa

statistics [Boletín trimestral de estadísticas del cacao de noviembre de 2022]. <https://www.icco.org/november-2022-quarterly-bulletin-of-cocoa-statistics/>.

James, C. E. y Stanley, K. N. (2013). How to culture fastidious anaerobes [Cómo cultivar anaerobios fastidiosos]. *Journal of microbiology and biotechnology*, 23(6), 797-809. <https://doi.org/10.4014/jmb.1302.02065>.

Jayabalan, R., Malbaša, R.V., Lončar, E.S., Vitas, J.S. y Sathishkumar, M. (2014). A Review on Kombucha Tea—Microbiology, Composition, Fermentation, Beneficial Effects, Toxicity, and Tea Fungus [Revisión de la microbiología, composición, fermentación, efectos beneficiosos, toxicidad y hongos del té Kombucha]. *Comprehensive Reviews in Food Science and Food Safety*, 13(4), 538- 550. <https://doi.org/10.1111/1541-4337.12073>.

Jeong, S. y Chun, J. (2007). Kimchi and Other Widely Consumed Traditional Fermented Foods of Korea: A Review [Kimchi y otros alimentos fermentados tradicionales de Corea: Una revisión]. *Frontiers in Microbiology*, 8, 1-11. <https://doi.org/10.3389/fmicb.2017.00786>.

Johnson, J. S., Spakowicz, D. J., Hong, B. Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E. y Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis [Evaluación de la secuenciación del gen ARNr 16S para el análisis del microbioma a nivel de especie y de cepa]. *Nature communications*, 10(1), 5029. <https://doi.org/10.1038/s41467-019-13036-1>.

Jung, J., Lee, S., Kim, J. Park, M. S., Bae, J. W., Hahn, Y., Madsen, E. L. y Jeon, C. O. (2011). Metagenomic analysis of kimchi, a Traditional Korean fermented food [Análisis metagenómico del kimchi, un alimento fermentado tradicional coreano]. *Applied and Environmental Microbiology*, 77(7). <https://doi.org/10.1128/AEM.02157-10>.

Kanehisa, M., Sato, Y. y Kawashima, M. (2022). KEGG mapping tools for uncovering hidden features in biological data [Herramientas de cartografía KEGG para descubrir características ocultas en datos biológicos]. *Protein Science*, 31(1), 47-53. <https://doi.org/10.1002/pro.4172>.

Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. y Tanabe, M. (2019). New approach for understanding genome variations in KEGG [Nuevo enfoque para comprender las variaciones del genoma en KEGG]. *Nucleic Acids Research*, 47(D1), D590-D595. <https://doi.org/10.1093/nar/gky962>.

Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities [MetaBAT, una herramienta eficaz para reconstruir con precisión genomas individuales de comunidades microbianas complejas]. *PeerJ*, 3, e1165. <https://doi.org/10.7717/peerj.1165>.

Kim, M. y Chun, J. (2005). Bacterial community structure in kimchi, a Korean 40

fermented vegetable food, as revealed by 16S rRNA gene analysis. *International Journal of Food Microbiology*, 103. <https://doi.org/10.1016/j.ijfoodmicro.2004.11.030>.

