



Universidad Internacional de La Rioja
Facultad de Ciencias de la Salud

Máster Universitario en Bioinformática

Análisis de la microbiota y del resistoma de muestras metagenómicas fecales de individuos con SIBO

Trabajo fin de Estudio presentado por:	Laura Ruiz Ripa
Tipo de trabajo:	Redacción de trabajo original relacionado con bases de datos
Director/a:	Francisco Javier Ruiz Ojeda
Fecha:	09/07/2025

Resumen

Este estudio analiza la microbiota y el resistoma en muestras fecales de individuos diagnosticados con sobrecrecimiento bacteriano del intestino delgado (SIBO), mediante el análisis de datos de metagenómica de amplicón y *shotgun*. A partir de datos públicos del American Gut Project, se seleccionaron y emparejaron individuos con y sin SIBO para comparar la diversidad microbiana y la presencia de genes de resistencia a antibióticos (ARGs). El análisis bioinformático reveló que, aunque no se observaron diferencias significativas en los índices de diversidad alfa global, sí se identificaron alteraciones en la equidad (índice de *evenness*) y en la diversidad beta, lo que sugiere diferencias estructurales en las comunidades microbianas entre ambos grupos. A nivel taxonómico, se detectó una mayor abundancia de la familia Enterobacteriaceae y del filo Pseudomonadota en individuos con SIBO. En cuanto al resistoma, no se encontraron diferencias estadísticamente significativas en la abundancia relativa global de ARGs entre grupos, salvo en el gen *ant(6)-Ia*, cuya abundancia fue significativamente menor en individuos con SIBO. Estos hallazgos aportan evidencia de una disbiosis intestinal asociada al SIBO y subrayan la utilidad de los enfoques metagenómicos para su caracterización.

Palabras clave: SIBO; Microbiota; Resistoma

Abstract

This study analyzes the microbiota and resistome in fecal samples from individuals diagnosed with small intestinal bacterial overgrowth (SIBO), using amplicon and shotgun metagenomic data analysis. Public data from the American Gut Project were used to select and match individuals with and without SIBO in order to compare microbial diversity and the presence of antibiotic resistance genes (ARGs). Bioinformatic analysis revealed that, although no significant differences were observed in global alpha diversity indices, alterations in evenness and beta diversity were identified, suggesting structural differences in the microbial communities between groups. At the taxonomic level, a higher abundance of the Enterobacteriaceae family and the Pseudomonadota phylum was detected in individuals with SIBO. Regarding the resistome, no statistically significant differences were found in the overall relative abundance of ARGs between groups, except for the *ant(6)-Ia* gene, whose abundance was significantly lower in individuals with SIBO. These findings provide evidence of intestinal dysbiosis associated with SIBO and highlight the value of metagenomic approaches for its characterization.

Keywords: SIBO; Microbiota; Resistome

Índice de contenidos

1.	Introducción	9
1.1.	Justificación	9
1.2.	Planteamiento del problema	9
1.3.	Objetivos.....	10
2.	Marco teórico	12
2.1.	El microbioma intestinal	12
2.1.1.	Papel de la microbiota intestinal en la salud humana	12
2.1.2.	Composición de la microbiota intestinal	14
2.2.	Sobrecrecimiento bacteriano del intestino delgado	15
2.2.1.	Definición y características clínicas	15
2.2.2.	Epidemiología	17
2.2.3.	Tratamiento	18
2.3.	Resistencia a antibióticos	19
2.3.1.	Dianas de los antibióticos y mecanismos de resistencia a antibióticos	19
2.3.2.	El intestino como reservorio de genes de resistencia a antibióticos	22
2.4.	Análisis del microbioma.....	22
2.4.1.	Ciencias ómicas aplicadas al estudio del microbioma.....	22
2.4.2.	Metagenómica.....	23
3.	Marco metodológico	29
3.1.	Diseño del estudio	29
3.1.1.	Tipo de estudio	29
3.1.2.	Base de datos.....	29
3.1.3.	Grupos de estudio	29
3.1.4.	Tipos de muestras.....	31

3.2.	Obtención de los datos	31
3.2.1.	Lecturas metagenómicas	32
3.2.2.	Metadatos asociados	32
3.3.	Análisis bioinformático de la microbiota	33
3.3.1.	Preprocesamiento de las lecturas	33
3.3.2.	Asignación taxonómica y análisis filogenético	34
3.3.3.	Análisis de diversidad y composición microbiana	34
3.4.	Análisis bioinformático del resistoma	36
3.4.1.	Detección y cuantificación de ARGs	36
3.4.2.	Análisis estadísticos	38
4.	Resultados	40
4.1.	Análisis bioinformático de la microbiota	40
4.1.1.	Preprocesamiento de las lecturas	40
4.1.2.	Análisis de diversidad y composición microbiana	40
4.2.	Análisis bioinformático del resistoma	46
4.2.1.	Detección y cuantificación de ARGs	46
4.2.2.	Análisis estadísticos	50
5.	Discusión	54
5.1.	Análisis de la microbiota	54
5.1.1.	Diversidad alfa y beta	54
5.1.2.	Composición taxonómica	56
5.2.	Análisis del resistoma	57
5.3.	Limitaciones	58
5.4.	Prospectiva	59
6.	Conclusiones	61

7. Referencias bibliográficas	62
8. Anexos	67
Anexo A. Comandos utilizados para el análisis taxonómico con QIIME2	67
Anexo B. Comandos utilizados para la detección de genes de resistencia a antibióticos	
78	
9. Índice de siglas y acrónimos	83

Índice de figuras

Figura 1. Principales tipos de análisis y flujos de trabajo ómicos. Adaptado de Arıkan et al., 2023 (16).....	23
Figura 2. Regiones conservadas, variables e hipervariables del gen bacteriano ARNr 16S, así como los cebadores utilizados para su amplificación por PCR. Adaptado de Shahi et al., 2017 (22).....	25
Figura 3. Flujo de trabajo en metagenómica shotgun (izquierda) y metagenómica de amplicón (derecha). Elaboración propia	26
Figura 4. Comparación del índice de Shannon entre los grupos con y sin diagnóstico de SIBO. Elaboración propia.....	41
Figura 5. Comparación del número de características observadas entre los grupos con y sin diagnóstico de SIBO. Elaboración propia	41
Figura 6. Comparación del índice de Faith entre los grupos con y sin diagnóstico de SIBO. Elaboración propia.....	42
Figura 7. Comparación del índice de evenness entre los grupos con y sin diagnóstico de SIBO. Elaboración propia.....	43
Figura 8. PCoA de las distancias de Bray-Curtis, Jaccard, weighted UniFrac y unweighted UniFrac de los grupos con y sin diagnóstico de SIBO. Elaboración propia.....	43
Figura 9. Distribución de la abundancia relativa de la familia Enterobacteriaceae y del filo Pseudomonadota en el grupo SIBO y Control. Elaboración propia	45
Figura 10. Prevalencia de ARGs en las muestras metagenómicas del grupo SIBO y Control. Elaboración propia.....	48
Figura 11. Prevalencia de ARGs por clase. Elaboración propia	49
Figura 12. Abundancia relativa (log RPKM) de los diferentes ARGs de los grupos SIBO y Control. Elaboración propia.....	51
Figura 13. Abundancia relativa (log RPKM) de los diferentes ARGs agrupada por familia de antibióticos a los que confieren resistencia de los grupos SIBO y Control. Elaboración propia	53

Índice de tablas

Tabla 1. Principales características y géneros bacterianos más representativos de los filos Bacillota, Bacteroidota, Actinomycetota y Pseudomonadota.	15
Tabla 2. Principales familias de antibióticos, ejemplos de antibióticos, mecanismo de acción, mecanismos de resistencia y ejemplos de genes de resistencia.....	20
Tabla 3. Principales diferencias entre la metagenómica de amplicón y la metagenómica <i>shotgun</i>	24
Tabla 4. Covariables utilizadas para la selección de controles mediante la función <code>matchit()</code>	30
Tabla 5. Abundancia relativa promedio de los filos predominantes en los grupos SIBO y Control.	44
Tabla 6. Estadísticos de la prueba de Wilcoxon de la abundancia relativa de los ARGs entre grupos.....	50
Tabla 7. Estadísticos de la prueba de Wilcoxon de la abundancia relativa de los ARGs agrupados por familia de antibiótico entre grupos.....	52

1. Introducción

En las últimas décadas, el estudio del microbioma intestinal ha cobrado gran relevancia debido a su estrecha relación con múltiples aspectos de la salud humana. Particularmente, el sobrecrecimiento bacteriano del intestino delgado (SIBO) representa una condición clínica de creciente interés, caracterizada por una proliferación anómala de bacterias en regiones del intestino donde normalmente su densidad es reducida. Esta alteración puede generar síntomas gastrointestinales inespecíficos como hinchazón, diarrea y dolor abdominal, lo que complica su diagnóstico y tratamiento adecuado.

1.1. Justificación

El interés en la investigación de la microbiota y el resistoma intestinal en individuos con SIBO surge de la necesidad de comprender mejor la disbiosis subyacente en esta patología, así como las posibles consecuencias del uso repetido de tratamientos antibióticos, práctica común en estos pacientes. A pesar de que diversos estudios han documentado desequilibrios microbianos en otras enfermedades gastrointestinales, como el síndrome del intestino irritable (SII) o la enfermedad inflamatoria intestinal (EII), existe una notable escasez de estudios metagenómicos que caractericen de forma integral tanto la microbiota como el resistoma en individuos con SIBO.

La caracterización precisa del ecosistema microbiano intestinal tendría un gran impacto en el diagnóstico y tratamiento de la enfermedad. Además, en un contexto clínico donde el uso indiscriminado de antibióticos ha contribuido significativamente al aumento de la resistencia, conocer el resistoma en estos pacientes puede ofrecer datos clave para mejorar la gestión y reducir el riesgo de fracaso terapéutico, así como para contener la propagación de genes de resistencia a antibióticos (ARGs) en la comunidad.

1.2. Planteamiento del problema

La escasez de estudios metagenómicos que analicen de forma comparativa la microbiota intestinal y los perfiles de resistencia a antibióticos en individuos con SIBO frente a individuos sin esta condición representa una barrera para profundizar en la comprensión de los

mecanismos implicados en el desarrollo y la persistencia de la enfermedad, así como en su respuesta terapéutica.

La pregunta central que guía este trabajo fin de máster es: **¿Existen diferencias significativas en la microbiota y en el resistoma fecal entre individuos con SIBO y controles sin esta condición?** A partir de esta cuestión se plantea una serie de análisis específicos, que incluyen la identificación de taxones diferenciales, el estudio de la diversidad microbiana y la detección de ARGs. Con este enfoque, se busca generar conocimiento científico que permita una mejor caracterización del SIBO desde una perspectiva microbiológica.

1.3.Objetivos

El objetivo general de este trabajo es analizar la microbiota y el resistoma presente en muestras metagenómicas fecales de individuos con SIBO, en comparación con controles sin esta condición, con el fin de identificar diferencias en la composición bacteriana y en la abundancia de ARGs entre ambos grupos.

Los objetivos específicos son:

1. **Caracterizar la microbiota fecal bacteriana:**

- a) A partir de los datos crudos de muestras metagenómicas de heces, realizar el preprocesamiento de las lecturas y determinar la composición taxonómica de la microbiota bacteriana.
- b) Evaluar la diversidad alfa y la diversidad beta de la microbiota fecal bacteriana de los individuos con SIBO y controles.
- c) Realizar un análisis diferencial de la composición de la microbiota fecal entre individuos con SIBO y controles para identificar posibles biomarcadores bacterianos asociados a la presencia de SIBO.

2. **Caracterizar el resistoma fecal:**

- a) Identificar ARGs en las muestras fecales mediante alineamiento con bases de datos especializadas.

- b) Realizar un análisis comparativo de la abundancia relativa de ARGs entre individuos con SIBO y controles.

2. Marco teórico

2.1.El microbioma intestinal

Bidell *et al.*, (1) definen el microbioma intestinal como un ecosistema complejo formado por aproximadamente 100 billones de microorganismos y su contenido genético. Por su parte, la microbiota hace referencia al conjunto de microorganismos que conforman el microbioma, incluyendo bacterias, virus, arqueas y organismos eucariotas, como los hongos (1).

2.1.1. Papel de la microbiota intestinal en la salud humana

La microbiota intestinal tiene una gran relevancia en la salud y en diferentes patologías humanas. En un estado de homeostasis, es decir, cuando existe un equilibrio entre los microorganismos que habitan en el intestino y el organismo, así como entre las distintas especies microbianas entre sí, la microbiota contribuye activamente al metabolismo, la digestión y la modulación del sistema inmunológico. Por el contrario, la alteración de esta homeostasis, conocida como disbiosis, se ha asociado con infecciones, enfermedades y diversas condiciones médicas (1,2).

Se considera que la diversidad microbiana intestinal juega un papel clave en su capacidad para resistir alteraciones y mantener el estado de homeostasis. Por lo tanto, una disminución en la variedad de especies o cambios en la cantidad de ciertos filos bacterianos podrían ser señales de un estado de disbiosis (1).

A continuación se desarrollan los procesos más importantes en los que interviene la microbiota intestinal.

2.1.1.1. Funciones de protección

Resistencia a la colonización

Las bacterias comensales protegen al huésped impidiendo la proliferación de microorganismos patógenos. Esto lo logran mediante la producción de bacteriocinas, unas proteínas antimicrobianas que destruyen o inhiben el crecimiento de las bacterias invasoras a través de mecanismos como la degradación de ácidos nucleicos o la perforación de la

membrana celular. Filos como Bacteroidota y Pseudomonadota incluyen especies capaces de producir bacteriocinas. También compiten con otros microorganismos por nutrientes y espacio, limitando el crecimiento de especies nocivas (1).

Modulación de la inmunidad

La microbiota intestinal participa en la regulación de la inmunidad innata y adaptativa. Ayuda al sistema inmunológico a distinguir entre microorganismos beneficiosos y patógenos, promoviendo respuestas proinflamatorias y antiinflamatorias. Bacterias como *Clostridium* spp., *Bacteroides fragilis* y *Bifidobacterium infantis* favorecen la producción de inmunoglobulina A, y algunas especies de filo Bacillota y la especie *Faecalibacterium prausnitzii* contribuyen a disminuir citoquinas proinflamatorias, lo que protege el epitelio intestinal (1–3).

2.1.1.2. Funciones metabólicas

Homeostasis de los ácidos biliares

Los ácidos biliares son fundamentales para la digestión de grasas y la absorción de vitaminas liposolubles. Géneros como *Clostridium*, *Lactobacillus*, *Enterococcus* y *Bacteroides* poseen enzimas que transforman los ácidos biliares primarios conjugados en su forma no conjugada, favoreciendo su recirculación. Además, bacterias como *Clostridium scindens* convierten estos ácidos en su forma secundaria, que, junto con los primarios, tienen propiedades antimicrobianas y ayudan a prevenir infecciones, como las causadas por *Clostridioides difficile* (1,2).

Producción de ácidos grasos de cadena corta

Mediante la fermentación de la fibra dietética, bacterias del filo Bacillota, como *Roseburia* spp. y *Ruminococcus* spp., producen ácidos grasos de cadena corta, incluyendo butirato, propionato y acetato. Además de su función energética, los ácidos grasos de cadena corta tienen efectos beneficiosos sobre el sistema inmunológico y la integridad del epitelio intestinal (1,2).

Producción de vitaminas

La microbiota intestinal también contribuye a la producción de vitaminas. Especies de los filos Bacillota, Actinomycetota y Pseudomonadota participan en la síntesis de vitamina B12, crucial para funciones metabólicas. Otras bacterias como *Bacteroides fragilis*, *Eubacterium lentum*, *Enterobacter agglomerans*, *Serratia marcescens* y *Enterococcus faecium* participan en la producción vitamina K, fundamental para la coagulación sanguínea y la salud ósea. Por ello, cualquier cambio en estas poblaciones pueden afectar negativamente la biodisponibilidad de estas vitaminas (1,3).

2.1.2. Composición de la microbiota intestinal

En los seres humanos, el intestino es la región del cuerpo que alberga la mayor cantidad y diversidad de especies (4,5). Su composición varía en función de numerosos factores, como la genética, la edad, la dieta, el medio ambiente y el uso de medicamentos, lo que hace que el microbioma de cada individuo sea único. Sin embargo, existen ciertas características comunes que definen un microbioma intestinal saludable (1,4).

La microbiota bacteriana constituye la fracción predominante del microbioma intestinal, estimándose que, en condiciones de salud, está formada por más de 1.000 especies diferentes, concentrándose la mayoría en el colon (1), que alberga un hasta 10^{12} unidades formadoras de colonias (UFCs)/mL (4,6). A lo largo del tracto digestivo, la composición microbiana varía, siendo más escasa en el estómago y el intestino delgado (4,6). En este último se estima que hay menos de 10^3 UFCs/mL (7).

Alrededor del 99% de las bacterias intestinales son anaerobias, aunque en el colon también se encuentra una alta densidad de bacterias aerobias (4,5). De los aproximadamente 50 filos presentes en la microbiota intestinal humana sana, los filos *Bacillota* y *Bacteroidota* representan el 90% y si sumamos los filos *Actinomycetota* y *Pseudomonadota* alcanza hasta el 99% (1,4,5). Los géneros predominantes son *Bacteroides*, *Clostridium*, *Peptococcus*, *Bifidobacterium*, *Eubacterium*, *Ruminococcus* y *Faecalibacterium*. De estos, *Bacteroides* es el más abundante, representando aproximadamente el 30% de las bacterias intestinales, lo que sugiere su función clave en el ecosistema intestinal (4). En la Tabla 1 se puede ver las principales características y ejemplos de géneros bacterianos de estos filos.

Tabla 1. Principales características y géneros bacterianos más representativos de los filos Bacillota, Bacteroidota, Actinomycetota y Pseudomonadota.

	Principales características	Géneros bacterianos
Bacillota	Gram-positivos; normalmente anaerobios estrictos o facultativos; géneros muy diversos	<i>Clostridium, Enterococcus, Eubacterium, Faecalibacterium, Lactobacillus, Roseburia, Ruminococcus, Streptococcus</i>
Bacteroidota	Gram-negativos; normalmente anaerobios estrictos	<i>Bacteroides, Prevotella</i>
Actinomycetota	Gram-positivos; normalmente anaerobios estrictos o facultativos	<i>Bifidobacterium, Corynebacterium, Eggerthella</i>
Pseudomonadota	Gram-negativos; principalmente anaerobios facultativos; incluye numerosas especies patógenas	<i>Enterobacter, Escherichia, Klebsiella, Serratia</i>

Adaptado de Bidel *et al.* 2022 (1)

2.2. Sobrecrecimiento bacteriano del intestino delgado

2.2.1. Definición y características clínicas

2.2.1.1. Definición

El SIBO se define como una forma de disbiosis caracterizada por el aumento excesivo y anormal del número de bacterias que colonizan el intestino delgado, así como por variaciones en su composición taxonómica (6–9).