- Kim, K. H., Lee, J. H., Kim, T. S., Lee, H. J., Kim, S. y Lee, S. (2016). Metabolomics analysis for the identification of differential metabolites in Alzheimer's disease [Análisis metabolómico para la identificación de metabolitos diferenciales en la enfermedad de Alzheimer]. *Molecular bioSystems*, 12(8), 2290-2299. <https://doi.org/10.1039/C6MB00202C>.
- Knight, R., Jansson, J., Field, D., Fierer, N., Desai, N., Fuhrman, J. A., Hugenholtz, P., van der Lelie, D., Meyer, F., Stevens, R., Bailey, M. J., Gordon, J. I., Kowalchuk, G. A. y Gilbert, J. A. (2012) Unlocking the potential of metagenomics through replicated experimental design [Liberar el potencial de la metagenómica mediante un diseño experimental replicado]. *Nat Biotechnol*, 30(6): 513-20.
- Koffi, O., Samagaci, L., Goualie, B. y Niamke, S. (2018). Screening of potential yeast starters with high ethanol production for a small-scale cocoa fermentation in Ivory Coast [Selección de posibles levaduras iniciadoras con alta producción de etanol para una fermentación de cacao a pequeña escala en Costa de Marfil]. *Food and Environment Safety Journal*, 17(2).
- Kolde R. (2019). *pheatmap: Pretty Heatmaps* [pheatmap: Mapas de calor bonitos]. R package version 1.0.12, <https://CRAN.R-project.org/package=pheatmap>.
- Kresnowati, M.T.A.P., Gunawan, A.Y. y Muliadini, W. (2015). Kinetics model development of cocoa bean fermentation [Desarrollo de un modelo cinético de fermentación del cacao en grano]. <https://www.science.gov/topicpages/c/cocoa+bean+fermentation.html>
- Lacroix-Gueu, P., Briandet, R., y Leveque-Fort, S. (2015). Exploring bacterial biofilm diversity in water networks with different disinfection treatments [Exploración de la diversidad de biopelículas bacterianas en redes de agua con diferentes tratamientos de desinfección]. *Water research*, 68, 249-260. <https://doi.org/10.1016/j.watres.2014.10.051>.
- Langille, M. G.I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G. y Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences [Perfiles funcionales predictivos de comunidades microbianas mediante secuencias de genes marcadores del ARNr 16S]. *Nature Biotechnology*, 8, 1-10. <https://doi.org/10.1038/nbt.2676>.
- Lenz, E. M., Bright, J., Wilson, I. D., Hughes, A., Morrisson, J. y Lindberg, H. (2004). Metabonomics, dietary influences and cultural differences: a 1H NMR-based study of urine samples obtained from healthy British and Swedish subjects [Metabonomía, influencias dietéticas y diferencias culturales: un estudio basado en 1H RMN de muestras de orina obtenidas de sujetos sanos británicos y suecos]. *Journal of pharmaceutical and biomedical analysis*, 36(5), 841-849.

<https://doi.org/10.1016/j.jpba.2004.08.013>.

- Lewis, M. J. y Young, T. W. (2002). *Brewing* (2nd ed.) [Elaboración de cerveza]. Springer Science & Business Media.
- Li, D., Liu, C.M., Luo, R., Sadakane, K. y Lam, T.W. (2015). MEGAHIT: an ultra- fast single-node solution for large and complex metagenomics assembly via succinctde Bruijn graph [MEGAHIT: una solución ultrarrápida de un solo nodo para el ensamblaje de metagenómica grande y compleja mediante el sucinto grafo de Bruijn]. *Bioinformatics*, 31(10): 1674-6. <https://doi.org/10.1093/bioinformatics/btv033>.
- Li, L., Stoeckert Jr, C. J. y Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes [OrthoMCL: identificación de grupos ortólogos para genomas eucarióticos]. *Genome research*, 13(9), 2178-2189. <https://doi.org/10.1101/gr.1224503>.
- Liu, J., Zhou, H., Yang, Z., Wang, X., Chen, H., Zhong, L. & Bian, X. (2021). Rational construction of genome-reduced Burkholderiales chassis facilitates efficient heterologous production of natural products from proteobacteria [La construcción racional de chasis de Burkholderiales de genoma reducido facilita la producción heteróloga eficiente de productos naturales a partir de proteobacterias]. *Nature Communications*, 12(1), 4347. <https://doi.org/10.1038/s41467-021-24645-0>.
- López Suárez, A. F., González Barrios, A., Herrera Rocha, F.E. y Reyes Muñoz, (2021). Análisis metataxonómico de la fermentación espontánea de cacao de finoaroma en dos granjas industriales, Colombia. <http://hdl.handle.net/1992/51669>.
- Lozupone, C. A., & Knight, R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities [UniFrac: Un nuevo método filogenético para comparar comunidades microbianas]. *Applied and Environmental Microbiology*, 71(12), 8228-8235. <https://doi.org/10.1128/AEM.71.12.8228-8235.2005>.
- MAGA. (2014). *Perfil comercial cacao*. Gobierno de Guatemala.
- Magurran, A. E. (2004). *Measuring Biological Diversity* [Medición de la diversidad biológica]. Blackwell Publishing, Oxford.
- Marco, M. L. y Heeney, D. (2017). Bifidobacteria and their role in fermented foods. In *Lactobacillus and bifidobacterium* (pp. 297-318) [Las bifidobacterias y su papel en los alimentos fermentados. En *Lactobacillus y bifidobacterium*]. Springer, Cham.
- Marco, M. L. y Heeney, D. (Eds.). (2019). *Fermented foods in health and disease prevention* [Los alimentos fermentados en la salud y la prevención de enfermedades]. Academic Press.
- Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics [El impacto de la tecnología de secuenciación de nueva generación en la genética]. *Trends in Genetics*, 24(3), 133-141.

<https://doi.org/10.1016/j.tig.2007.12.007>.

- McGinnis, S. y Madden, T.L. (2004). BLAST: at the core of a powerful and diverse set of sequence analysis tools [BLAST: en el núcleo de un potente y variado conjunto de herramientas de análisis de secuencias]. *Nucleic Acids Research*, 32. <https://doi.org/10.1093/nar/gkh435>.
- McSweeney, P.L.H., & Sousa, M.J. (2000). Biochemical pathways for the production of flavor compounds in cheeses during ripening: A review [Vías bioquímicas para la producción de compuestos de sabor en quesos durante la maduración: Una revisión]. *Lait*, 80(3), 293-324. <https://doi.org/10.1051/lait:2000127>.
- Meersman, E., Steensels, J., Mathawan, M., Wittocs, P. J., Saels, V., Struyf, N., Bernaert, H., rancken, G. y Verstrepen, K. J. (2013). Detailed analysis of the microbial population in Malaysian spontaneous cocoa pulp fermentations reveals a core and variable microbiota [El análisis detallado de la población microbiana en las fermentaciones espontáneas de pulpa de cacao de Malasia revela una microbiota central y variable]. *PLoS ONE*, 8(12). <https://doi.org/10.1371/journal.pone.0081559>.
- Metzker, M. L. (2010). Sequencing technologies—the next generation [Tecnologías de secuenciación: la próxima generación]. *Nature Reviews Genetics*, 11(1), 31-46. <https://doi.org/10.1038/nrg2626>.
- Ministerio de Agricultura, Ganadería y Alimentación de Guatemala (2014). Perfil comercial cacao.
- Mota-Gutierrez, J., Barbosa-Pereira, L., Ferrocino, I. y Cocolin, L. (2019). Traceability of Functional Volatile Compounds Generated on Inoculated Cocoa Fermentation and Its Potential Health Benefits [Trazabilidad de los compuestos volátiles funcionales generados en la fermentación inoculada del cacao y sus posibles beneficios para la salud]. *Nutrients*, 11(884). <https://doi.org/10.3390/nu11040884>.
- Mota-Gutierrez, J., Botta, C., Ferrocino, I. et al. (2018). Dynamics and Biodiversity of Bacterial and Yeast Communities during Fermentation of Cocoa Beans [Dinámica y Biodiversidad de las Comunidades de Bacterias y Levaduras durante la Fermentación del Cacao en Grano]. *Applied and Environmental Microbiology*, 84(19). <https://doi.org/10.1128/AEM.01164-18>.
- Muhamad Rizal, N. S., Neoh, H. M., Ramli, R., A, et al. (2020). Advantages and limitations of 16S rRNA next-generation sequencing for pathogen identification in the diagnostic microbiology laboratory: perspectives from a middle-income country [Ventajas y limitaciones de la secuenciación de próxima generación del ARNr 16S para la identificación de patógenos en el laboratorio de microbiología diagnóstica: perspectivas de un país de renta media]. *Diagnostics*, 10(10), 816. <https://doi.org/10.3390/diagnostics10100816>.
- National Center for Biotechnology Information (NCBI) [Internet] (1988). Bethesda (MD):

National Library of Medicine (US) [Bethesda (MD): Biblioteca Nacional de Medicina (EE.UU.)], National Center for Biotechnology Information. <https://www.ncbi.nlm.nih.gov/>.