2.2.1.2. Diagnóstico

Para el diagnóstico del SIBO, uno de los métodos que ha sido ampliamente utilizado es el cultivo del fluido aspirado endoscópicamente del lumen del intestino delgado, considerando como umbral diagnóstico un recuento igual o superior a 10^3 UFCs/mL de aspirado duodenal o

yeyunal (6–9). Sin embargo, esta técnica tiene sus limitaciones, ya que implica un procedimiento invasivo, conlleva un coste elevado, se restringe a la identificación del sobrecrecimiento bacteriano en regiones proximales del intestino y puede no detectar microorganismos difíciles de cultivar en condiciones de laboratorio (7,9).

Actualmente, la prueba más común que se realiza en la práctica médica diaria para el diagnóstico del SIBO es la prueba de aliento. Esta prueba mide la concentración de gases, como hidrógeno o metano, en el aire exhalado después de que el paciente ingiere un sustrato carbohidratado, normalmente lactulosa o glucosa. Se basa en el principio de que las células del cuerpo humano no producen los gases mencionados, por lo que un aumento de sus niveles en el aire exhalado indica una mayor actividad metabólica de los carbohidratos no digeridos, los cuales son fermentados por los microorganismos presentes en el intestino delgado (7–9). Los umbrales de aumento de los niveles de gases en la prueba de aliento para un diagnóstico positivo de SIBO dependen del sustrato utilizado (9).

Aunque la prueba de aliento es no invasiva y relativamente económica, su sensibilidad y especificidad son limitadas (7,10). Un estudio concluyó que, según el sustrato utilizado, lactulosa o glucosa, la sensibilidad de esta prueba fue del 42% y 54% y la especificidad del 70% y 83%, respectivamente (10). Además, hay que tener en cuenta que la lactulosa puede ocasionar falsos positivos, mientras que la glucosa solo detecta el SIBO proximal (9).

Las pruebas de cultivo y de aliento para diagnosticar el SIBO no siempre arrojan resultados coincidentes, con una concordancia aproximada del 65% (8). Esto indica que un único método de prueba diagnóstica podría no ser suficiente para un diagnóstico definitivo y que es recomendable pruebas adicionales cuando se sospeche de SIBO.

2.2.1.3. Síntomas clínicos comunes

El SIBO puede causar una amplia variedad manifestaciones clínicas inespecíficas e incluso cursar de manera asintomática. Los síntomas más comunes, descritos por dos tercios de los pacientes con SIBO, y que varían en su frecuencia, duración y gravedad, son dolor y distensión abdominal, acumulación de gas excesiva y flatulencia, sentimiento de hinchazón y hábitos intestinales alterados (predominantemente diarrea, pero también estreñimiento) (7–9).

En casos más severos, puede producirse pérdida de peso, esteatorrea y signos de malabsorción, como deficiencia de vitamina D o B12. Estas deficiencias pueden provocar síntomas como fatiga, debilidad, ataxia y calambres musculares (7,8).

2.2.2. Epidemiología

2.2.2.1. Prevalencia

La prevalencia del SIBO en la población general no está claramente establecida, debido en gran parte a la inespecificidad y similitud de sus síntomas con los de otras afecciones gastrointestinales y a que puede cursar de forma asintomática (7,8). Algunos estudios han estimado su prevalencia en un rango que va del 2% al 22% (8), aunque estas cifras pueden variar considerablemente según el método diagnóstico utilizado y la población analizada.

2.2.2.2. Factores predisponentes

El equilibrio microbiano en el intestino delgado está regulado por varios mecanismos fisiológicos que, cuando se ven alterados, pueden favorecer el desarrollo de SIBO. Estos mecanismos protectores se agrupan en tres categorías principales (8):

- 1. Defensas antimicrobianas gastrointestinales:** Están relacionados con la secreción de ácido gástrico, bilis, enzimas pancreáticas, etc. Su disfunción puede deberse a condiciones como gastritis atrófica, pancreatitis crónica, fibrosis quística o inmunodeficiencias (8).
- 2. Motilidad intestinal:** Los trastornos de motilidad afectan a la limpieza del intestino delgado, por lo que trastornos del complejo motor migratorio y peristalsis retrógrada, comunes en enfermedades como la diabetes, enfermedad de Parkinson o tras el uso de fármacos como opiáceos y anticolinérgicos, se asocian con una mayor prevalencia de SIBO (8,9).
- 3. Alteraciones anatómicas:** Las anomalías del tracto gastrointestinal, como divertículos del intestino delgado, estenosis o fístulas, y las cirugías, como la colectomía, contribuyen a la aparición de SIBO al modificar la arquitectura del intestino (8).

Además de estos factores, la edad avanzada, el sexo femenino y ciertas comorbilidades sistémicas (como SII, diabetes, enfermedad celíaca, hipotiroidismo o EII) también se han vinculado al desarrollo de SIBO. En particular, se ha documentado una elevada prevalencia de SIBO en pacientes con SII, estimándose que entre el 50 % y 70 % de estos pacientes presentan SIBO, una tasa significativamente mayor que en la población general (8,9).

2.2.3. Tratamiento

El objetivo principal del tratamiento en individuos con SIBO es erradicar el crecimiento bacteriano en el intestino delgado para aliviar los síntomas (8,9). Asimismo, es fundamental identificar y tratar la causa subyacente o predisponente que favorece su aparición (7–9). El abordaje terapéutico también debe incluir el mantenimiento de la remisión, la prevención de recaídas y la corrección de posibles déficits nutricionales (8,9).

Los antibióticos orales juegan un papel clave en la remisión del SIBO al modular la microbiota intestinal y mejorar los síntomas (6,8). Aunque la evidencia sobre su eficacia es limitada debido a la falta de ensayos clínicos a gran escala (6,8), se han utilizado varios antibióticos de amplio espectro para el tratamiento del SIBO, entre ellos norfloxacin, ciprofloxacino, metronidazol, tetraciclina, doxiciclina, neomicina, cotrimoxazol, amoxicilina-ácido clavulánico y rifaximina, siendo esta última la más estudiada y utilizada. (6–8). La rifaximina destaca por su eficacia y seguridad: no se absorbe en el tracto gastrointestinal, tiene un amplio espectro antibacteriano, escasos efectos secundarios y actúa como un eubiótico, protegiendo la microbiota beneficiosa. Además, no induce resistencia bacteriana transferible, lo que permite su uso repetido (6). Un desafío importante respecto al tratamiento antibiótico es la alta tasa de recurrencia, presente en hasta la mitad de los casos tras un tratamiento exitoso (6,7). Esto puede requerir retratamientos, que conllevan riesgos como resistencia bacteriana, diarrea (incluida la provocada por *Clostridioides difficile*), intolerancia y disbiosis intestinal (6).

El uso de probióticos en el tratamiento del SIBO ha sido explorado por su capacidad para modular la microbiota intestinal y competir con bacterias patógenas (8,11). Si bien algunos estudios y metaanálisis han mostrado que pueden reducir la producción de hidrógeno en las pruebas de aliento, los efectos varían según la cepa utilizada (8).

Finalmente, puede emplearse de forma complementaria una dieta baja en carbohidratos fermentables que sirven como sustrato para las bacterias, reduciendo así su proliferación y la producción de gases (8,11). No obstante, su aplicación debe ser temporal y no debe repetirse si no resulta eficaz (8).

2.3. Resistencia a antibióticos

El descubrimiento de los antibióticos marcó un antes y un después en la medicina al permitir tratar eficazmente infecciones y reducir significativamente la mortalidad asociada enfermedades bacterianas. Su uso ha sido clave para el desarrollo de procedimientos clínicos complejos y el avance de la medicina moderna (12). Sin embargo, su eficacia se ve cada vez más amenazada por la creciente resistencia a estos compuestos, favorecida por su uso excesivo y por la capacidad de los microorganismos para adquirir resistencia. Actualmente, la resistencia a antibióticos representa una de las principales amenazas para la salud humana y animal (12–14).

2.3.1. Dianas de los antibióticos y mecanismos de resistencia a antibióticos

La resistencia puede deberse a la presencia de ARGs o a mutaciones en los genes que son diana de los antibióticos. Es importante distinguir entre resistencia intrínseca, presente de forma natural en ciertas especies bacterianas, y resistencia adquirida, que representa un riesgo particular por su capacidad de diseminarse entre especies mediante mecanismos de transferencia horizontal (12,14). Esta diseminación se produce porque los genes de resistencia adquirida se localizan en elementos genéticos móviles, como plásmidos, integrones y transposones, que facilitan la propagación de ARGs en distintos ambientes, incluyendo el tracto intestinal (12,14).

En la Tabla 2 se resumen las familias de antibióticos que se emplean comúnmente en medicina humana y animal, su mecanismo de acción, los principales mecanismos de resistencia bacteriana y los genes de resistencia asociados.

Tabla 2. Principales familias de antibióticos, ejemplos de antibióticos, mecanismo de acción, mecanismos de resistencia y ejemplos de genes de resistencia.

Familia de antibióticos	Ejemplos de antibióticos	Mecanismo de acción	Mecanismos de resistencia	Ejemplos de genes de resistencia más comunes o de mayor interés
β-lactámicos	Penicilina, cefalosporinas, carbapenémicos	Inhibición de la síntesis de la pared celular	Inactivación enzimática	<i>blaTEM</i> , <i>blaCTX-M</i> , <i>blaOXA</i> , <i>cfxA</i> , <i>blaZ</i>
			Modificación de la diana	<i>mecA</i> , <i>mecC</i>
			Bombas de eflujo	<i>AcrAB-TolC</i> , <i>MexAB-OprM</i> ,
Tetraciclinas	Tetraciclina, doxiciclina, minociclina	Inhibición de la síntesis proteica (subunidad 30S ribosomal)	Bombas de eflujo	<i>tet(L)</i> , <i>tet(B)</i> , <i>tet(40)</i>
			Protección ribosomal	<i>tet(O)</i> , <i>tet(W)</i> , <i>tet(Q)</i>
			Modificación enzimática	<i>tet(X)</i>
Macrólidos	Eritromicina, azitromicina	Inhibición de la síntesis proteica (subunidad 50S ribosomal)	Modificación de la diana	<i>erm(B)</i> , <i>erm(F)</i> , <i>erm(G)</i> , <i>erm(X)</i>
			Bombas de eflujo	<i>mefA</i> , <i>msrD</i> , <i>msrE</i>
Lincosamidas	Clindamicina	Inhibición de la síntesis proteica (subunidad 50S ribosomal)	Modificación de la diana	<i>erm(B)</i> , <i>erm(F)</i> , <i>erm(G)</i>
			Bombas de eflujo	<i>mefA</i>
			Inactivación enzimática	<i>Inu(A)</i> , <i>Inu(B)</i> , <i>Inu(C)</i>
Aminoglucósidos	Gentamicina, estreptomicina, tobramicina	Inhibición de la síntesis proteica (subunidad 30S ribosomal)	Modificación enzimática	<i>aac(6')-Ib</i> , <i>aph(3')-IIIa</i> , <i>ant(6)-Ia</i> , <i>ant(3'')-Ia</i>

Familia de antibióticos	Ejemplos de antibióticos	Mecanismo de acción	Mecanismos de resistencia	Ejemplos de genes de resistencia más comunes o de mayor interés
Quinolonas	Ciprofloxacino, levofloxacino, norfloxacino	Inhibición de la ADN girasa y topoisomerasa IV	Modificación de la diana	Mutaciones en ADN girasa/topoisomerasa
			Protección de la diana	<i>qnrA</i> , <i>qnrS</i> ,
			Inactivación enzimática	<i>aac(6')-Ib-cr</i>
Rifamicinas	Rifampicina, rifaximina	Inhibición de la ARN polimerasa dependiente de ADN	Modificación de la diana	Mutaciones en <i>rpoB</i>
Glicopéptidos	Vancomicina	Inhibición de la síntesis de la pared celular	Modificación de la diana	<i>vanA</i> , <i>vanB</i>
Fenicoles	Cloranfenicol, florfenicol	Inhibición de la síntesis proteica (subunidad 50S ribosomal)	Inactivación enzimática	<i>cat</i>
			Bombas de eflujo	<i>fexA</i> , <i>fexB</i>
			Modificación de la diana	<i>cfr</i> , <i>cfr(B)</i> , <i>cfr(C)</i> , <i>cfr(D)</i>
			Protección ribosomal	<i>poxA</i> , <i>optrA</i>
Sulfonamidas y trimetoprim	Cotrimoxazol	Inhibición de la síntesis de folato	Modificación de la diana	<i>sul1</i> , <i>sul2</i> , <i>dfrA</i> , <i>dfrD</i>
Oxazolidinodas	Linezolid, tedizolid	Inhibición de la síntesis proteica (subunidad 50S ribosomal)	Modificación de la diana	Mutaciones en el ARNr 23S y cambios aminoacídicos en las proteínas ribosomales L3, L4 y L22 <i>cfr</i> , <i>cfr(B)</i> , <i>cfr(C)</i> , <i>cfr(D)</i>

Elaboración propia

2.3.2. El intestino como reservorio de genes de resistencia a antibióticos

El uso de antibióticos puede alterar significativamente la microbiota intestinal, cambiando su composición taxonómica y funcional. Este fenómeno genera un desequilibrio en las poblaciones microbianas del intestino, abriendo oportunidades para la colonización de patógenos y microorganismos resistentes a antibióticos (13,14). Esto se debe a que los antibióticos, además de eliminar las bacterias patógenas, también afectan a microorganismos beneficiosos que desempeñan funciones esenciales en el mantenimiento de la salud intestinal. La pérdida de estas bacterias puede dar lugar a una mayor susceptibilidad a infecciones y a la proliferación de organismos resistentes. En este sentido, las bacterias que adquieren ARGs tienen una ventaja selectiva, lo que puede llevar a una expansión del resistoma en microbiomas de pacientes que están expuestos al uso de antibióticos (14,15). Así, la microbiota comensal del intestino actúa como un reservorio de ARGs, de forma que bacterias potencialmente patógenas pueden adquirir rápidamente resistencia durante una infección o tras la exposición a antibióticos (13–15). En este contexto, el análisis metagenómico ha sido una herramienta crucial para ampliar el conocimiento sobre la diversidad de ARGs presentes en el intestino (13,14).

2.4. Análisis del microbioma

2.4.1. Ciencias ómicas aplicadas al estudio del microbioma

El estudio del microbioma ha experimentado una profunda transformación con el auge de las ciencias ómicas de alto rendimiento, lo que han permitido superar las limitaciones de los métodos clásicos basados en cultivo y observación microscópica (16,17). Estas tecnologías ofrecen un análisis integral de las comunidades microbianas desde distintos niveles moleculares, como se puede ver en la Figura 1, proporcionando una visión holística de su composición, funciones y dinámica ecológica.

Entre ellas, la metagenómica permite identificar y caracterizar el material genético de todos los microorganismos presentes en una muestra, sin necesidad de cultivo previo (18). Por su parte, la metatranscriptómica se centra en los perfiles de ARN mensajero, ofreciendo información sobre los genes que están activos y permitiendo inferir funciones metabólicas en un momento dado (16,19). La metaproteómica analiza las proteínas expresadas por la

comunidad, reflejando funciones y procesos en curso (16), mientras que la metabolómica se ocupa de los productos finales del metabolismo microbiano, brindando información sobre la actividad bioquímica global de la muestra (11,12).

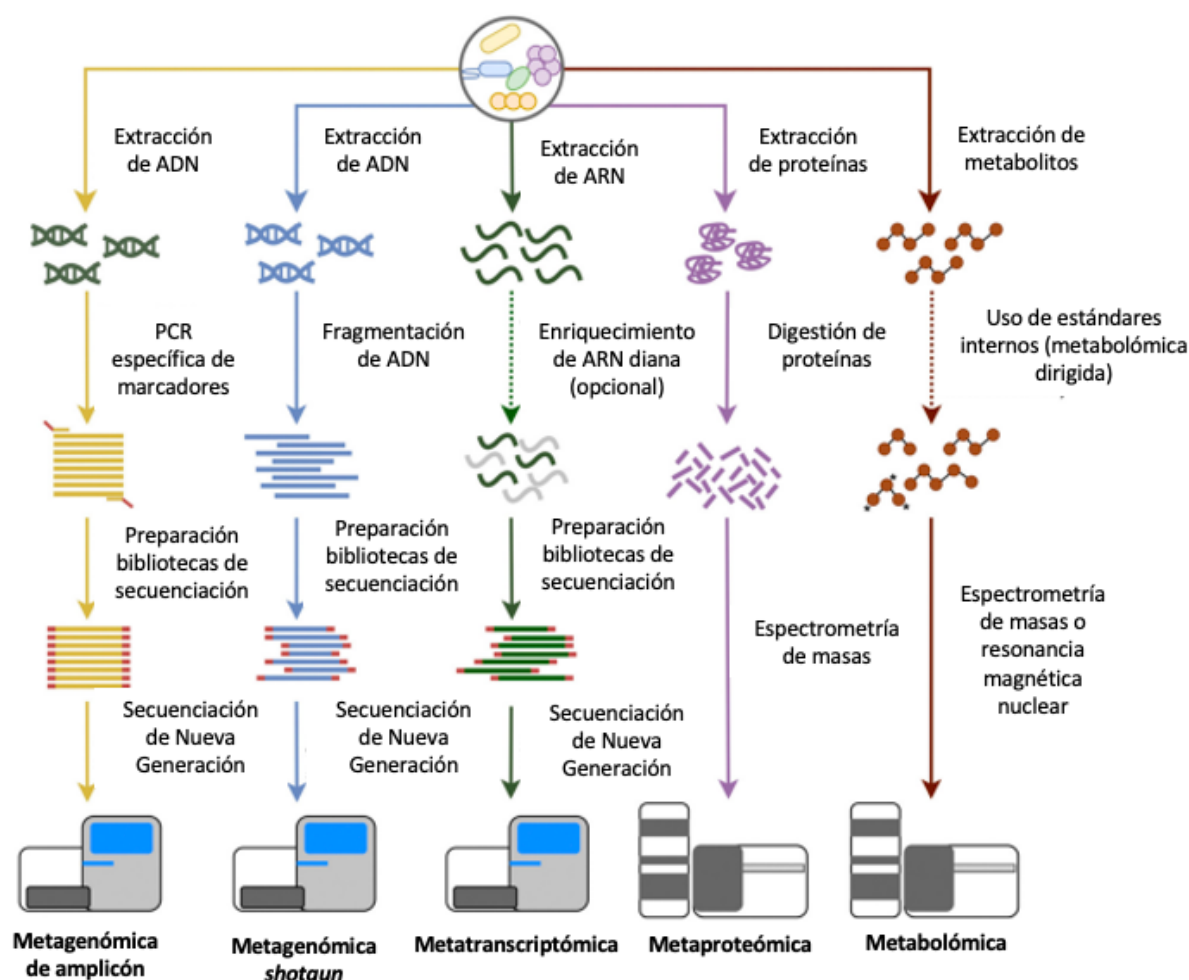


Figura 1. Principales tipos de análisis y flujos de trabajo ómicos. Adaptado de Arıkan et al., 2023 (16)

2.4.2. Metagenómica

El término metagenómica se refiere al estudio del material genético de los microorganismos presentes en una comunidad mediante secuenciación. Puede dividirse en dos enfoques principales: la metagenómica de amplicón y la metagenómica tipo *shotgun* (Tabla 3) (17,18,20). La metagenómica ha transformado el análisis de comunidades microbianas al

permitir la caracterización de su composición taxonómica y funciones metabólicas, incluso en entornos complejos como el intestino humano.

Tabla 3. Principales diferencias entre la metagenómica de amplicón y la metagenómica *shotgun*.

	Metagenómica de amplicón	Metagenómica <i>shotgun</i>
¿Cuál es su alcance?	Focalizado en un grupo de organismos (bacterias, hongos, etc.)	Análisis de todas las especies y genes presentes
¿Qué se secuencía?	Región específica (ARNr 16S, ITS, etc.)	Todo el ADN de una muestra
¿Qué se estudia?	Taxonomía	Taxonomía y funciones
Coste y complejidad	Menos costoso y complejo debido a que está focalizado en la región amplificada	Más costoso y complejo debido al mayor volumen de datos

Elaboración propia

2.4.2.1. Metagenómica de amplicón

La metagenómica de amplicón se centra en el estudio de la diversidad taxonómica (18). Esta técnica consiste en la amplificación, mediante la reacción en cadena de la polimerasa (PCR), de genes específicos altamente conservados que pueden incluir ARNr 16S, ARNr 18S e ITS, que permiten la identificación taxonómica de bacterias, organismos eucariotas y hongos, respectivamente (5,17,18).

Para la identificación taxonómica de bacterias se utiliza el gen ARNr 16S debido a su elevado grado de conservación y su tasa de mutación constante (5,21). En este gen se pueden encontrar nueve regiones hipervariables (V1 a V9), como muestra la Figura 2. Estas regiones, especialmente V1–V2, V3–V4 y V3–V5, se utilizan para la clasificación de bacterias a nivel de

familia, género o incluso de especie (5,17,18). Existen cebadores basados en la regiones conservadas de este gen que se utilizan para su amplificación y que se han optimizado para asegurar una alta eficiencia y especificidad (18).

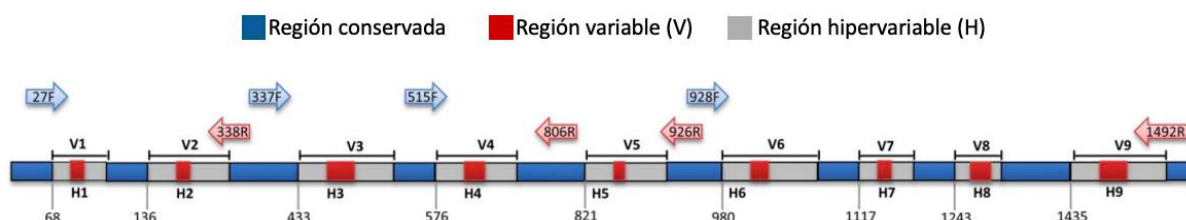


Figura 2. Regiones conservadas, variables e hipervariables del gen bacteriano ARNr 16S, así como los cebadores utilizados para su amplificación por PCR. Adaptado de Shahi et al., 2017 (22)

Esta técnica es muy útil para identificar y caracterizar comunidades microbianas en diversas muestras biológicas. Sin embargo, presenta limitaciones en términos de resolución taxonómica entre especies muy cercanas y en la detección de organismos que no se encuentran bien representados por los cebadores utilizados (5).

2.4.2.2. Metagenómica *shotgun*

La secuenciación metagenómica tipo *shotgun* representa una alternativa robusta para el estudio de comunidades microbianas, superando las limitaciones de los métodos dirigidos como la secuenciación del gen ARNr 16S (17). Este enfoque permite una caracterización integral de las comunidades microbianas mediante la secuenciación del contenido genético total presente en una muestra. A diferencia del enfoque basado en metagenómica de amplicón, la metagenómica *shotgun* permite identificar microorganismos con mayor resolución taxonómica, llegando hasta el nivel de especie, y ofrece además la posibilidad de analizar el potencial funcional de los genes presentes en la comunidad (5,18,23,24).

No obstante, este tipo de análisis conlleva una serie de desafíos técnicos y bioinformáticos que deben ser considerados. Uno de los principales retos es el gran volumen y la complejidad de los datos generados, lo que dificulta los procesos de ensamblado, anotación e identificación

precisa, especialmente en muestras con alta diversidad microbiana o con organismos poco representados en las bases de datos de referencia (23,24). A ello se suma el hecho de que, aunque los costes de secuenciación han disminuido en los últimos años, la metagenómica *shotgun* sigue siendo una técnica más costosa que otros enfoques como la secuenciación de amplicones (23,24).

2.4.2.3. Flujo de trabajo en metagenómica

En la Figura 3 se puede ver el flujo de trabajo en metagenómica de amplicón y metagenómica *shotgun*. El proceso comienza con la extracción del ADN total de la muestra, una etapa crítica que requiere especial atención para garantizar la integridad y representatividad del material genético. A continuación, se realiza la cuantificación del ADN extraído y un control de calidad para asegurarse de que sea adecuado tanto en términos de cantidad como de pureza para las etapas posteriores (17,20). A partir de este punto, el flujo de trabajo se bifurca en función del enfoque metodológico seleccionado.



Figura 3. Flujo de trabajo en metagenómica *shotgun* (izquierda) y metagenómica de amplicón (derecha). Elaboración propia

En metagenómica *shotgun*, el ADN genómico se fragmenta aleatoriamente y los fragmentos resultantes se preparan como librerías de secuenciación mediante la adición de adaptadores específicos. La secuenciación genera grandes volúmenes de datos complejos que se someten a un control de calidad riguroso utilizando herramientas como Trimmomatic o FastQC, para eliminar secuencias contaminantes, de baja calidad o procedentes del hospedador. Posteriormente, las secuencias de alta calidad se ensamblan mediante programas como MEGAHIT o metaSPAdes para reconstruir *contigs*, permitiendo la recuperación de genomas microbianos parciales o completos. Este proceso se puede complementar con la anotación funcional mediante herramientas como Prokka, para obtener información sobre las funciones metabólicas y ecológicas de los microorganismos detectados. Posteriormente, se realizan los análisis necesarios en función de los objetivos (17,20,25).

En el caso de la metagenómica de amplicón, se utiliza la PCR para amplificar regiones altamente conservadas de genes marcadores. Los productos de la PCR se preparan también en librerías y se someten a un control de calidad similar al de metagenómica *shotgun*. El análisis bioinformático posterior incluye la eliminación de secuencias de baja calidad, la inferencia de variantes de secuencia de amplicón (*Amplicon Sequence Variant*, ASVs) y la asignación taxonómica utilizando bases de datos como SILVA o Greengenes, dependiendo del tipo de muestra. Este flujo se gestiona habitualmente con plataformas integradas como QIIME 2, que también permiten realizar análisis de diversidad alfa y beta, evaluar la estructura comunitaria y generar visualizaciones interactivas (17,20,25).

2.4.2.4. Tecnologías de secuenciación

En metagenómica, las tecnologías de secuenciación más utilizadas actualmente son Illumina, Oxford Nanopore Technologies (ONT) y Pacific Biosciences (PacBio). Cada una se basa en principios distintos y ofrece ventajas específicas según el enfoque.

Illumina, basada en secuenciación por síntesis, incorpora nucleótidos fluorescentes y detecta la señal emitida ciclo a ciclo (26). Genera lecturas cortas de alta precisión (150–300 pb), con bajo coste por base y alto rendimiento, lo que la convierte en la plataforma predominante en ambos enfoques (25,26). En metagenómica de amplicón, sistemas como Illumina MiSeq permiten secuenciar regiones específicas con lecturas pareadas de hasta 300 pb. En

metagenómica *shotgun*, plataformas como Illumina NextSeq o NovaSeq generan cientos de millones de lecturas, ideales para estudios cuantitativos y ensamblaje de genomas microbianos (17).

Las tecnologías de lecturas largas como ONT y PacBio están ganando protagonismo debido a su capacidad para secuenciar fragmentos extensos del genoma sin necesidad de ensamblajes tan fragmentados (17). ONT permite obtener lecturas muy largas, incluso de más de 100 Kb, mediante la detección de cambios en la corriente eléctrica al pasar las moléculas por un nanoporo. Aunque su precisión ha sido tradicionalmente menor, los avances recientes la han mejorado significativamente (17,27). PacBio, mediante su tecnología SMRT, produce lecturas largas y precisas, de hasta 50 Kb y con una precisión superior al 99,9%. Esto la hace muy adecuada para analizar regiones genómicas complejas y obtener genomas completos (17,28).

3. Marco metodológico

3.1. Diseño del estudio

3.1.1. Tipo de estudio

Se trata de un estudio comparativo transversal basado en el análisis bioinformático de datos metagenómicos primarios, tanto de secuenciación *shotgun* como de amplicón del gen ARNr 16S, provenientes de muestras fecales de pacientes con SIBO y controles sin esta condición, obtenidas del repositorio público del American Gut Project (AGP).

3.1.2. Base de datos

El AGP, impulsado por la Universidad de San Diego, es el mayor estudio sobre el microbioma humano realizado hasta la fecha, que tiene como objetivo explorar la diversidad microbiana en la población humana a nivel global. Con un enfoque de acceso abierto, cuenta con más de 40.000 muestras de microbiomas humanos, que abarcan tanto metagenomas *shotgun*, como metagenomas de amplicón de ITS, ARNr 16S y 18S, que provienen de heces, piel y cavidad oral, entre otras. Además de los datos metagenómicos, el AGP recopila información detallada sobre el estilo de vida, hábitos alimenticios, tratamientos médicos y condiciones de salud de los participantes, lo que proporciona una valiosa información para el análisis de correlaciones entre la microbiota y factores ambientales, genéticos y clínicos.

3.1.3. Grupos de estudio

Este estudio se basa en dos grupos de participantes, diferenciados en función de su diagnóstico clínico de SIBO en base a los metadatos proporcionados por el AGP.

3.1.3.1. Grupo SIBO

Este grupo está compuesto por muestras de personas diagnosticados con SIBO por un profesional médico. Estos individuos están identificados en la base de datos del AGP con la etiqueta "Diagnosed by a medical professional (doctor, physician assistant)" en la casilla correspondiente a SIBO de los metadatos.

3.1.3.2. Grupo control

Para evaluar diferencias en el microbioma y resistoma intestinal entre individuos diagnóstico de SIBO y sujetos sin esta condición, fue necesario definir un grupo adecuado que permitiera minimizar sesgos.

Se partió de un grupo de individuos no diagnosticados con SIBO, identificados en la base de datos del AGP con la etiqueta “I do not have this condition” en la casilla de SIBO de los metadatos. Dado que existen diversas variables que pueden afectar significativamente tanto la composición del microbioma como el resistoma intestinal, se optó por un procedimiento de emparejamiento basado en covariables. Se aplicó un emparejamiento por puntuación de propensión (*Propensity Score Matching*) con el fin de reducir el sesgo por confusión y aumentar la comparabilidad entre grupos. Para ello, se utilizó la función `matchit()` del paquete `MatchIt` (versión 4.7.1) (29) de R. Se construyó un modelo logístico multivariable para estimar la probabilidad de pertenecer al grupo SIBO en función de un conjunto de variables clínicas, demográficas y relacionadas con el estilo de vida, que forman parte de los metadatos del AGP y que se muestran en la Tabla 4. A partir de esta puntuación de propensión, se aplicó un emparejamiento por el método del vecino más cercano (*nearest neighbor matching*) con una relación 1:1, es decir, cada individuo con SIBO fue emparejado con un único sujeto control. Se aplicó un *caliper* de 0.2 desviaciones estándar para restringir el emparejamiento y mejorar el balance entre grupos.

Este enfoque permitió obtener un subconjunto emparejado que incluía todas las muestras disponibles del grupo SIBO, cada una con su correspondiente control más similar según las covariables seleccionadas.

Tabla 4. Covariables utilizadas para la selección de controles mediante la función `matchit()`.

Categoría	Covariables
Demográficos	Categoría de edad, país de residencia, país de nacimiento, raza, sexo

Categoría	Covariables
Estilo de vida y entorno	Frecuencia de consumo de alcohol, frecuencia de ejercicio, frecuencia de consumo de tabaco, tipo de dieta, fuente de consumo de agua, consumo de alimentos fermentados
Salud física	Enfermedad autoinmune, categoría de índice de masa corporal, cáncer, infección por <i>Clostridioides difficile</i> , diabetes, crecimiento excesivo de hongos, EII, SII, embarazo
Suplementos dietéticos y medicamentos	Historial de uso de antibióticos, contracepción, frecuencia de probióticos

Elaboración propia

3.1.4. Tipos de muestras

En este estudio se han utilizado dos tipos de muestras metagenómicas: amplicón del ARNr 16S y *shotgun*.

Para realizar una comparación detallada de la composición y diversidad de la microbiota intestinal entre los grupos SIBO y Control, con el fin de identificar posibles diferencias asociadas al diagnóstico de SIBO, se han utilizado las muestras de metagenómica de amplicón del ARNr 16S.

Para el análisis del resistoma y explorar posibles diferencias en la abundancia de ARGs entre los dos grupos de estudio, se han empleado muestras de metagenómica *shotgun*.

3.2. Obtención de los datos

Los datos generados por el AGP están disponibles en el European Nucleotide Archive (ENA), bajo el número de proyecto PRJEB11419, y en la plataforma Qiita, con el ID 10317.

3.2.1. Lecturas metagenómicas

Las muestras analizadas en este trabajo fueron descargadas en formato .fastq.gz desde el repositorio del ENA del AGP.

Según informa el AGP, el ADN utilizado en los análisis de metagenómica de amplicón del gen ARNr 16S y metagenómica *shotgun* fue extraído a partir de 0,1 gramos de muestra fecal, mediante un protocolo estandarizado desarrollado para garantizar la reproducibilidad y comparabilidad entre muestras.

3.2.1.1. Muestras de metagenómica de amplicón del ARNr 16S

Las muestras de metagenómica de amplicón del gen ARNr 16S se obtuvieron mediante la amplificación por PCR de la región V4 del gen ARNr 16S, empleando los siguientes *primers* específicos:

- *Forward* (515F): GTGCCAGCMGCCGCGGTAA
- *Reverse* (806Rbc): GGACTACHVGGGTWTCTAAT

La construcción de las librerías de secuenciación se realizó a partir de estos productos amplificados. Posteriormente, la secuenciación se llevó a cabo utilizando plataformas Illumina MiSeq o HiSeq, con una longitud de lectura de 150 pb, en formato *single-end* o *paired-end*, según el lote y el momento del procesamiento.

3.2.1.2. Muestras de metagenómica *shotgun*

Las librerías de metagenómica *shotgun* fueron secuenciadas utilizando plataformas Illumina HiSeq 2500 o NovaSeq 6000, con una longitud de lectura de 150 pb en modo *paired-end* o *single-end*, según el lote y el momento del procesamiento.

3.2.2. Metadatos asociados

Se descargó la tabla completa de metadatos de todas las muestras disponibles desde la plataforma Qiita, lo que permitió acceder a la información necesaria para seleccionar las

muestras de interés y clasificarlas adecuadamente. La gestión de esta tabla de metadatos se realizó con RStudio (versión 2024.12.0).

3.3. Análisis bioinformático de la microbiota

Se partió de un total de 275 muestras de metagenómica de amplicón del ARNr 16S de heces de sujetos diagnosticados con SIBO y 275 de controles sin esta condición. El análisis bioinformático se realizó utilizando QIIME 2 (versión 2020.8) (30) en un entorno Conda (versión 25.3.1) y ejecutado en un equipo con sistema operativo macOS Monterey (versión 12.7.4) con procesador 1,8 GHz Intel Core i5 de doble núcleo. La visualización de los archivos en formato .qzv generados en QIIME 2 se realizó utilizando QIIME 2 View. Para todos los análisis estadísticos se consideró un nivel de significación de $p < 0,05$.

En el Anexo A se incluyen todos los comandos utilizados durante este proceso.

3.3.1. Preprocesamiento de las lecturas

3.3.1.1. Importación y control de calidad inicial

Las secuencias crudas fueron importadas a QIIME 2 a través de un archivo de tipo *manifest*. Se evaluó la calidad de las lecturas mediante visualizaciones que sirvieron para establecer los parámetros de truncamiento y filtrado por calidad utilizados en pasos posteriores.

Se eliminaron las secuencias de adaptadores utilizando *Cutadapt* y se aplicó un filtro de calidad, estableciendo un umbral PHRED mínimo de 30.

3.3.1.2. *Denoising* y construcción de la tabla de frecuencias

Para el proceso de *denoising* se utilizó DADA2, aplicando un truncamiento de las lecturas a 140 pares de bases para eliminar la parte terminal de menor calidad y se utilizaron 1.000.000 de lecturas para entrenar el modelo de error. Este proceso produjo una tabla de ASVs, un conjunto de secuencias representativas y un informe con las estadísticas del proceso de depuración.

Para asegurar una cobertura adecuada y evitar sesgos derivados de muestras con escasas lecturas, se excluyeron aquellas muestras que contenían menos de 5.000 lecturas en la tabla final. Este umbral se seleccionó buscando un equilibrio entre retención de muestras y calidad de los datos. Hay que tener en cuenta que se eliminaron conjuntamente las muestras pareadas de los grupos SIBO y Control.

3.3.2. Asignación taxonómica y análisis filogenético

La asignación taxonómica de las lecturas se realizó mediante el clasificador Naive Bayes entrenado con la base de datos Greengenes (versión 13_8, agrupamiento al 99%).