- Nielsen, S. S. (2019). Chocolate and cocoa beans: Overview of their history, health benefits, and processing [Chocolate y cacao en grano: Reseña de su historia, beneficios para la salud y elaboración]. *Food Analysis* (5th ed., pp. 647-662). Springer.
- Noga, M. J., Dane, A., Shi, S., Attali, A., & van Aken, H. (2016). Metabolomics for plant stress response [Metabólica para la respuesta de las plantas al estrés]. *Physiologia plantarum*, 157(3), 295-306. <https://doi.org/10.1111/j.1399-3054.2007.01025.x>.
- Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., O'Hara, R.B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., ... Weedon, J. (2022). *vegan: Community Ecology Package (vegan: paquete de ecología comunitaria)*. R package version 2.6-4, <https://CRAN.R-project.org/package=vegan>.
- Ozturkoglu-Budak, S. y Akman, P. K. (2020). Food fermentation: A tool for improving the nutritional quality of foods [Fermentación de alimentos: Una herramienta para mejorar la calidad nutricional de los alimentos]. *Critical Reviews in Food Science and Nutrition*, 60(16), 2726-2743. <https://doi.org/10.1080/10408398.2019.1638127>
- Pacheco, K. M. (2022). Estudios sobre historia y clima. Argentina, Colombia, Chile, España, Guatemala, México y Venezuela. Secuencia. *Revista de Historia y Ciencias Sociales*.
- Pacheco-Montealegre, M.E., Dávila-Mora, L.L., Botero-Rute, L.M., Reyes, A. y Caro-Quintero, A. (2019). Fine grained longitudinal analysis of cocoa bean fermentation provides insights into the dynamics of microbial populations [El análisis longitudinal detallado de la fermentación del cacao en grano permite comprender la dinámica de las poblaciones microbianas]. *Microbiome dynamics during cocoa bean fermentation*. : <https://doi.org/10.1101/702225>.
- Papalexandratou, Z., Vrancken, G., De Bruyne, K., Vandamme, P. y De Vuyst, L. (2011). Spontaneous organic cocoa bean box fermentations in Brazil are characterized by arestricted species diversity of lactic acid bacteria and acetic acid bacteria [Las fermentaciones orgánicas espontáneas de cajas de cacao en grano en Brasil se caracterizan por una diversidad restringida de especies de bacterias lácticas y acéticas]. *Food Microbiology*, 28(7), 1326-1338. <https://doi.org/10.1016/j.fm.2011.06.003>.
- Park, K. Y., Jeong, J. K., Lee, Y. E. y Daily III, J. W. (2014). Health benefits of kimchi (Korean fermented vegetables) as a probiotic food [Beneficios para la salud del kimchi (verduras fermentadas coreanas) como alimento probiótico]. *Journal of medicinal food*, 17(1), 6-20. <https://doi.org/10.1089/jmf.2013.3083>.

- Payne, M. J., Hurst, W. J., Miller, K. B., Rank, C. y Stuart, D. A. (2010). Impact of fermentation, drying, roasting, and Dutch processing on epicatechin and catechin content of cacao beans and cocoa ingredients [Efectos de la fermentación, el secado, el tostado y el procesamiento holandés en el contenido de epicatequina y catequina de los granos de cacao y los ingredientes del cacao]. *Journal of Agricultural and Food Chemistry*, 58(19), 10518-10527. <https://doi.org/10.1021/jf102391q>.
- Perea Villamil, J., Villamizar Quiñonez, C., Calle Hoyos, L., Espinosa Safar, A. (1998). *Fermente bien su Cacao: el tambor rotatorio una buena opción*. <http://hdl.handle.net/20.500.12324/20538>.
- Pereira, G., Soccol, V. y Soccol, C. (2016). Current state of research on cocoa and coffee fermentations [Estado actual de la investigación sobre las fermentaciones del cacao y el café]. *Current opinion in food science*, 7. <https://doi.org/10.1016/j.cofs.2015.11.001>.
- Phillips-Mora, W., & Wilkinson, M. J. (2007). Cryopreservation of cocoa (*Theobroma cacao* L.) seeds [Crioconservación de semillas de cacao (*Theobroma cacao* L.)]. *CryoLetters*, 28(6), 435-444.
- Phillips-Mora, W., Wilkinson, M. J., Lass, R. A., Rivera, R. L., Orozco, M. A. y Mockaitis, K. (2009). Extraction of DNA from mucilaginous seeds of *Theobroma cacao* L. for use in PCR [Extracción de ADN de semillas mucilaginosas de *Theobroma cacao* L. para su uso en PCR]. *Plant Molecular Biology Reporter*, 27(2), 183-189.
- Posit team (2023). *RStudio: Integrated Development Environment for R* [RStudio: Entorno de desarrollo integrado para R]. Posit Software, PBC, Boston, MA. URL <http://www.posit.co/>.
- Priest, F. G. y Stewart, G. G. (2006). *Handbook of brewing* (2nd ed.) [Manual de elaboración de la cerveza]. CRC Press.
- Quail, M. A., Smith, M., Coupland, P., Otto, T., Harris, S., Connor, T., Bertoni, A., Swerdlow, H. y Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers [Historia de tres plataformas de secuenciación de nueva generación: comparación de los secuenciadores Ion Torrent, Pacific Biosciences e Illumina MiSeq]. *BMC Genomics*, 13(1), 341.
- Quintero, M., Santos, C., Lima, N. y Bemquerer, M. (2013). Microbiota dynamics and lignocellulosic enzymes of the cocoa fermentation process [Dinámica de la microbiota y enzimas lignocelulósicas del proceso de fermentación del cacao]. *International Journal of Food Microbiology*, 163(2-3), 103-113.
- R Core Team (2021). *R: A language and environment for statistical computing* [R: Lenguaje y entorno de cálculo estadístico]. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rebelo Lima, L. J. (2012). *Microbial ecology of the cocoa chain. Quality aspects and insight into heat-resistant bacterial spores* [Ecología microbiana de la cadena