Para realizar análisis filogenéticos basados en la similitud evolutiva de las lecturas, se generó un árbol utilizando un alineamiento múltiple mediante transformación rápida de Fourier (MAFFT), enmascaramiento de regiones hipervariables, construcción del árbol filogenético con FastTree y enraizado mediante el método de raíz media (*midpoint-root*).

3.3.3. Análisis de diversidad y composición microbiana

3.3.3.1. Diversidad alfa

La diversidad alfa permite evaluar la complejidad ecológica del microbioma presente dentro de una misma muestra, e integra aspectos de riqueza (número de especies) y equidad (distribución de las abundancias) (31).

El análisis de diversidad alfa se realizó utilizando el árbol filogenético previamente construido y la tabla de frecuencias. Se calcularon los siguientes índices:

- **Índice de Shannon:** Mide la diversidad considerando tanto el número de especies como su distribución. Valores más altos reflejan mayor diversidad y uniformidad (31).
- **Número de ASVs:** Representa la riqueza taxonómica, es decir, el número de ASVs. Valores más altos implican mayor riqueza (31).
- **Índice de diversidad filogenética de Faith:** Mide la diversidad evolutiva basada en filogenia. Se expresa en unidades de longitud filogenética y valores más altos indican una comunidad microbiana filogenéticamente más diversa (31).

- **Índice de *evenness*:** Evalúa la uniformidad de distribución de las abundancias de las especies dentro de la muestra. Este índice varía entre 0 y 1, donde 1 representa una distribución perfectamente equitativa, mientras que una muestra dominada por pocas especies tendrá un índice de *evenness* bajo (31).

Las diferencias entre el grupo SIBO frente al Control se analizaron mediante la prueba no paramétrica de Kruskal-Wallis integrada en QIIME 2, lo que permitió determinar la significancia estadística de las diferencias observadas en la diversidad alfa.

3.3.3.2. Diversidad beta

La diversidad beta mide las diferencias en la composición microbiana entre muestras, es decir, cuán similares o distintas son las comunidades microbianas de diferentes grupos (31).

Para explorar las diferencias en la composición microbiana entre el grupo SIBO y Control, se calcularon varias métricas de diversidad beta a partir de la tabla de frecuencias y del árbol filogenético previamente construido:

- **Bray-Curtis:** Se basa en la abundancia relativa de taxones. Varía entre 0 (idénticas) y 1 (completamente diferentes) (31).
- **Jaccard:** Se basa en la presencia/ausencia de taxones, cuantificando la proporción de especies compartidas entre dos muestras. Varía entre 0 (idénticas) y 1 (completamente diferentes) (31).
- ***Unweighted* UniFrac:** Se basa en la presencia/ausencia de taxones incorporando información filogenética (31).
- ***Weighted* UniFrac:** Se basa en la abundancia relativa de taxones incorporando información filogenética (31).

Se generaron matrices de distancia y se aplicaron análisis de coordenadas principales (PCoA) para visualizar las relaciones entre muestras en un espacio bidimensional. Estas visualizaciones permitieron identificar patrones de agrupamiento entre los distintos grupos

de estudio. Las diferencias en la diversidad beta entre grupos fueron evaluadas mediante el análisis de varianza multivariante permutacional (PERMANOVA).

3.3.3.3. Composición taxonómica

Posteriormente, se generaron gráficos de barras interactivos que representan la distribución de los principales taxones en cada muestra. Estos gráficos facilitaron la exploración visual de posibles patrones de disbiosis o alteraciones microbianas asociadas al SIBO.

Para complementar este análisis descriptivo, se realizó un análisis de composiciones microbianas (ANCOM), tras la adición de pseudocontajes y el colapsado de la tabla de frecuencias al nivel taxonómico de filo, familia, género y especie.

3.4. Análisis bioinformático del resistoma

Se partió de un total de 50 muestras de metagenómica de amplicón del gen ARNr 16S de heces de individuos diagnosticados con SIBO y 50 del grupo Control. El análisis bioinformático para la detección y cuantificación de ARGs en muestras de metagenómica *shotgun* se realizó en un entorno Conda (versión 25.3.1) y el análisis estadístico con RStudio (versión 2024.12.0). Se utilizó un equipo con sistema operativo macOS Monterey (versión 12.7.4), equipado con un procesador Intel Core i5 de doble núcleo a 1,8 GHz. Para todos los análisis estadísticos se consideró un nivel de significación de $p < 0,05$.

En el Anexo B se incluyen todos los comandos utilizados durante este análisis.

3.4.1. Detección y cuantificación de ARGs

3.4.1.1. Filtrado y control de calidad de las lecturas

Los archivos de lecturas crudas en formato .fastq.gz fueron evaluados con FastQC (versión 0.12.1) (32) para revisar su calidad general e identificar adaptadores, entre otros.

A continuación, las lecturas crudas fueron procesadas con Trimmomatic (versión 0.39) (33), aplicando los siguientes criterios de filtrado:

- Se recortaron las bases de baja calidad en el extremo 3'.

- Se recortaron las bases de baja calidad en el extremo 5'.
- Recorte cuando el promedio de calidad PHRED en una ventana de 4 bases cae por debajo de 20.
- Se eliminaron lecturas con longitud inferior a 50 pb.

No se eliminaron adaptadores, ya que no estaban presentes.

Tras el filtrado, se realizó un nuevo control de calidad con FastQC sobre las lecturas filtradas para verificar la mejora en la calidad.

3.4.1.2. Ensamblaje metagenómico

Se empleó el ensamblador MEGAHIT (versión 1.2.9) (34) para el ensamblaje de los genomas microbianos a partir de las lecturas filtradas obtenidas con Trimmomatic. Se empleó la selección automática de tamaños de k-meros y se mantuvo la longitud por defecto de 200 pb, ya que el objetivo principal del estudio es la detección de ARGs, los cuales pueden estar fragmentados en *contigs* pequeños.

MEGAHIT generó un archivo en formato .fasta con los *contigs* ensamblados de cada muestra.

3.4.1.3. Evaluación de la calidad de los ensamblajes

Se evaluó la calidad de los ensamblajes mediante Quast (versión 5.3.0) (35), el cual proporciona métricas clave como el número de *contigs* y su tamaño.

3.4.1.4. Detección ARGs

La identificación de ARGs se realizó utilizando Abricate (versión 1.0.1) (36) mediante la comparación con la base de datos de Resfinder (37,38). Se configuró la búsqueda con los siguientes parámetros:

- Identidad mínima del 80% para permitir la detección de posibles variantes de genes conocidos sin comprometer la especificidad del análisis.

- Cobertura mínima del 30% para permitir la detección de genes de resistencia fragmentados en diferentes *contigs*.

3.4.1.5. Cálculo de la abundancia génica

Para cada muestra, las lecturas filtradas con Trimmomatic fueron alineadas contra el ensamblaje metagenómico generado por MEGAHIT utilizando Bowtie2 (versión 2.5.3) (39). El archivo de salida en formato .sam se convirtió a .bam y posteriormente a .bed, utilizando Samtools (versión 1.6) y Bedtools (versión 2.31.1).

Paralelamente, los archivos .tab generados por Abricate, que contienen los ARGs detectados, fueron convertidos a formato .bed. De esta manera, se pudo realizar un conteo del número de lecturas alineadas contra los ARGs detectados.

A partir de esta matriz de conteos brutos obtenida con Bowtie2, se utilizó el método de normalización lecturas por kilobase de lecturas en la muestra-por millón de lecturas mapeadas (RPKM). El valor RPKM que corrige por la longitud del gen y por el número total de lecturas secuenciadas por muestra. La fórmula general utilizada fue la siguiente:

$$RPKM = \frac{10^9 \times \text{Número de lecturas mapeadas al gen (pb)}}{\text{Número total de lecturas en la muestra} \times \text{Longitud del gen (pb)}}$$

El resultado fue una matriz de abundancia relativa en unidades de RPKM, con genes como filas y muestras como columnas. Además, se construyó también una matriz de abundancia relativa en función de la familia de antibióticos a las que confieren resistencia los ARGs.

3.4.2. Análisis estadísticos

3.4.2.1. Análisis diferencial de la abundancia relativa de ARGs entre el grupo SIBO y Control

Para visualizar y comparar la abundancia relativa de los ARGs entre las muestras de individuos con SIBO y controles, se utilizó la matriz de abundancias en unidades de RPKM con una transformación logarítmica para generar un *heatmap*, utilizando la función `pheatmap()` del paquete Pheatmap (40).

Se realizó un análisis estadístico para identificar que ARGs presentan diferencias significativas en su abundancia relativa entre grupos. Para ello, se seleccionaron únicamente aquellos genes presentes en al menos 10 muestras para evitar sesgos derivados de genes con muy baja prevalencia. Se evaluó la normalidad de la distribución de los genes mediante el test de Shapiro-Wilk y, como ninguno presentó una distribución normal, se aplicó la prueba no paramétrica de Wilcoxon-Mann-Whitney.

3.4.2.2. Análisis diferencial de la abundancia relativa de ARGs por familia de antibiótico entre el grupo SIBO y Control

Para visualizar la abundancia relativa de genes de resistencia agrupados por familia de antibiótico a los que confieren resistencia, se construyó *heatmap* utilizando la matriz de abundancia en RPKM con una transformación logarítmica, utilizando la función `pheatmap()` del paquete `Pheatmap` (40).

Se realizó un análisis estadístico para identificar familias de antibióticos que presentan diferencias significativas en la abundancia relativa de ARGs entre individuos con SIBO y controles. Se evaluó la normalidad de la distribución de los grupos de antibióticos mediante el test de Shapiro-Wilk y, como ninguno presentó una distribución normal, se aplicó la prueba no paramétrica de Wilcoxon-Mann-Whitney.

4. Resultados

4.1. Análisis bioinformático de la microbiota

4.1.1. Preprocesamiento de las lecturas

La evaluación de la calidad de las secuencias reveló una calidad aceptable en la mayoría de las posiciones de lectura, aunque con una caída progresiva a partir del ciclo 140, por lo que se optó por truncar las secuencias en la posición 140 y eliminar las bases con un índice PHRED menor de 30.

El proceso de *denoising* mediante DADA2 permitió la detección y corrección de errores de secuenciación, así como la eliminación de quimeras. Se eliminaron 44 muestras por presentar un recuento de lecturas inferior al umbral de 5.000, considerado insuficiente para una caracterización fiable del microbioma intestinal.

Se mantuvieron un total de 506 muestras, con 8.176 ASVs detectadas. La frecuencia total fue de 14.699.522 lecturas. La distribución de frecuencias por muestra mostró una mediana de 20.074 lecturas, con un rango que fue desde 5.339 hasta 498.490 lecturas por muestra. La frecuencia media fue de aproximadamente 29.050 lecturas por muestra, lo cual indica una buena profundidad de secuenciación, aunque con cierta dispersión (Q1 = 15.068; Q3 = 28.679).

4.1.2. Análisis de diversidad y composición microbiana

4.1.2.1. Diversidad alfa

Los valores del índice de Shannon oscilaron aproximadamente entre 0,5 y 6,5 en ambos grupos, con una mediana ligeramente superior en el grupo Control (Figura 4), de aproximadamente 4,8. Sin embargo, el análisis estadístico mediante la prueba de Kruskal-Wallis no reveló diferencias significativas entre grupos ($H = 2,33$; $p = 0,127$). Este resultado sugiere que ambos grupos presentan una diversidad global interna comparable en términos de número y distribución relativa de taxones.

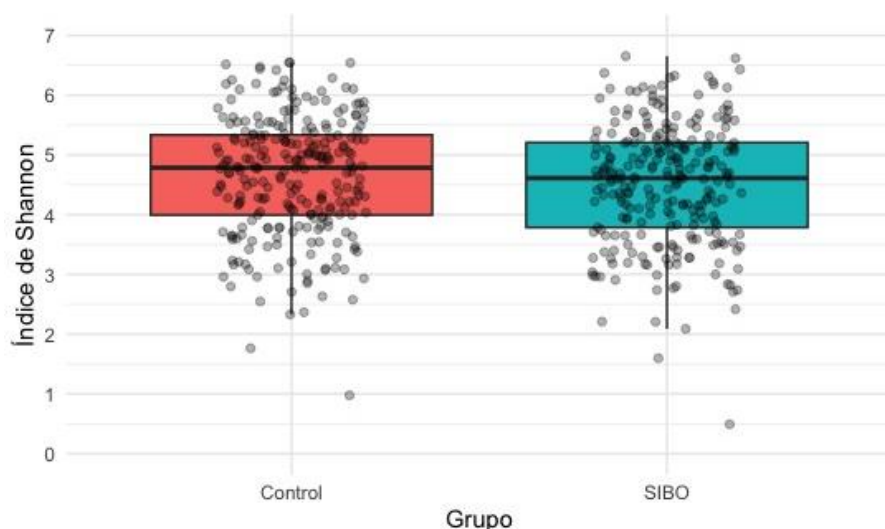


Figura 4. Comparación del índice de Shannon entre los grupos con y sin diagnóstico de SIBO.

Elaboración propia

Los valores correspondientes al número de ASVs variaron ampliamente en ambos grupos, oscilando entre 15 a más de 300 (Figura 5). La mediana fue ligeramente superior en el grupo sin Control, con 125 ASVs, aunque las diferencias entre grupos no fueron estadísticamente significativas según la prueba de Kruskal-Wallis ($H = 0,06$; $p = 0,805$). Este resultado indica que ambos grupos muestran una riqueza comparable en cuanto al número total de taxones detectados.

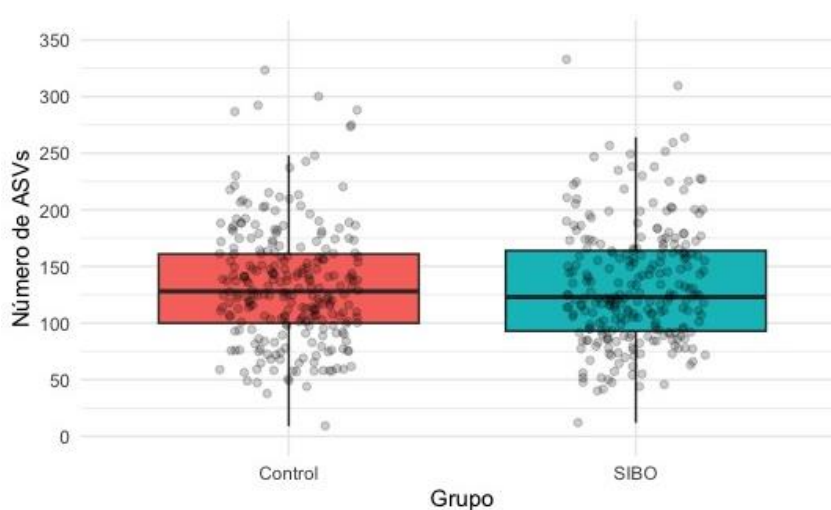


Figura 5. Comparación del número de características observadas entre los grupos con y sin diagnóstico de SIBO. Elaboración propia

Los valores del índice de diversidad filogenética de Faith mostraron una variación considerable en ambos grupos (Figura 6), con valores que oscilaron entre 2 y 65. La mediana fue similar en ambos grupos, con un valor de aproximadamente 15, por lo que las diferencias entre grupos no fueron estadísticamente significativas según la prueba de Kruskal-Wallis ($H = 0,06$; $p = 0,809$). Este resultado sugiere que no existen diferencias relevantes en la diversidad filogenética global entre las personas con y sin diagnóstico de SIBO.

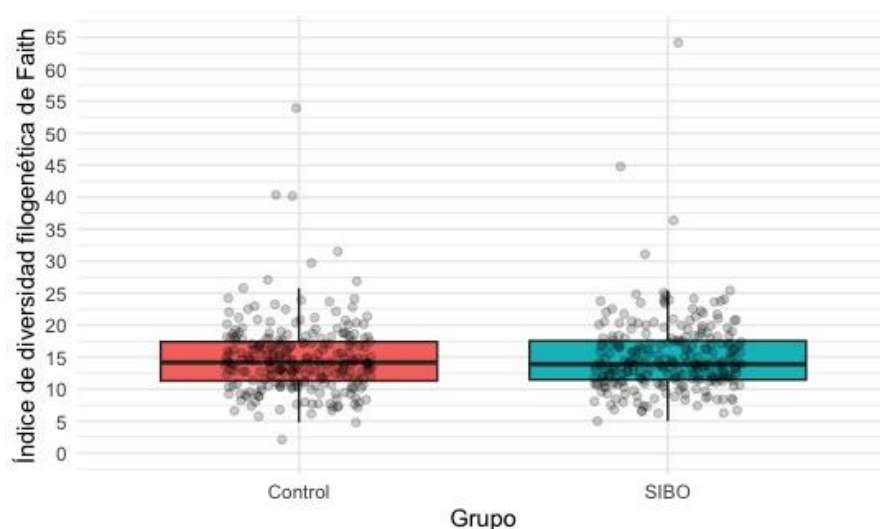


Figura 6. Comparación del índice de Faith entre los grupos con y sin diagnóstico de SIBO.

Elaboración propia

Los valores del índice de *evenness* mostraron diferencias más marcadas entre los grupos analizados (Figura 7). La mediana fue ligeramente mayor en el grupo Control (0,70) en comparación con el grupo diagnosticado con SIBO (0,67), lo que indica una distribución más uniforme de los taxones en el grupo Control. La prueba de Kruskal-Wallis reveló diferencias estadísticamente significativas entre ambos grupos ($H = 5,23$; $p = 0,022$), lo que sugiere que la microbiota intestinal de los individuos con SIBO tiende a estar menos equilibrada, con ciertos taxones potencialmente más dominantes que otros.

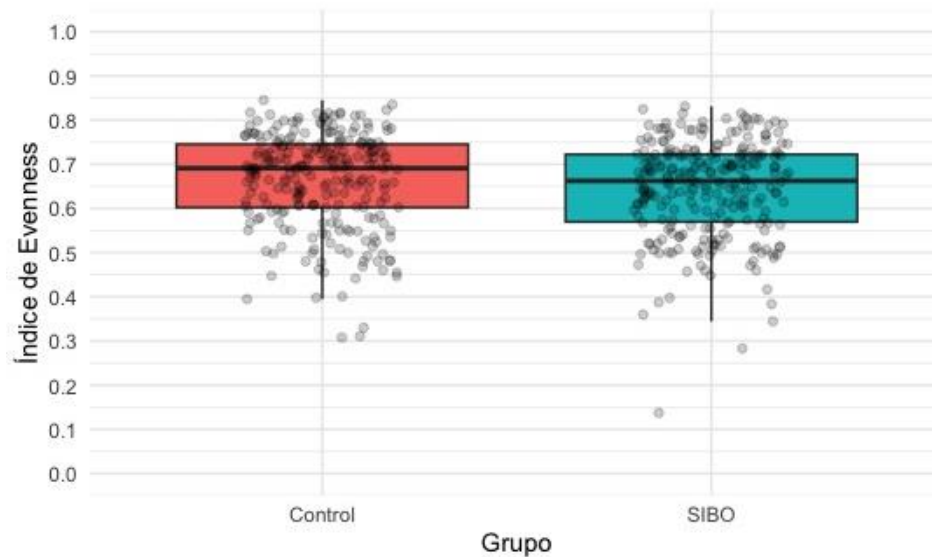


Figura 7. Comparación del índice de evenness entre los grupos con y sin diagnóstico de SIBO.