- del cacao. Aspectos de calidad y conocimiento de las esporas bacterianas termorresistentes]. [Tesis] Wageningen University, Wageningen, N.L. <https://edepot.wur.nl/240014>
- Rezac, S., Kok, C. R. y Heermann, M. (2018). Fermented foods as a dietary source of live organisms [Los alimentos fermentados como fuente dietética de organismos vivos]. *Frontiers in Microbiology*, 9, 1785. <https://doi.org/10.3389/fmicb.2018.01785>
- Rizzatti, G., Lopetuso, L. R., Gibiino, G., Binda, C. y Gasbarrini, A. (2017). Proteobacteria: a common factor in human diseases [Proteobacterias: un factor común en las enfermedades humanas]. *BioMed research international*. <https://doi.org/10.1155/2017/9351507>.
- Rohsius, C. y Engelmann, F. (2007). Manual for the Application of Cocoa Pulp Sampling and Fermentation Inception Protocols [Manual para la Aplicación de Protocolos de Muestreo de Pulpa de Cacao e Inicio de Fermentación]. *Bioversity International*, Montpellier, France.
- Sandoval, E. Y. A., Merchán, P. J. A., Rodríguez, A. F. B., Díaz, E. P. y Serrano-Cely, P. A. (2020). Estado actual de la cacaocultura: una revisión de sus principales limitantes. *Revista Ciencia y agricultura*, 17(2), 1-11. <https://doi.org/10.19053/01228420.v17.n2.2020.10729>.
- Sanger, F., Nicklen, S. y Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors [Secuenciación del ADN con inhibidores de la terminación encadena]. *Proceedings of the National Academy of Sciences*, 74(12), 5463-5467. <https://doi.org/10.1073/pnas.74.12.5463>.
- Sarkar D (2008). *Lattice: Multivariate Data Visualization with R (Lattice: Multivariate Data Visualization with R)*. Springer, New York. ISBN 978-0-387- 75968-5, <http://lmdvr.r-forge.r-project.org>.
- Sauer, K., Stoodley, P., Goeres, D. M., Hall-Stoodley, L., Burmølle, M., Stewart, P. S. y Bjarnsholt, T. (2022). The biofilm life cycle: Expanding the conceptual model of biofilm formation [El ciclo de vida del biofilm: Ampliación del modelo conceptual de formación de biopelículas]. *Nature Reviews Microbiology*, 20(10), 608-620. <https://doi.org/10.1038/s41579-022-00767-0>.
- Schuster, S. C. (2008). Next-generation sequencing transforms today's biology [La secuenciación de nueva generación transforma la biología actual]. *Nature Methods*, 5(1), 16-18. <https://doi.org/10.1038/nmeth1156>.
- Schwan, R. F. y Wheals, A. E. (2004). The microbiology of cocoa fermentation and its role in chocolate quality [La microbiología de la fermentación del cacao y su papel en la calidad del chocolate]. *Critical Reviews in Food Science and Nutrition*, 44(4), 205-221. <https://doi.org/10.1080/10408690490464104>.
- Schwartz, A. y Taras, D. (2014). *Microbiota of the human body: implications in health*

and disease [Microbiota del cuerpo humano: implicaciones en la salud y la enfermedad.]. Springer Science & Business Media.