Elaboración propia

4.1.2.2. Diversidad beta

En la Figura 8 se muestran los PCoA de las diferentes métricas de distancia entre los grupos SIBO y Control. Se observa una superposición considerable entre grupos en todos los gráficos y la métrica *weighted* UniFrac presenta una nube de puntos altamente solapada.

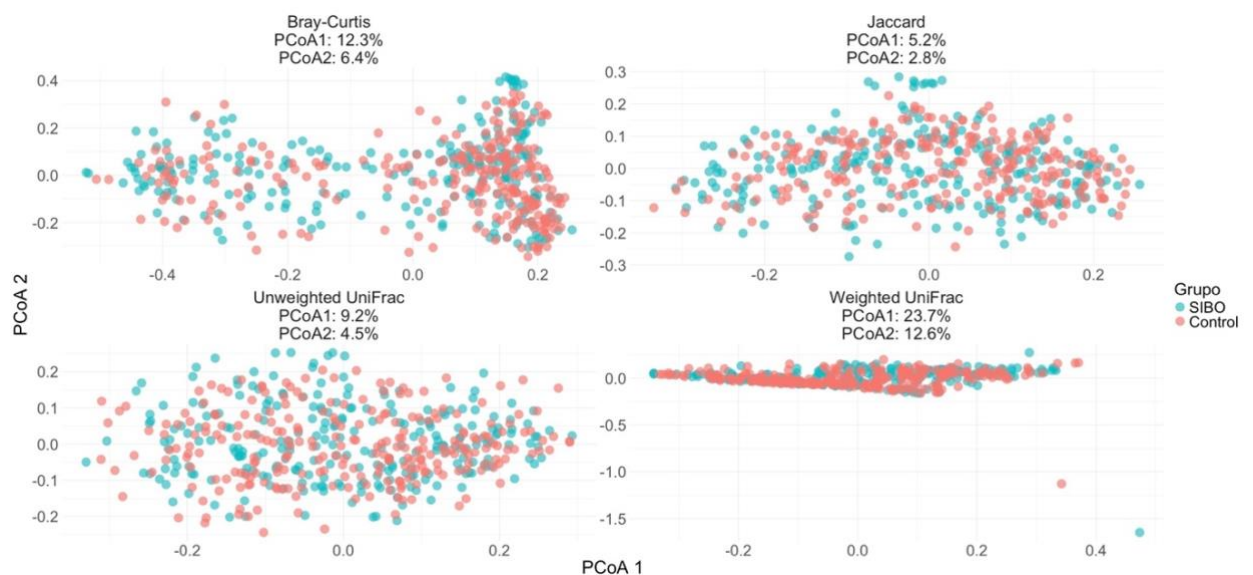


Figura 8. PCoA de las distancias de Bray-Curtis, Jaccard, *weighted* UniFrac y *unweighted* UniFrac de los grupos con y sin diagnóstico de SIBO. Elaboración propia

La métrica Bray-Curtis reveló diferencias significativas entre grupos (pseudo-F = 2,554; $p = 0,002$), lo que sugiere alteraciones cuantitativas en la composición microbiana asociadas al SIBO. De forma similar, el índice Jaccard también mostró diferencias significativas (pseudo-F = 1,471; $p = 0,003$), indicando que la composición cualitativa de las comunidades microbianas difiere entre grupos.

Las métricas filogenéticas UniFrac reflejaron resultados consistentes. En particular, *unweighted* UniFrac mostró diferencias significativas (pseudo-F = 1,430; $p = 0,029$), lo que indica una divergencia filogenética en la presencia de taxones. En contraste, *weighted* UniFrac (que considera abundancia y filogenia) no alcanzó significación estadística (pseudo-F = 1,859; $p = 0,062$), aunque mostró una tendencia hacia la diferenciación.

En conjunto, estos resultados indican que existen diferencias significativas en la estructura de las comunidades microbianas intestinales entre los grupos, tanto a nivel de abundancia como de presencia/ausencia, y que dichas diferencias tienen un componente filogenético.

4.1.2.3. Análisis diferencial de taxones

Los filos predominantes, con una abundancia relativa promedio superior al 1%, son los que aparecen en la Tabla 5.

Tabla 5. Abundancia relativa promedio de los filos predominantes en los grupos SIBO y Control.

Filo	Grupo SIBO	Grupo Control
Bacillota	38.76%	37.42%
Bacteroidota	34.86%	38.73%
Pseudomonadota	19.98%	17.72%
Actinomycetota	1.21%	1.14%
Verrucomicrobiota	1.17%	1.05%

Elaboración propia

Los resultados del ANCOM revelaron que a nivel de especie y de género no se detectaron diferencias significativas entre los grupos SIBO y Control.

A nivel de familia, se identificó que Enterobacteriaceae, con un valor de $W = 144$, presentó diferencias consistentes en su abundancia relativa (Figura 9). El valor W del estadístico de ANCOM refleja la robustez del hallazgo, indicando el número de comparaciones en las que un determinado taxón fue significativamente diferente respecto al resto. Aunque los valores máximos son similares entre ambos grupos (incluso ligeramente superiores en el grupo Control), la mediana de abundancia de Enterobacteriaceae es considerablemente mayor en el grupo con SIBO (1378 vs. 267). Esto indica que, en promedio, los individuos diagnosticados con SIBO tienden a tener una proliferación más elevada de esta familia bacteriana.

A nivel de filo, Pseudomonadota mostró diferencias consistentes en su abundancia relativa entre los grupos estudiados (Figura 9). El valor de W ($W=16$) es mucho menor que el observado para Enterobacteriaceae a nivel familiar, lo que sugiere una evidencia menos robusta. La mediana de abundancia de Proteobacteria en el grupo SIBO es de 3036, mientras que en el grupo Control la mediana es menor, con un valor de 1059. Esto indica que, en promedio, los individuos con SIBO tienden a presentar una mayor abundancia relativa de este filo bacteriano.

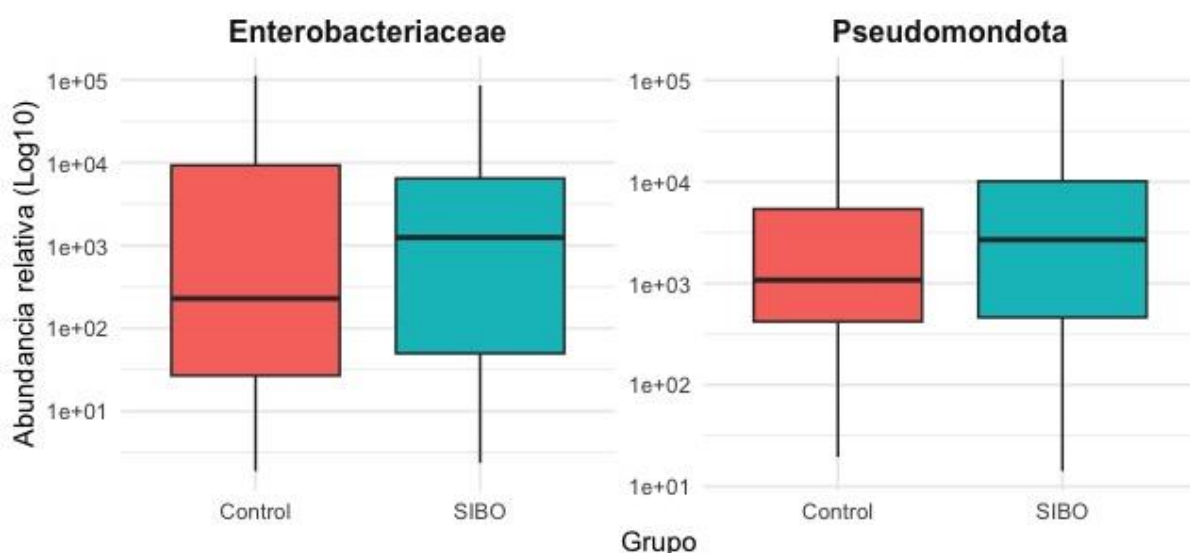


Figura 9. Distribución de la abundancia relativa de la familia Enterobacteriaceae y del filo Pseudomonadota en el grupo SIBO y Control. Elaboración propia

4.2. Análisis bioinformático del resistoma

4.2.1. Detección y cuantificación de ARGs

4.2.1.1. Evaluación del filtrado de las lecturas

Antes del filtrado, se observó una amplia variabilidad en la cantidad de lecturas entre las distintas muestras, con valores que oscilaron desde aproximadamente 169.000 hasta más de 16 millones. Tras aplicar el control de calidad con Trimmomatic, se registró una reducción en el número de lecturas que varió dependiendo del volumen de datos inicial en cada muestra. En todos los casos se mantuvieron más del 95% de las lecturas, por lo que todas ellas retuvieron un número suficiente de lecturas que permitieron continuar con los análisis posteriores.

4.2.1.2. Evaluación de calidad de los ensamblajes

En el análisis de la calidad de los ensamblajes metagenómicos de las diversas muestras realizados con Quast se utilizaron varias métricas que permitieron evaluar su integridad y calidad general.

Se evidenció una considerable heterogeneidad entre las muestras, reflejando diferencias tanto en la complejidad de las comunidades microbianas como en la eficiencia del proceso de ensamblaje.

El número total de *contigs* con longitud mayor a 0 bp varió ampliamente entre ensamblajes, alcanzando en algunos casos más de 160.000. En relación con los *contigs* de al menos 1.000 pb también se observó una dispersión significativa, con ensamblajes que contenían desde poco más de 1.000 hasta más de 40.000.

La longitud total de los ensamblajes también mostró una gran variabilidad, oscilando entre menos de 30 Mb a más de 240 Mb, lo cual sugiere diferencias sustanciales en la cantidad y complejidad del material genético presente en las muestras. El valor de N50, una métrica que refleja la continuidad del ensamblaje, presentó una distribución igualmente dispersa, con valores inferiores a 10 kb en ensamblajes más fragmentados y superiores a 50 kb en ensamblajes más completos. El tamaño del *contig* más largo en cada muestra alcanzó en

algunos casos longitudes superiores a 800 kb, mientras que en otras se mantuvo en rangos mucho menores.

En cuanto al contenido de GC, este se mantuvo estable entre muestras, con valores comprendidos entre el 44% y el 49%.

4.2.1.3. Detección ARGs

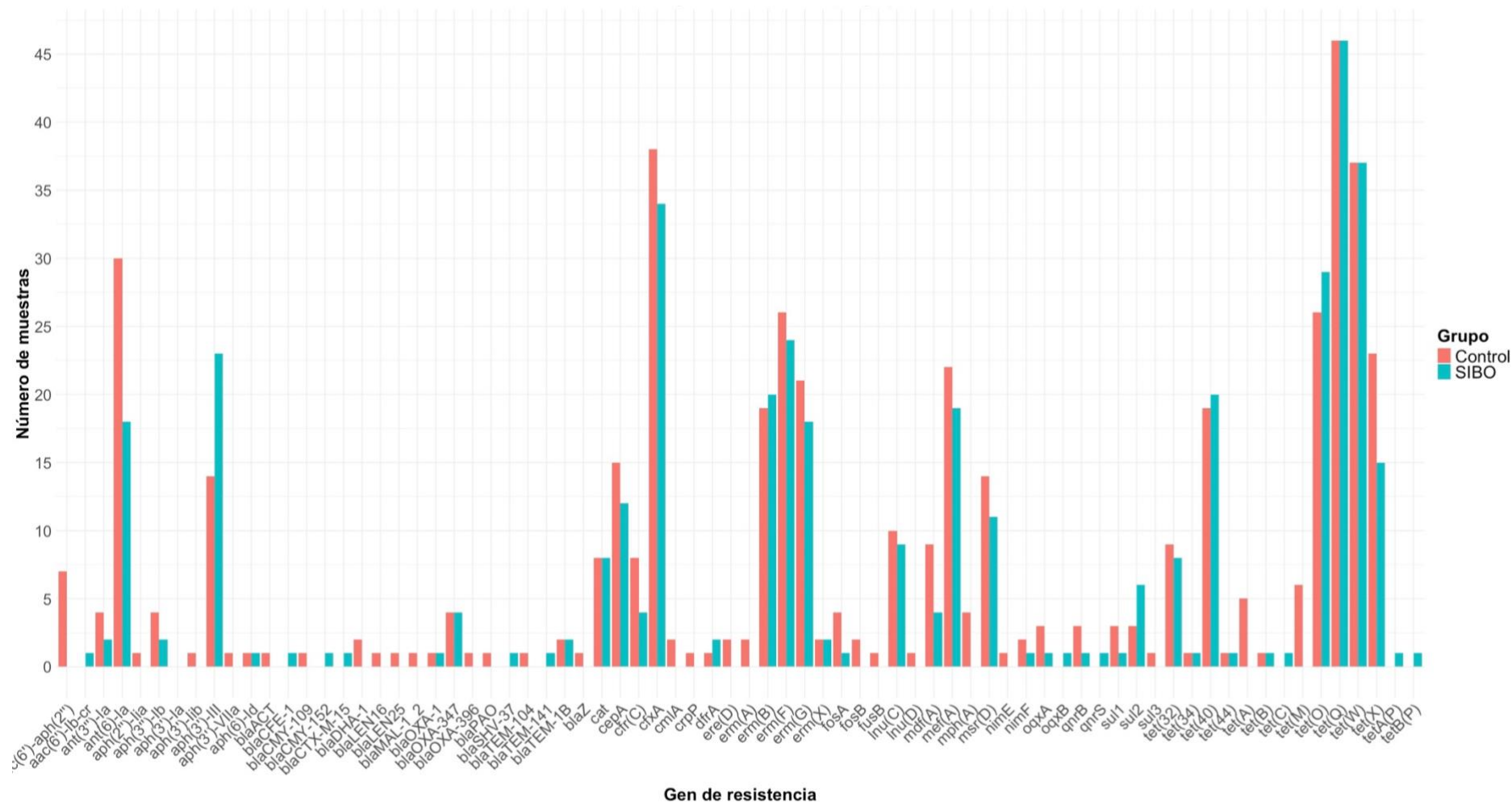
A partir de los ensamblajes obtenidos de las muestras metagenómicas de los individuos del grupo SIBO y Control, se identificaron de 74 ARGs que confieren resistencia a diferentes familias farmacológicas, incluyendo tetraciclinas, macrólidos, β -lactámicos, aminoglucósidos y quinolonas, entre otros. En la Figura 10 se representa el número de muestras en las que se detectaron los ARGs para los grupos SIBO y Control.

En el caso de los genes de resistencia a aminoglucósidos, destacan el gen *aph(3')-III*, con mayor frecuencia en el grupo SIBO, y *ant(6)-Ia*, más prevalente en el grupo Control. También se observa que el gen *aac(6')-aph(2'')* aparece exclusivamente en siete muestras del grupo Control.

En relación con los betalactámicos, aunque en menos de cinco muestras cada uno, se identificaron múltiples genes clínicamente relevantes como *blaDHA-1*, *blaCMY-109*, *blaCMY-152*, *blaTEM-1B*, *blaCTX-M-15*, *blaOXA-1* y *blaSHV-37*. Es especialmente prevalente la betalactamasa *cfxA*, presente en 38 y 34 muestras de los grupos Control y SIBO, respectivamente, que confiere resistencia a penicilinas y cefalosporinas.

Respecto a la resistencia a macrólidos, lincosamidas y/o estreptograminas (MLS) se identificó una alta prevalencia de los genes *erm(B)*, *erm(F)*, *erm(G)*, *mef(A)* y *msr(D)*, todos ellos con presencia comparable en ambos grupos pero, en general, mayor en el grupo Control.

La diversidad y abundancia de genes relacionados con la resistencia a tetraciclinas (variantes *tet*) fue la más alta en ambos grupos, siendo genes como *tet(0)*, *tet(Q)*, *tet(W)*, *tet(X)* y *tet(40)* especialmente prevalentes en ambos, llegando el gen *tet(Q)* y el *tet(W)* a detectarse en 46 y 37 muestras de cada grupo, respectivamente.



La resistencia a quinolonas mediada por ARGs (*aac(6')*-Ib-cr, *qnrB* y *qnrS*) presenta una baja prevalencia en ambos grupos, al igual que ocurre con los genes relacionados con resistencia a sulfonamidas *sul1*, *sul2* y *sul3*, y con los genes *fosA* y *fosB*, vinculados a la resistencia a fosfomicina.

Es importante destacar la detección de genes como *cfr(C)*, *oqxA* y *oqxB*, que confiere un fenotipo de multiresistencia (MDR), que implica la resistencia a más de tres familias de antibióticos.

Si se agrupan los ARGs en función de las familias de antibióticos a los que confieren resistencia (Figura 11), los dos grupos presentan patrones muy similares. En particular, la resistencia a aminoglucósidos, tetraciclinas, β -lactámicos y MLS fue alta en los pacientes con SIBO y controles sin esta condición. No obstante, aunque con muy baja prevalencia, se observaron algunas diferencias en la distribución de la presencia de genes entre grupos, como la resistencia a ácido fusídico, que se detecta exclusivamente en el grupo Control, o ARGs que confieren resistencia combinada a aminoglucósidos y quinolonas, que se detectan únicamente en individuos con SIBO.