- Shahanas, E., Panjikaran, S. T., Aneena, E. R., Sharon, C. L. y Remya, P. R. (2019). Health benefits of bioactive compounds from cocoa (*Theobroma cacao*) [Beneficios para la salud de los compuestos bioactivos del cacao (*Theobroma cacao*)]. *Agricultural Reviews*, 40(2), 143-149. <https://doi.org/10.18805/ag.R-1851>.
- Shannon, C. E. (1948). A Mathematical Theory of Communication [Teoría matemática de la comunicación]. *Bell System Technical Journal*, 27(3), 379-423. DOI: <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D. y Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks [Cytoscape: un entorno informático para modelos integrados de redes de interacción biomolecular]. *Genome research*, 13(11), 2498- 2504. <https://doi.org/10.1101/gr.1239303>.
- Shendure, J., Porreca, G. J., Reppas, N. B., et al. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome [Secuenciación polónica multiplex precisa de un genoma bacteriano evolucionado]. *Science*, 309(5741), 1728-1732. <https://doi.org/10.1126/science.1117389>.
- Silva, G. D. C., Kitano, I. T., Ribeiro, I. A. D. F. y Lacava, P. T. (2022). The potential use of actinomycetes as microbial inoculants and biopesticides in agriculture [Uso potencial de los actinomicetos como inoculantes microbianos y biopesticidas en la agricultura]. *Frontiers in Soil Science*, 2. <https://doi.org/10.3389/fsoil.2022.833181>.
- Simpson, G. (2022). *permute: Functions for Generating Restricted Permutations of Data* [permute: Funciones para generar permutaciones restringidas de datos]. R package version 0.9-7, <https://CRAN.R-project.org/package=permute>.
- Smit, G., Smit, B.A. y Engels, W.J. (2005). Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products [Formación de aromas por bacterias lácticas y perfil bioquímico de aromas en productos de queso]. *FEMS Microbiology Reviews*, 29(3), 591-610. <https://doi.org/10.1016/j.femsre.2004.09.002>.
- Sorrenti, V., Ali, S., Mancin, L., Davinelli, S., Paoli, A. y Scapagnini, G. (2020). Cocoa polyphenols and gut microbiota interplay: bioavailability, prebiotic effect, and impact on human health [Interacción entre los polifenoles del cacao y la microbiota intestinal: biodisponibilidad, efecto prebiótico e impacto en la salud humana]. *Nutrients*, 12(7), 1908. <https://doi.org/10.3390/nu12071908>.
- Steiner, E. y Hieronymus, S. (2014). *Yeast: The practical guide to beer fermentation* [Levadura: Guía práctica de la fermentación de la cerveza]. BrewersPublications.
- Steinkraus, K. H. (1997). Fermentations in world food processing [Las fermentaciones en la elaboración de alimentos en el mundo]. *Comprehensive Reviews in Food Science*