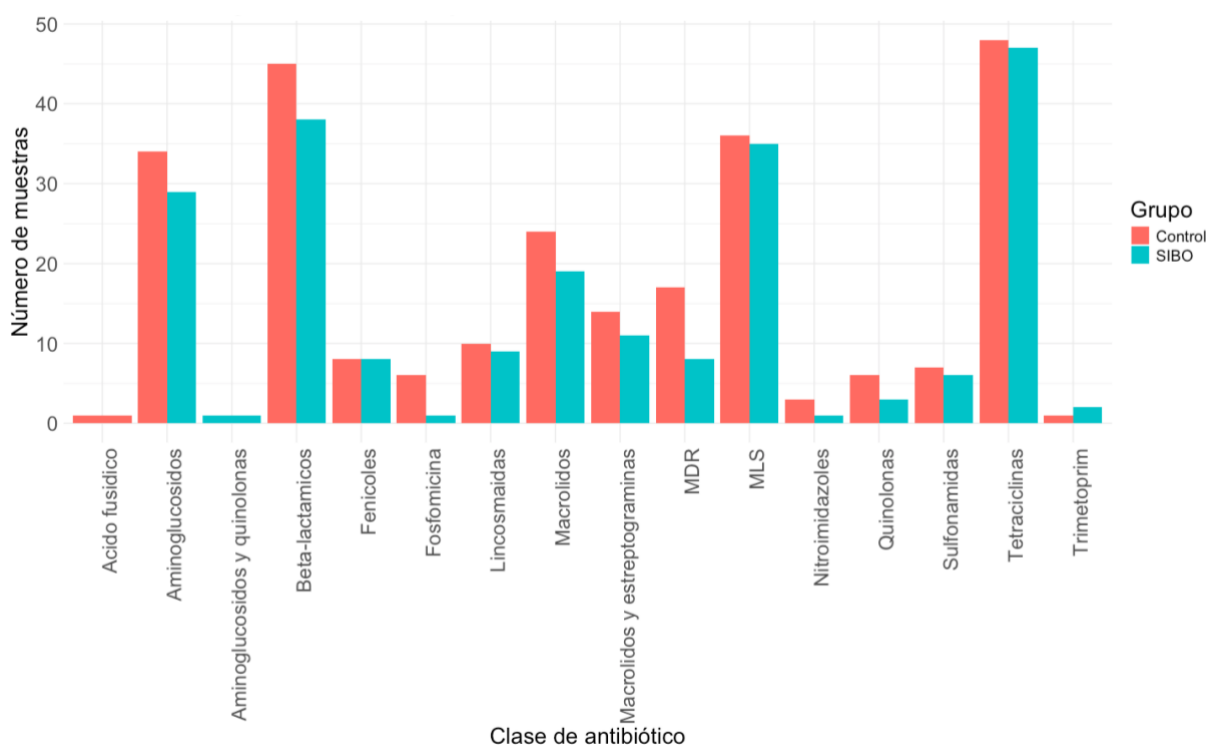


Figura 11. Prevalencia de ARGs por clase. Elaboración propia

4.2.2. Análisis estadísticos

4.2.2.1. Análisis diferencial de la abundancia relativa de ARGs entre el grupo SIBO y Control

En la Figura 12 se muestra el *heatmap* la distribución de la abundancia relativa de los ARGs en los dos grupos. Los resultados de la prueba de Wilcoxon-Mann-Whitney para evaluar si existían diferencias significativas en la abundancia de los ARGs entre los grupos SIBO y Control se muestran en la Tabla 6. Para el único gen que se encontraron diferencias significativas fue el gen *ant(6)-Ia*, que mostró una disminución en su abundancia media en el grupo SIBO ($\log_{FC} = -0,87$; $p = 0,027$).

Tabla 6. Estadísticos de la prueba de Wilcoxon de la abundancia relativa de los ARGs entre grupos.

Gen	logFC	p-valor	Gen	logFC	p-valor
<i>ant(6)-Ia</i>	-0,866	0,026	<i>mdf(A)</i>	-0,502	0,151
<i>aph(3')-III</i>	0,856	0,060	<i>mef(A)</i>	-0,589	0,335
<i>cat</i>	-0,062	0,978	<i>msr(D)</i>	-0,218	0,627
<i>cepA</i>	-0,153	0,551	<i>tet(32)</i>	0,015	0,862
<i>cfr(C)</i>	-0,372	0,219	<i>tet(40)</i>	0,285	0,585
<i>cfxA</i>	-0,309	0,745	<i>tet(O)</i>	0,217	0,630
<i>erm(B)</i>	0,297	0,674	<i>tet(Q)</i>	-0,496	0,210
<i>erm(F)</i>	-0,458	0,454	<i>tet(W)</i>	-0,036	0,994
<i>erm(G)</i>	-0,330	0,572	<i>tet(X)</i>	-0,675	0,112
<i>lnu(C)</i>	-0,051	0,852			

Elaboración propia

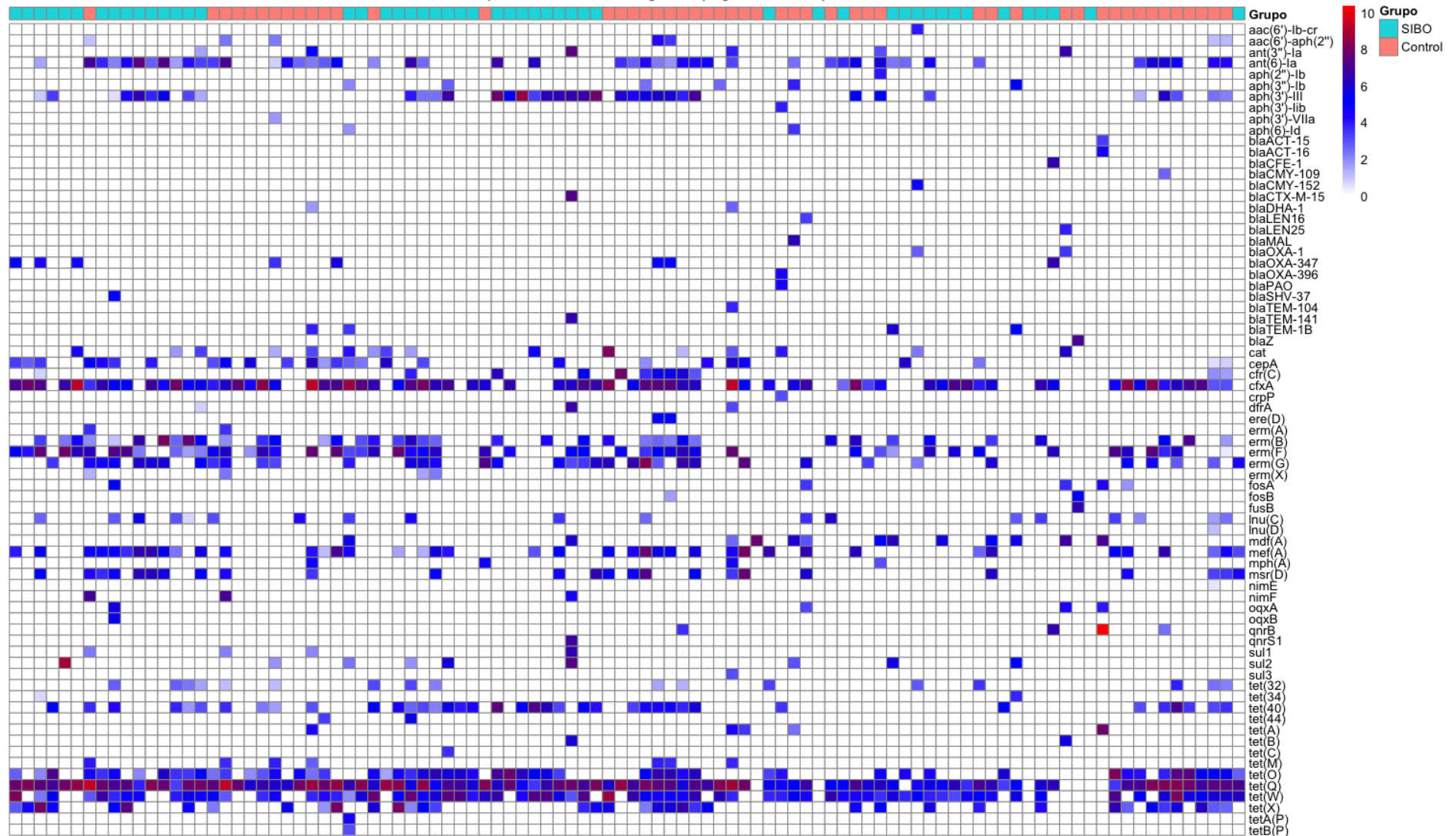


Figura 12. Abundancia relativa (log RPKM) de los diferentes ARGs de los grupos SIBO y Control. Elaboración propia

4.2.2.2. Análisis diferencial de la abundancia relativa de ARGs por familia de antibiótico entre el grupo SIBO y Control

En la Figura 13 se muestra el *heatmap* sobre distribución de la abundancia relativa de genes de resistencia agrupados por familia de antibióticos en las muestras de individuos con SIBO y controles sin esta condición. Los resultados de la prueba de Wilcoxon-Mann-Whitney para evaluar si existían diferencias significativas en la abundancia de los ARGs entre grupos se muestran en la Tabla 7. El análisis estadístico determinó que no existen diferencias significativas para ninguna familia de antibióticos. Las categorías MDR y fosfomicina, siendo ambas abundancias mayores en el grupo Control, fueron las que mostraron un menor p-valor, aunque no alcanzaron significación estadística.

Tabla 7. Estadísticos de la prueba de Wilcoxon de la abundancia relativa de los ARGs agrupados por familia de antibiótico entre grupos.

Clase de Antibiótico	logFC	p-valor
MDR	-0,834	0,057
Fosfomicina	-0,838	0,058
Macrólidos	-0,517	0,163
Tetraciclinas	-0,376	0,343
MLS	0,461	0,453
Aminoglucósidos	-0,426	0,430
Quinolonas	-0,531	0,341
β -lactámicos	-0,761	0,701
Macrólidos y estreptograminas	-0,216	0,627
Lincosamidas	-0,052	0,852
Sulfonamidas	0,058	0,882
Fenicoles	-0,034	0,978

Elaboración propia

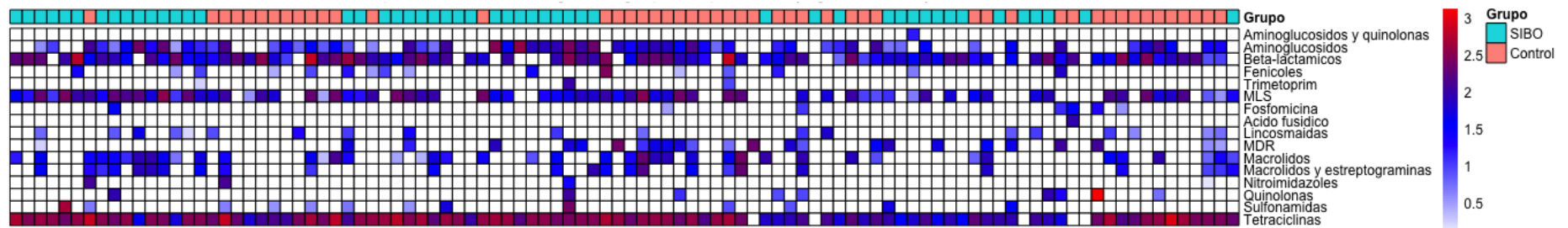


Figura 13. Abundancia relativa (log RPKM) de los diferentes ARGs agrupada por familia de antibióticos a los que confieren resistencia de los grupos SIBO y Control. Elaboración propia

5. Discusión

En el presente trabajo se realizó un análisis comparativo de la microbiota y del resistoma fecal entre individuos con y sin diagnóstico de SIBO, utilizando datos metagenómicos públicos del AGP. Mediante el análisis de datos de secuenciación de amplicón del gen ARNr 16S y de metagenómica tipo *shotgun*, se caracterizó tanto la diversidad microbiana como la presencia de ARGs.

5.1. Análisis de la microbiota

5.1.1. Diversidad alfa y beta

La diversidad de la microbiota fecal se evaluó desde dos enfoques complementarios: diversidad alfa, que describe la complejidad ecológica dentro de cada muestra, y diversidad beta, que permite comparar la composición microbiana entre grupos.

En este estudio, se aplicaron varios índices para capturar distintas dimensiones la diversidad alfa: Shannon, número de ASVs, índice de Faith y *evenness*.

No se observaron diferencias significativas en el índice de Shannon, que integra riqueza y equidad de la comunidad microbiana, aunque la mediana fue ligeramente mayor en el grupo Control. De igual modo, el número de ASVs, indicador de riqueza taxonómica, no mostró diferencias entre grupos, sugiriendo que la cantidad total de taxones se mantiene. Esto contrasta con estudios que reportan disminución en diversidad microbiana en sujetos con SIBO (41,42).

El índice filogenético de Faith, que incorpora la distancia evolutiva entre taxones, tampoco presentó diferencias significativas, indicando que la diversidad evolutiva se mantiene estable. Esto contrasta con otras disbiosis, como en la EII o la enfermedad de Crohn, donde se ha documentado una reducción de este índice (43,44).

El índice de *evenness* fue el único que mostró diferencias estadísticamente significativas, siendo menor en pacientes con SIBO. Esto indica que la distribución de las especies es menos equitativa en estos individuos, lo que implica una dominancia excesiva de ciertos taxones, fenómeno que se ha descrito previamente en estudios donde bacterias oportunistas proliferan en la microbiota intestinal de individuos, desplazando otras poblaciones

comensales (45). Esta menor equidad sugiere una disbiosis funcional caracterizada por pérdida de nichos ecológicos y un debilitamiento del control competitivo entre especies (46).

Tras analizar la diversidad alfa, se exploró la diversidad beta con el objetivo de evaluar posibles diferencias en la composición global del microbioma entre los grupos estudiados. Se emplearon índices basados en la distancia de Bray-Curtis, Jaccard, *weighted* UniFrac y *unweighted* UniFrac.

Los resultados mostraron diferencias significativas en las métricas Bray-Curtis y Jaccard, indicando que tanto la abundancia como la presencia o ausencia de ciertos taxones varían entre individuos con SIBO y controles sin esta condición. Esto sugiere que el microbioma en pacientes con SIBO no solo difiere en qué especies están presentes, sino también en sus proporciones relativas, lo que implica una reestructuración profunda del ecosistema microbiano, como ha sido reportado previamente en estudios metagenómicos (9).

Asimismo, el análisis con *unweighted* UniFrac, que considera la distancia evolutiva entre los taxones presentes sin ponderar su abundancia, también mostró diferencias significativas. Esto revela que los grupos se diferencian en la presencia de linajes bacterianos filogenéticamente distintos, apuntando a una alteración en la composición microbiana no solo a nivel taxonómico, sino también evolutivo. Este tipo de cambio podría reflejar la aparición o desaparición de grupos bacterianos específicos vinculados con el estado patológico.

Por otro lado, la métrica *weighted* UniFrac, que pondera las diferencias filogenéticas según la abundancia de cada taxón, no alcanzó significación estadística. Este hallazgo indica que aunque existen diferencias en los taxones presentes entre los grupos, estas no están acompañadas por cambios drásticos en su abundancia relativa. En otras palabras, las especies más abundantes tienden a mantenerse relativamente constantes, mientras que las diferencias se encuentran en taxones menos dominantes.

En resumen, el microbioma intestinal en individuos con SIBO muestra una menor equidad y una composición microbiana distinta, pero sin pérdida generalizada de riqueza ni diversidad filogenética, alineándose con estudios que sugieren una menor estabilidad ecológica y resiliencia del microbioma frente a perturbaciones externas como infecciones o tratamientos antibióticos, asociadas a esta condición (47,48).

5.1.2. Composición taxonómica

En consonancia con lo reportado en la literatura, las bacterias predominantes en ambos grupos pertenecen a los filos Bacillota, Bacteroidota, Actinomycetota y Pseudomonadota, que en conjunto constituyen la mayor parte de la microbiota intestinal humana (1,4,5).

El análisis taxonómico de las muestras fecales evidenció diferencias significativas entre los grupos tanto a nivel de filo como de familia bacteriana. En particular, el grupo de individuos con SIBO presentó una mayor abundancia relativa de bacterias del filo Pseudomonadota y de la familia Enterobacteriaceae, lo cual es consistente con un patrón de disbiosis caracterizado por la expansión de taxones oportunistas y potencialmente patógenos. Este hallazgo ha sido descrito previamente en estudios clínicos y metagenómicos, donde se ha documentado una colonización excesiva del intestino delgado por bacterias aerobias o anaerobias facultativas, muchas de ellas pertenecientes a Enterobacteriaceae, como *Escherichia coli* y *Klebsiella* spp. (49,50).

En la literatura consultada (7,41), la mayor proporción de bacterias del filo Pseudomonadota en individuos con SIBO se ha asociado con una reducción relativa del filo Bacillota, típicamente relacionado con un estado de eubiosis. En este estudio, aunque las diferencias no fueron estadísticamente significativas, se observó una menor abundancia relativa promedio del filo Bacteroidota en pacientes con SIBO (34,86 %) en comparación con el grupo Control (38,73 %), mientras que la del filo Bacillota fue similar entre ambos grupos (38,76 % en el grupo SIBO y 37,42 % en el grupo Control).

La dominancia de Enterobacteriaceae en el grupo SIBO podría tener implicaciones clínicas relevantes. Esta familia, que incluye géneros como *Escherichia*, *Klebsiella*, *Enterobacter* y *Citrobacter*, posee una alta capacidad fermentativa de carbohidratos no absorbidos, lo que puede derivar en una producción excesiva de gases, hinchazón y distensión abdominal, explicando parcialmente la fisiopatología de la enfermedad (7). Además, estos géneros albergan múltiples factores de virulencia, como adhesinas, toxinas (por ejemplo, enterotoxinas de *E. coli*) y endotoxinas lipopolisacáridas, que pueden dañar la barrera epitelial intestinal, aumentar su permeabilidad y contribuir a procesos inflamatorios locales o sistémicos (51).

Este patrón de alteración en la composición taxonómica en el que determinadas familias bacterianas ganan predominancia sobre otras, se alinea con la menor equidad observada en los análisis de diversidad alfa y beta, reforzando la hipótesis de que el SIBO se asocia más con desequilibrios funcionales de la microbiota que con una pérdida global de riqueza taxonómica.

5.2. Análisis del resistoma

El análisis del resistoma fecal mostró una alta prevalencia ARGs en ambas cohortes. Se detectaron más de 70 ARGs que confieren resistencia a diversas familias de antibióticos, lo cual concuerda con estudios previos que han demostrado que el intestino humano constituye un reservorio importante de resistencia a antibióticos, reflejo del impacto del uso y exposición a antibióticos en la microbiota intestinal (13–15).

Entre los ARGs más frecuentes se encontraron genes que confieren resistencia a MLS, principalmente variantes del gen *erm*, así como a tetraciclinas, mediante diferentes variantes del gen *tet*. Esta alta prevalencia ha sido ampliamente documentada y se atribuye al uso extensivo de estos antibióticos tanto en medicina humana como veterinaria, lo que ha favorecido su persistencia y diseminación en la microbiota intestinal (52,53).

Asimismo, se detectó el gen *cfxA* en más de 30 muestras de ambos grupos. Este gen, que codifica una β -lactamasa, confiere resistencia a penicilinas y a cefalosporinas de primera y segunda generación. Su alta frecuencia en muestras fecales es esperada, ya que suele encontrarse en bacterias anaerobias del género *Bacteroides*, miembros habituales de la microbiota intestinal humana, según lo descrito previamente por Hasegawa et al. (2002) (54).

De particular relevancia clínica es la detección de β -lactamasas de espectro extendido como *blaCTX-M-15* y *blaDHA-1*, ya que estas enzimas confieren resistencia a cefalosporinas de tercera generación, antibióticos clave en el tratamiento de infecciones por patógenos multirresistentes (55). Del mismo modo, la presencia del gen *cfr(C)*, que confiere resistencia a linezolid, un antibiótico de último recurso frente a bacterias Gram-positivas multirresistentes, representa un hallazgo preocupante desde el punto de vista terapéutico (56).

La comparación de abundancias relativas de ARGs, tanto individuales como agrupadas por familia antibiótica, no mostró diferencias significativas generalizadas entre grupos. Solo un gen, *ant(6)-Ia*, que confiere resistencia a estreptomicina, presentó una reducción significativa

en el grupo SIBO. Esta limitada detección de diferencias puede atribuirse, al menos en parte, al umbral mínimo de detección aplicado (exclusión de genes presentes en menos de 10 muestras), lo cual puede haber reducido la sensibilidad del análisis.

Es importante destacar que la resistencia a rifaximina, antibiótico de primera línea en el tratamiento del SIBO, no suele estar mediada por genes adquiridos como los analizados en este estudio. En su lugar, está relacionada con mutaciones puntuales en el gen *rpoB* (57). Esta limitación también podría explicar la ausencia de una mayor carga diferencial de ARGs en el grupo SIBO.

Otro aspecto metodológico relevante fue la cuidadosa selección de la cohorte Control. Las muestras fueron emparejadas por diversas variables, incluyendo el historial reciente de consumo de antibióticos, con el objetivo de minimizar sesgos atribuibles a exposiciones antimicrobianas previas. Si bien esta estrategia fortalece la validez comparativa entre grupos, también puede haber reducido la magnitud de las diferencias observadas, al introducir cierta homogeneidad en la presión antibiótica previa. No obstante, este diseño asegura que las diferencias detectadas reflejen alteraciones inherentes a la ecología microbiana asociada al SIBO, más que a exposiciones medicamentosas recientes.

En conjunto, estos hallazgos subrayan la complejidad de la dinámica del resistoma intestinal en condiciones clínicas como el SIBO, así como la necesidad de considerar tanto los mecanismos de resistencia mediados por genes transferibles como aquellos derivados de mutaciones cromosómicas.

5.3.Limitaciones

A pesar de la solidez metodológica del presente trabajo, es importante reconocer varias limitaciones que podrían haber influido en los resultados obtenidos. El diagnóstico de SIBO se basa en autodeclaraciones con la etiqueta "Diagnosed by a medical professional (doctor, physician assistant)", lo que mejora la fiabilidad frente al autodiagnóstico, pero no garantiza homogeneidad en los métodos empleados (como pruebas de aliento o cultivos yeyunales), introduciendo riesgo de clasificación errónea. Además, la especificidad y sensibilidad de estas pruebas diagnósticas varían, lo que puede afectar la precisión en la identificación de los casos. Por otro lado, el SIBO puede presentarse de forma

asintomática, lo que dificulta aún más su diagnóstico y posible subestimación en la cohorte estudiada.

Desde el punto de vista técnico, aunque se aplicaron procedimientos robustos de procesamiento y análisis bioinformático, existen limitaciones inherentes al uso de diferentes tecnologías de secuenciación y lotes de procesamiento.

Asimismo, cabe destacar que las muestras analizadas provienen de heces y no del intestino delgado, principal sitio afectado en el SIBO, lo que podría limitar la representatividad de la microbiota y del resistoma intestinal.

Por último, debido a su disponibilidad, el tamaño muestral del análisis del resistoma fue relativamente reducido, lo que podría haber limitado la potencia estadística para detectar diferencias sutiles en genes de baja prevalencia. Además, el uso de herramientas como Abricate con ResFinder, aunque ampliamente validadas, depende de la calidad del ensamblaje y la cobertura de los *contigs*, lo que puede afectar la sensibilidad en la detección de ARGs.

A pesar de estas limitaciones, el presente estudio aporta evidencia relevante sobre el microbioma y resistoma asociados al SIBO, y sienta las bases para futuros trabajos con un enfoque más dirigido, prospectivo y clínicamente validado.

5.4. Prospectiva

Los hallazgos de este trabajo abren diversas líneas de investigación futura en el ámbito del microbioma y el resistoma intestinal en pacientes con SIBO.

En primer lugar, sería fundamental realizar estudios longitudinales que permitan analizar la evolución del microbioma y del resistoma antes, durante y después del tratamiento, con el objetivo de establecer relaciones causales y evaluar su utilidad como biomarcadores pronósticos o de respuesta terapéutica.

Desde un enfoque más integrador, la combinación de datos metagenómicos con otros datos ómicos (metabolómica, transcriptómica o proteómica) podría proporcionar una visión más completa del ecosistema intestinal en el contexto del SIBO. Esta integración permitiría no solo detectar la presencia de ARGs, sino también evaluar su expresión activa y su relevancia

funcional en el entorno intestinal, así como identificar rutas funcionales alteradas con mayor precisión.

Por último, la aplicación de modelos computacionales avanzados y técnicas de aprendizaje automático sobre grandes conjuntos de datos podría facilitar la identificación de biomarcadores robustos y el desarrollo de algoritmos predictivos para el diagnóstico o la selección de tratamientos personalizados.

En conjunto, este trabajo refuerza la evidencia sobre el papel del microbioma en la salud intestinal, destacando su relevancia en el SIBO y abriendo nuevas vías para el desarrollo de herramientas diagnósticas y terapéuticas más precisas.

6. Conclusiones

1. No se detectaron diferencias significativas en los índices de diversidad alfa global de la microbiota fecal (Shannon, número de ASVs y diversidad filogenética de Faith), lo que sugiere una riqueza y diversidad interna similares entre los grupos con y sin diagnóstico de SIBO.
2. El índice de *evenness* fue significativamente menor en individuos con SIBO, lo que refleja un ecosistema intestinal menos equilibrado, dominado por ciertos taxones, y sugiere una menor estabilidad ecológica y una mayor susceptibilidad a perturbaciones.
3. La microbiota fecal de individuos con SIBO presenta alteraciones estructurales respecto al grupo control, con diferencias significativas en la diversidad beta (Bray-Curtis, Jaccard y *unweighted* UniFrac).
4. Se observó un aumento en la abundancia relativa de la familia Enterobacteriaceae y del filo *Pseudomonadota* en el grupo SIBO, ambas asociadas a disbiosis intestinal, procesos inflamatorios y producción de gases.
5. En el análisis del resistoma, no se observaron diferencias significativas en la abundancia relativa de genes de resistencia a antibióticos entre grupos, excepto para el gen *ant(6)-Ia*, cuya presencia fue menor en individuos con SIBO.
6. Se identificaron genes de resistencia de relevancia clínica, como *blaCTX-M-15*, *blaDHA-1* y *cfr(C)*, lo que refuerza el papel del intestino como reservorio potencial de resistencia frente a antibióticos de uso crítico.
7. Este estudio confirma la utilidad del enfoque metagenómico para caracterizar integralmente la microbiota y el resistoma intestinal, proporcionando información valiosa para mejorar el diagnóstico, pronóstico y el diseño de estrategias terapéuticas personalizadas en enfermedades como el SIBO.
8. Futuras investigaciones deberían incluir estudios longitudinales y funcionales, integrando otras ciencias ómicas.

7. Referencias bibliográficas

1. Bidell MR, Hobbs ALV, Lodise TP. Gut microbiome health and dysbiosis: A clinical primer. *Pharmacotherapy*. 2022;42(11):849–57.
2. Fan Y, Pedersen O. Gut microbiota in human metabolic health and disease. *Nat Rev Microbiol*. 2021;19(1):55–71.
3. De Vos WM, Tilg H, Van Hul M, Cani PD. Gut microbiome and health: mechanistic insights. *Gut*. 2022;71(5):1020–32.
4. Gomaa EZ. Human gut microbiota/microbiome in health and diseases: a review. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol*. 2020;113(12):2019–40.
5. Athanasopoulou K, Adamopoulos PGA, Scorilas A. Unveiling the Human Gastrointestinal Tract Microbiome: The Past, Present, and Future of Metagenomics. *Biomedicines*. 2023;11(3):827–43.
6. Alcedo González J, Estremera-Arévalo F, Cobián Malaver J, Santos Vicente J, Alcalá-González LG, Naves J, et al. Common questions and rationale answers about the intestinal bacterial overgrowth syndrome (SIBO). *Gastroenterol Hepatol*. 2024;48(2):502216.
7. Sharabi E, Rezaie A. Small Intestinal Bacterial Overgrowth. *Curr Infect Dis Rep*. 2024;26:227–33.
8. Skrzydło-Radomańska B, Cukrowska B. How to Recognize and Treat Small Intestinal Bacterial Overgrowth? *J Clin Med*. 2022;11(20):1–9.
9. Rao SSC, Bhagatwala J. Small Intestinal Bacterial Overgrowth: Clinical Features and Therapeutic Management. *Clin Transl Gastroenterol*. 2019;10(10):e00078.
10. Losurdo G, Leandro G, Ierardi E, Perri F, Barone M, Principi M, et al. Breath tests for the non-invasive diagnosis of small intestinal bacterial overgrowth: A systematic review with meta-analysis. *J Neurogastroenterol Motil*. 2020;26(1):16–28.
11. Goździewska M, Łyszczarz A, Kaczoruk M, Kolarzyk E. Relationship between SIBO and other bowel diseases and a common eating pattern for them. Part III. *Ann Agric Environ Med*. 2024;31(3):322–8.
12. Munita JM, Arias CA. Mechanisms of antibiotic resistance. *Virulence Mech Bact Pathog*. 2016;4(2):481–511.

13. Anthony WE, Burnham CAD, Dantas G, Kwon JH. The gut microbiome as a reservoir for antimicrobial resistance. *J Infect Dis*. 2021;223(12 Suppl 2):S209–13.
14. Fredriksen S, de Warle S, van Baarlen P, Boekhorst J, Wells JM. Resistome expansion in disease-associated human gut microbiomes. *Microbiome*. 2023;11(1):166–76.
15. Crits-Christoph A, Hallowell HA, Koutouvalis K, Suez J. Good microbes, bad genes? The dissemination of antimicrobial resistance in the human microbiome. *Gut Microbes*. 2022;14(1):2055944–58.
16. Arian M, Muth T. Integrated multi-omics analyses of microbial communities: a review of the current state and future directions. *Mol Omi*. 2023;19(8):607–23.
17. Bharti R, Grimm DG. Current challenges and best-practice protocols for microbiome analysis. *Brief Bioinform*. 2021;22(1):178–93.
18. Nam NN, Dang H, Do K, The K, Trinh L, Lee NY. Metagenomics: An Effective Approach for Exploring Microbial. *Foods*. 2023;12(11):2140–63.
19. Di Carlo P, Serra N, Alduina R, Guarino R, Craxì A, Giammanco A, et al. A systematic review on omics data (metagenomics, metatranscriptomics, and metabolomics) in the role of microbiome in gallbladder disease. *Front Physiol*. 2022;13:1–14.
20. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol*. 2017;35(9):833–44.
21. Clarridge JE. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin Microbiol Rev*. 2004;17(4):840–62.
22. Shahi SK, Freedman SN, Mangalam AK. Gut microbiome in multiple sclerosis: The players involved and the roles they play. *Gut Microbes*. 2017;8(6):607–15.
23. Woalder, Ranjan. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Physiol Behav*. 2017;176(3):139–48.
24. Sharpton TJ. An introduction to the analysis of shotgun metagenomic data. *Front Plant Sci*. 2014;5:1–14.
25. Liu YX, Qin Y, Chen T, Lu M, Qian X, Guo X, et al. A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell*. 2021;12(5):315–30.
26. Goodwin S, McPherson JD, McCombie WR. Coming of age: Ten years of next-generation sequencing technologies. *Nat Rev Genet*. 2016;17(6):333–51.
27. Overholt WA, Hölzer M, Geesink P, Diezel C, Marz M, Küsel K. Inclusion of Oxford Nanopore long reads improves all microbial and viral metagenome-assembled

- genomes from a complex aquifer system. *Environ Microbiol.* 2020;22(9):4000–13.
28. Xie H, Yang C, Sun Y, Igarashi Y, Jin T, Luo F. PacBio Long Reads Improve Metagenomic Assemblies, Gene Catalogs, and Genome Binning. *Front Genet.* 2020;11:516269–77.
 29. Ho DE, King G, Stuart EA, Imai K. MatchIt: Nonparametric Preprocessing for. *J Stat Softw.* 2011;42(8):1–28.
 30. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol.* 2019;37(8):852–7.
 31. Roswell M, Dushoff J, Winfree R. A conceptual guide to measuring species diversity. *Oikos.* 2021;130(3):321–38.
 32. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Internet]. 2010. [cited 2025 Jun 10]. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
 33. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20.
 34. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics.* 2015;31(10):1674–6.
 35. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics.* 2013;29(8):1072–5.
 36. Seemann T. Abricate [Internet]. [cited 2025 Jun 12]. Available from: <https://github.com/tseemann/abricate>
 37. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother.* 2020;75(12):3491–500.
 38. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: Architecture and applications. *BMC Bioinformatics.* 2009;10:1–9.
 39. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–9.
 40. Kolde R. pheatmap: Pretty Heatmaps [Internet]. 2025 [cited 2025 Jun 6]. Available from: <https://cran.r-project.org/web/packages/pheatmap/index.html>
 41. Leite G, Morales W, Weitsman S, Celly S, Parodi G, Mathur R, et al. The duodenal

- microbiome is altered in small intestinal bacterial overgrowth. *PLoS One*. 2020;15(7):1–17.
42. Saffouri GB, Shields-Cutler RR, Chen J, Yang Y, Lekatz HR, Hale VL, et al. Small intestinal microbial dysbiosis underlies symptoms associated with functional gastrointestinal disorders. *Nat Commun*. 2019;10(1):1–11.
 43. Park SK, Kim HN, Choi CH, Im JP, Cha JM, Eun CS, et al. Differentially abundant bacterial taxa associated with prognostic variables of crohn's disease: Results from the impact study. *J Clin Med*. 2020;9(6):1–17.
 44. Ott SJ, Musfeldt M, Wenderoth DF, Hampe J, Brant O, Fölsch UR, et al. Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut*. 2004;53(5):685–93.
 45. Wei L, Singh R, Ro S, Ghoshal UC. Gut microbiota dysbiosis in functional gastrointestinal disorders: Underpinning the symptoms and pathophysiology. *JGH Open*. 2021;5(9):976–87.
 46. Langdon A, Crook N, Dantas G. The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome Med*. 2016;8(1):39.
 47. Fassarella M, Blaak EE, Penders J, Nauta A, Smidt H, Zoetendal EG. Gut microbiome stability and resilience: Elucidating the response to perturbations in order to modulate gut health. *Gut*. 2021;70(3):595–605.
 48. Relman DA. The human microbiome: Ecosystem resilience and health. *Nutr Rev*. 2012;70(Suppl 1):S2-9.
 49. Barlow JT, Leite G, Romano AE, Sedighi R, Chang C, Celly S, et al. Quantitative sequencing clarifies the role of disruptor taxa, oral microbiota, and strict anaerobes in the human small-intestine microbiome. *Microbiome*. 2021;9(1):1–17.
 50. Yersin S, Vonaesch P. Small intestinal microbiota: from taxonomic composition to metabolism. *Trends Microbiol*. 2024;32(10):970–83.
 51. Dubreuil JD. Enterotoxigenic *Escherichia coli* targeting intestinal epithelial tight junctions: An effective way to alter the barrier integrity. *Microb Pathog*. 2017;113:129–34.
 52. Sóki J, Wybo I, Baaity Z, Stefán G, Jeverica S, Ulger N, et al. Detection of the antibiotic resistance genes content of intestinal *Bacteroides*, *Parabacteroides* and *Phocaeicola*

- isolates from healthy and carbapenem-treated patients from European countries. *BMC Microbiol.* 2024;24(1):1–9.
53. Hu Y, Yang X, Qin J, Lu N, Cheng G, Wu N, et al. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun.* 2013;4:2151.
 54. Hasegawa P, Hasegawa P, Fosse T, Giudicelli J, Shoemaker NB, Gluzman Y, et al. What Constitutes an Extended-Spectrum beta-lactamase? 2002;46(2):600–1.
 55. Bush K, Bradford PA. beta-lactams and beta-lactamase Inhibitors: An Overview. *Cold Spring Harb Perspect Med.* 2016;6(8):a025247.
 56. Schwarz S, Zhang W, Du XD, Krüger H, Feßler AT, Ma S, et al. Mobile oxazolidinone resistance genes in gram-positive and gram-negative bacteria. *Clin Microbiol Rev.* 2021;34(3):1–63.
 57. Kothary V, Scherl EJ, Bosworth B, Jiang ZD, DuPont HL, Harel J, et al. Rifaximin Resistance in *Escherichia coli* Associated with Inflammatory Bowel Disease Correlates with Prior Rifaximin Use, Mutations in *rpoB*, and Activity of Phe-Arg- β -Naphthylamide-Inhibitable Efflux Pumps. *Antimicrob Agents Chemother.* 2013;57(2):811–7.

8. Anexos

Anexo A. Comandos utilizados para el análisis taxonómico con QIIME2

Preprocesamiento de los datos

Las muestras se importaron a QIIME2 desde un archivo *manifest* en formato .csv. Se utilizó el siguiente comando:

```
qiime tools import --type 'SampleData[SequencesWithQuality]' --input-path  
samples.csv --output-path demux.qza --input-format  
SingleEndFastqManifestPhred33
```

Argumentos

- `--type`: Especifica el tipo de datos que se van a importar. Para lecturas *single-end* se utiliza 'SampleData[SequencesWithQuality]', mientras que para *paired-end* se utiliza 'SampleData[PairedEndSequencesWithQuality]'.
- `--input-path`: Archivo *manifest* en formato .csv que contiene las siguientes las columnas *sample-id*, *absolute-filepath*, *direction*.
- `--output-path`: Define el directorio y archivo donde se guardarán los resultados del artefacto de QIIME2 en formato .qza que contiene las secuencias asociadas a cada muestra y sus metadatos.
- `--input-format`: Indica el formato de las lecturas que se van a importar. SingleEndFastqManifestPhred33 indica que las lecturas son *single-end* y codificadas con calidad Phred+33, mientras que con las lecturas son *paired-end* el argumento es PairedEndFastqManifestPhred33.