- and Food Safety*, 1(1), 23-32. <https://doi.org/10.1111/j.1541-4337.2002.tb00002.x>.
- Tamang, J. P., Watanabe, K. y Holzapfel, W. H. (Eds.). (2016). *Fermented foods and beverages of the world* [Alimentos y bebidas fermentados del mundo]. CRC Press.
- Tamang, J. P., Watanabe, K. y Holzapfel, W. H. (2016). Review: Diversity of microorganisms in global fermented foods and beverages [Revisión: Diversidad de microorganismos en alimentos y bebidas fermentados a nivel mundial]. *Frontiers in Microbiology*, 7,377. <https://doi.org/10.3389/fmicb.2016.00377>
- Tikapunka, T. (2021). Changes in Physico-Chemical and Microbiological Properties in Thai Cocoa Bean Fermentation [Cambios en las propiedades físico- químicas y microbiológicas en la fermentación del cacao en grano tailandés]. *Walailak Journal of Science and Technology (WJST)*, 18(14), 21443-13.
- Truong, D. T., Tett, A., Pasolli, E., Huttenhower, C. y Segata, N. (2017). Microbial strain-level population structure and genetic diversity from metagenomes [Estructura de la población microbiana a nivel de cepa y diversidad genética a partir de metagenomas]. *Genome research*, 27(4), 626-638. <https://doi.org/10.1101/gr.216242.116>.
- USDA Foreign Agricultural Service (2019). Advances and opportunities of cacao production in Guatemala. <https://fas.usda.gov/data/guatemala-advances-and-opportunities-cacao-production-guatemala>.
- Viesser, J. A., de Melo Pereira, G. V., de Carvalho Neto, D. P., Favero, G. R., de Carvalho, J. C., Goés-Neto, A. y Soccol, C. R. (2021). Global cocoa fermentation microbiome: revealing new taxa and microbial functions by next generation sequencing technologies [Microbioma mundial de la fermentación del cacao: revelación de nuevos taxones y funciones microbianas mediante tecnologías de secuenciación de nueva generación]. *World Journal of Microbiology and Biotechnology*, 37, 1-17. <https://doi.org/10.1007/s11274-021-03079-2>.
- Whittaker, R. J., Willis, K. J. y Field, R. (2001). Scale and species richness: towards a general, hierarchical theory of species diversity [Escala y riqueza de especies: hacia una teoría general y jerárquica de la diversidad de especies]. *Journal of biogeography*, 28(4), 453-470. <https://doi.org/10.1046/j.1365-2699.2001.00563.x>.
- Wickham, H., François, R., Henry, L., Müller, K. y Vaughan, D. (2023). *dplyr: A Grammar of Data Manipulation* [dplyr: Una gramática de la manipulación de datos]. R package version 1.1.3, <https://CRAN.R-project.org/package=dplyr>.
- Wickham, H., Vaughan, D. y Girlich, M. (2023). *tidyr: Tidy Messy Data* [tidyr: Ordenar datos desordenados]. R package version 1.3.0, <https://CRAN.R-project.org/package=tidyr>.
- Wilson, I. G. (1997). Inhibition and facilitation of nucleic acid amplification [Inhibición y facilitación de la amplificación de ácidos nucleicos]. *Applied and Environmental*

Microbiology, 63(10), 3741-3751. <https://doi.org/10.1128/aem.63.10.3741-3751.1997>.

- Wishart, D. S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine [Nuevas aplicaciones de la metabolómica en el descubrimiento de fármacos y la medicina de precisión]. *Nature Reviews Drug Discovery*, 15(7), 473-484. <https://doi.org/10.1038/nrd.2016.32>
- Wood, D.E. y Salzberg, S.L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments [Kraken: clasificación ultrarrápida de secuencias metagenómicas mediante alineamientos exactos]. *Genome Biology*, 15(R46). <https://doi.org/10.1186/gb-2014-15-3-r46>
- World Bank Group. (2023). *Colombia Country Climate and Development Report* [Informe sobre el clima y el desarrollo en Colombia]. World Bank Group.
- Wouters, D., Grosch, R. y Soares, E.V. (2016). Lactic Acid Bacteria as a Starter Culture for Sauerkraut Fermentation [Bacterias lácticas como cultivo iniciador para la fermentación del chucrut]. In F. Mozzi, R.R. Raya, & G.M. Vignolo (Eds.), *Biotechnology of Lactic Acid Bacteria: Novel Applications* (pp. 251-264). John Wiley & Sons.
- Zachariassen, T., Petersen, A. Ø., Brejnrod, A., Vestergaard, G. A., Eklund, A. y Nielsen, H. B. (2023). Identification of representative species-specific genes for abundance measurements [Identificación de genes representativos de cada especie para las mediciones de abundancia]. *Bioinformatics Advances*, 3(1), <https://doi.org/10.1093/bioadv/vbad060>.
- Zareie, Z., Moayedi, A., Garavand, F., Tabar-Heydar, K., Khomeiri, M. y Maghsoudlou, Y. (2023). Probiotic Properties, Safety Assessment, and Aroma-Generating Attributes of Some Lactic Acid Bacteria Isolated from Iranian Traditional Cheese [Propiedades probióticas, evaluación de la seguridad y atributos aromatizantes de algunas bacterias lácticas aisladas del queso tradicional iraní]. *Fermentation*, 9(4), 338. <https://doi.org/10.3390/fermentation9040338>.
- Zong, P., Hu, W., Huang, Y., An, H., Zhang, Q., Chai, Z. y Cai, W. (2023). Effects of adding cocoa fermentation medium on cigar leaves in agricultural fermentation stage [Efectos de la adición de medio de fermentación de cacao sobre las hojas de puro en fase de fermentación agrícola]. *Frontiers in Bioengineering and Biotechnology*, 11. <https://doi.org/10.3389/fbioe.2023.1251413>.