Se genera un resumen para visualizarlo en QIIME 2 View con el fin de evaluar calidad, longitud de secuencias y número por muestra. El comando utilizado fue:

```
qiime demux summarize --i-data demux.qza --o-visualization demux-summary.qzv
```

Argumentos

- --i-data: Archivo .qza generado en el paso anterior.
- --o-visualization: Archivo de salida en formato .qzv para visualización en QIIME 2 View.

Se eliminan los posibles adaptadores de las lecturas mediante Cutadapt. Para ello, se utilizó el siguiente comando:

```
qiime cutadapt trim-single --i-demultiplexed-sequences demux.qza --o-trimmed-sequences trimmed-demux.qza
```

Argumentos

- trim-single o trim-paired: Indica si las lecturas son *single-end* o *paired-end*.
- --i-demultiplexed-sequences: Archivo en formato .qza que contiene las secuencias asociadas a cada muestra y sus metadatos generado en el primer paso.
- --o-trimmed-sequences: Archivo en formato .qza con las secuencias tras el recorte de adaptadores.

A continuación, se eliminan las secuencias con calidad por debajo de un umbral determinado. Se utilizó el siguiente comando:

```
qiime quality-filter q-score --i-demux trimmed-demux.qza --p-min-quality 30 -o-filtered-sequences filtered-demux.qza --o-filter-stats filtered-demux-stats.qza
```

Argumentos

- `--i-demux`: Archivo en formato .qza con las secuencias tras el recorte de adaptadores.
- `--p-min-quality 30`: Umbral de calidad establecido por base de 30 PHRED.
- `--o-filtered-sequences`: Archivo de salida en formato .qza que contiene las secuencias que cumplen el criterio de calidad establecido.
- `--o-filter-stats`: Archivo de salida en formato .qza que contiene estadísticas del filtrado (porcentaje de secuencias retenidas, eliminadas, quimeras, etc.).

Se utilizó DADA2 para el proceso de *denoising* que consiste en corregir errores de secuenciación e identificar las ASVs. El comando utilizado fue:

```
qiime dada2 denoise-single --i-demultiplexed-seqs filtered-demux.qza --o-table  
ASV_table.qza --o-representative-sequences rep-seqs.qza --o-denoising-stats  
denoising-stats.qza --p-n-reads-learn 1000000 --p-trunc-len 140
```

Argumentos

- `--i-demultiplexed-seqs`: Archivo de entrada en formato .qza generado en el paso anterior y que contiene las secuencias que cumplen el criterio de calidad establecido.
- `--o-table`: Archivo de salida en formato .qza que contiene la tabla de abundancia de ASVs.
- `--o-representative-sequences`: Archivo de salida en formato .qza que contiene las secuencias representativas de cada ASV.
- `--o-denoising-stats`: Archivo de salida en formato .qza que contiene las estadísticas del proceso de *denoising*.
- `--p-n-reads-learn 1000000`: Establece en 1000000 el número de lecturas que DADA2 utilizará para estimar el modelo de error.
- `--p-trunc-len 140`: Parámetro que indica el número de bases a las que se recortarán las lecturas. En este caso, se recortan las lecturas a 140 bases para eliminar las regiones de baja calidad al final de la lectura.

Posteriormente se filtran las muestras en función de su número de lecturas para asegurar que las muestras incluidas en el análisis tengan una cantidad suficiente de datos para ser representativas. El comando utilizado es:

```
qiime feature-table filter-samples --i-table ASV_table.qza --p-min-frequency 5000 --o-filtered-table filtered-table_5000.qza
```

Argumentos

- --i-table: Archivo en formato .qza generado en el paso anterior que contiene la tabla de abundancia de ASVs.
- --p-min-frequency 5000: Establece en 5000 el umbral mínimo de lecturas que debe tener cada muestra para ser incluida en el análisis.
- --o-filtered-table: Archivo de salida en formato .qza que contiene las muestras que superan el umbral establecido.

Posteriormente, se filtran las ASVs para mantener únicamente aquellas asociadas a muestras que han superado el umbral mínimo de lecturas:

```
qiime feature-table filter-seqs --i-data rep-seqs.qza --i-table filtered-table_5000.qza --o-filtered-data rep-seqs.qza
```

Argumentos

- --i-data: Archivo de entrada en formato .qza que contiene las secuencias representativas antes del filtrado.
- --i-table: Archivo en formato .qza con la tabla de abundancia filtrada, que incluye únicamente las muestras que superan el umbral de lecturas .
- --o-filtered-data: Archivo de salida en formato .qza que contiene únicamente las ASVs correspondientes a las muestras retenidas tras el filtrado.

Para obtener un resumen visual de la tabla de abundancia de ASVs, que incluya estadísticas descriptivas y distribución de frecuencias por muestra, se utilizó el siguiente comando:

```
qiime feature-table summarize --i-table filtered-table_5000.qza \
--o-visualization filtered-table_5000-summary.qzv
```

Argumentos

- --i-table: Archivo de entrada en formato .qza que contiene la tabla de abundancia de ASVs previamente filtrada.
- --o-visualization: Archivo de salida en formato .qzv que genera un resumen interactivo.

Asignación taxonómica y análisis filogenético

Para la asignación taxonómica se utilizó la base de datos Greengenes mediante el siguiente comando:

```
qiime feature-classifier classify-sklearn --i-classifier gg-13-8-99-nb-
classifier.qza --i-reads rep-seqs.qza --o-classification taxonomy.qza
```

Argumentos

- --i-classifier: Archivo en formato .qza que contiene el clasificador preentrenado.
- --i-reads: Archivo en formato .qza que contiene las secuencias representativas de las ASVs (salida del paso de DADA2).
- --o-classification: Archivo de salida en formato .qza que contendrá la tabla de asignación taxonómica para cada ASV.

Para la construcción del árbol filogenético, en primer lugar se realizó un alineamiento de las secuencias, mediante MAFFT, utilizando el siguiente comando:

```
qiime alignment mafft --i-sequences rep-seqs.qza --o-alignment aligned-rep-seqs.qza
```

Argumentos

- --i-sequences: Archivo en formato .qza con las ASVs.
- --o-alignment: Archivo de salida .qza que contiene las secuencias alineadas.

A continuación, se realizó el enmascaramiento del alineamiento:

```
qiime alignment mask --i-alignment aligned-rep-seqs.qza --o-masked-alignment masked-aligned-rep-seqs.qza
```

Argumentos

- --i-alignment: Archivo en formato .qza que contiene el alineamiento de secuencias generado en el paso anterior.
- --o-masked-alignment: Archivo de salida en formato .qza que contiene el alineamiento con posiciones poco informativas o altamente variables enmascaradas para mejorar la precisión del árbol filogenético.

En siguiente paso fue la construcción del árbol filogenético mediante el siguiente comando:

```
qiime phylogeny fasttree --i-alignment masked-aligned-rep-seqs.qza --o-tree unrooted-tree.qza
```

Argumentos

- --i-alignment: Archivo en formato .qza que contiene el alineamiento enmascarado generado en el paso anterior.

- --o-tree: Archivo de salida .qza con el árbol filogenético sin enraizar.

Por último, se realizó el enraizamiento del árbol:

```
qiime phylogeny midpoint-root --i-tree unrooted-tree.qza --o-rooted-tree rooted-tree.qza
```

Argumentos

- --i-tree: Archivo en formato .qza que contiene el árbol filogenético sin enraizar obtenido en el paso anterior.
- --o-rooted-tree: Archivo de salida en formato .qza con el árbol enraizado en el punto medio (*midpoint rooting*).

Diversidad alfa y beta

Se calcularon diferentes métricas de diversidad alfa y beta mediante el siguiente comando:

```
qiime diversity core-metrics-phylogenetic --i-table filtered_table.qza --i-phylogeny rooted-tree.qza --m-metadata-file metadata.tsv --p-sampling-depth 5000 --output-dir core-metrics-results
```

Argumentos

- --i-table: Archivo en formato .qza que contiene la tabla de características con las abundancias relativas.
- --i-phylogeny: Archivo .qza que contiene el árbol filogenético enraizado.
- --m-metadata-file: Archivo en formato .tsv con la información de metadatos de las muestras, necesario para relacionar las muestras con las variables.
- --p-sampling-depth: Profundidad de muestreo para normalizar la tabla; aquí se estableció en 5000 lecturas por muestra para controlar la heterogeneidad en el número de secuencias.

- --output-dir: Directorio donde se guardarán los resultados generados con las métricas de diversidad alfa y beta.

Para visualizar y analizar la significancia estadística de los índices de diversidad alfa por grupos, se utilizó el siguiente comando (usando como ejemplo el índice de Shannon), que genera gráficos de cajas junto con la prueba no paramétrica de Kruskal-Wallis para comparar grupos:

```
qiime diversity alpha-group-significance --i-alpha-diversity core-metrics-  
results/shannon_vector.qza --m-metadata-file metadata.tsv --o-visualization  
shannon.qzv>
```

Argumentos

- --i-alpha-diversity: Archivo .qza que contiene los valores de diversidad alfa previamente calculados para cada índice.
- --m-metadata-file: Archivo en formato.tsv con la información de metadatos de las muestras, necesario para relacionar las muestras con las variables.
- --o-visualization: Archivo de salida en formato .qzv que contiene los resultados visuales e inferenciales de la prueba estadística aplicada.

Para evaluar si existen diferencias significativas en la composición microbiana global entre los grupos estudiados mediante análisis PERMANOVA de los índices de diversidad beta, se aplicó el siguiente comando (se utiliza como ejemplo al distancia ded Bray-Curtis):

```
qiime diversity beta-group-significance --i-distance-matrix  
braycurtis_distance.qza --m-metadata-file metadata.tsv --m-metadata-column  
sibo --p-method permanova --o-visualization permanova_results_braycurtis.qzv
```

Argumentos

- --i-distance-matrix: Archivo .qza con la matriz de distancias del índice de diversidad beta, que refleja las diferencias en la composición microbiana entre muestras.

- --m-metadata-file: Archivo en formato .tsv con la información de metadatos de las muestras, necesario para relacionar las muestras con las variables.
- --m-metadata-column: Nombre de la columna del archivo de metadatos que contiene la variable categórica de interés, en este caso "sibo".
- --p-method: Método estadístico para evaluar la significancia. En este caso se utilizó PERMANOVA.
- --o-visualization: Archivo de salida .qzv con los resultados estadísticos y gráficos interpretables de la prueba aplicada.

Composición taxonómica

Para representar gráficamente la composición microbiana de las muestras a distintos niveles taxonómicos (filo, clase, orden, etc.), se utilizó el siguiente comando:

```
qiime taxa barplot --i-table filtered_table.qza --i-taxonomy taxonomy.qza --m-metadata-file metadata.tsv --o-visualization taxa-barplot.qzv
```

Argumentos

- --i-table: Archivo en formato .qza que contiene la tabla de características con las abundancias relativas.
- --i-taxonomy: Archivo en formato .qza que contiene la tabla de asignación taxonómica para cada ASV.
- --m-metadata-file: Archivo en formato .tsv con la información de metadatos de las muestras, necesario para relacionar las muestras con las variables.
- --o-visualization: Archivo .qzv de salida que permite visualizar, de forma interactiva en QIIME 2 View, gráficos de barras apiladas de abundancias relativas por muestra o grupo.

Con el fin de identificar taxones diferencialmente abundantes entre los grupos de estudio, se realizó un análisis composicional utilizando el método ANCOM.

En primer lugar, las secuencias se agruparon por nivel taxonómico:

```
qiime taxa collapse --i-table filtered_table.qza --i-taxonomy taxonomy.qza --p-level 6 --o-collapsed-table collapsed_table.qza
```

Argumentos

- --i-table: Archivo en formato .qza que contiene la tabla de características con las abundancias relativas.
- --i-taxonomy: Archivo en formato .qza que contiene la tabla de asignación taxonómica para cada ASV.
- --p-level: Nivel taxonómico deseado.
- --o-collapsed-table: Archivo en formato .qza con la tabla de salida colapsada por el nivel taxonómico indicado.

En segundo lugar, dado que ANCOM requiere que no existan ceros en la tabla, se aplicó un pseudoconteo uniforme a todos los valores:

```
qiime composition add-pseudocount --i-table collapsed_table.qza --o-composition-table compositional_table.qza
```

Argumentos

- --i-table: Archivo en formato .qza con la tabla de salida colapsada por el nivel taxonómico indicado.
- --o-composition-table: Archivo de salida en formato .qza con la tabla composicional por el nivel taxonómico indicado.

Finalmente, se ejecutó el ANCOM con el siguiente comando:

```
qiime composition ancom --i-table compositional_table.qza --m-metadata-file metadata.tsv --m-metadata-column sibo --o-visualization ancom.qzv
```

Argumentos

- --i-table: : Archivo en formato .qza con la tabla composicional por el nivel taxonómico indicado.
- --m-metadata-file: Archivo en formato .tsv con la información de metadatos de las muestras, necesario para relacionar las muestras con las variables.
- --m-metadata-column: Nombre de la columna del archivo de metadatos que contiene la variable categórica de interés, en este caso “sibo”.
- --o-visualization: Archivo de salida en formato .qzv con los resultados de ANCOM, incluyendo W-scores y visualizaciones interactivas.

Anexo B. Comandos utilizados para la detección de genes de resistencia a antibióticos

Control de Calidad con FastQC

Se utilizó FastQC con el fin de evaluar la calidad de las lecturas de las muestras antes del análisis. Para ello, se utilizó el siguiente comando:

```
fastqc reads.fastq --outdir=./fastqc_results
```

Argumentos

- <archivo_fastq>: Archivo de lecturas en formato .fastq a analizar.
- --outdir=: Define el directorio donde se guardarán los resultados generados por FastQC.

Se genera un informe detallado en formato .html sobre la calidad de las lecturas, incluyendo contenido en adaptadores y longitud y calidad de las lecturas. También se genera un archivo .zip con la misma información del .html con los archivos en formato .txt y .png, entre otros.

Filtrado por calidad con Trimmomatic

Se utilizó Trimmomatic para el filtrado de las lecturas con baja calidad y eliminación de adaptadores. El comando utilizado fue el siguiente, para lecturas *paired-end*:

```
trimmomatic PE -phred33 reads_R1.fastq reads_R2.fastq reads_R1_filtered.fastq  
reads_R1_unpaired.fastq reads_R2_filtered.fastq reads_R2_unpaired.fastq  
LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:50
```

Argumentos

- -phred33: formato de codificación de calidad

- Input: Archivos en formato .fastq que contienen las lecturas originales.
- -Output: Archivos de salida en formato .fastq con las lecturas filtradas (pareadas y no pareadas).
- LEADING:3: elimina bases con calidad <3 desde el inicio.
- TRAILING:3: elimina bases con calidad <3 desde el final.
- SLIDINGWINDOW:4:20: recorta cuando el promedio de calidad en una ventana de 4 bases cae por debajo de 20.
- MINLEN:50: Solo se mantienen las lecturas con al menos 50 bases.

Control de Calidad Posterior al Filtrado con FastQC

Se ejecutó nuevamente FastQC para asegurar que las lecturas limpias cumplieran con los estándares de calidad. El comando utilizado fue el mismo que en el primer paso.

Ensamblaje con Megahit

Para el ensamblaje de las lecturas filtradas, se utilizó Megahit. El comando ejecutado fue:

```
megahit -1 reads_R1_filtered.fastq -2 reads_R2_filtered.fastq -o
directorio_salida
```

Argumentos

- -1 -2: Archivos en formato .fastq que contienen las lecturas filtradas.
- -o: Directorio donde se almacenarán los *contigs* ensamblados.

Detección de genes de resistencia a antibióticos con Abricate

Para la identificación de genes de resistencia a antibióticos se utilizó Abricate con la base de datos ResFinder. El comando utilizado fue:

```
abricate --db resfinder --minid 80 --mincov 30 contigs.fasta resultados.tab
```

Argumentos

- --db resfinder: Especifica la base de datos ResFinder para la búsqueda de genes de resistencia.
- --minid 80: Define un umbral de identidad mínima del 80% para considerar un alineamiento como positivo.
- --mincov 30: Establece un umbral de cobertura mínima del 30% para los alineamientos.
- contigs.fasta: Archivo en formato .fasta que contiene los *contigs* ensamblados.
- resultados.tab: Archivo de salida en formato .tab con los resultados de la búsqueda de genes de resistencia.

Cuantificación de genes de resistencia a antibióticos

Para la cuantificación de genes de resistencia a antibióticos se mapearon las lecturas frente a los *contigs* ensamblados usando Bowtie2.

En primer lugar, se construyó el índice de Bowtie2 a partir del archivo de *contigs* ensamblados:

```
bowtie2-build contigs.fasta contigs_index
```

Argumentos

- contigs.fasta: Archivo de entrada en formato .fasta con los *contigs* ensamblados.
- contigs_index: Nombre base para los archivos de índice generados.

Se mapearon las lecturas filtradas contra los *contigs*:

```
bowtie2 -x contigs_index -1 <lecturas_filtradas_R1.fastq> -2  
<lecturas_filtradas_R2.fastq> -S salida.sam
```

Argumentos

- -x contigs_index: Índice generado a partir del ensamblaje.

- -1, -2: Archivos de lecturas emparejadas (*paired-end*) en formato .fastq filtradas. En el caso de lecturas single-end es -U.
- -S salida.sam: Archivo de salida en formato .sam con los resultados del mapeo.

Se realiza la conversión de formatos para análisis posteriores:

- De .sam a .bam con Samtools:

```
samtools view -Sb entrada.sam > salida.bam
```

- De .bam a .bed con Bedtools:

```
bedtools bamtobed -i entrada.bam > salida.bed
```

Argumentos

- -Sb salida.sam: Archivo de entrada en formato .sam con los resultados del mapeo.
- salida.bam: Archivo de salida en formato .bam con los resultados del mapeo.
- -i entrada.bam: Archivo de entrada en formato . con los resultados del mapeo.
- salida.bed: Archivo de salida en formato .bed con los resultados del mapeo.

Se realiza la conversión de resultados de ABRicate a formato .bed para su integración en el análisis de cobertura mediante en comando awk:

```
awk 'BEGIN{OFS="\t"} NR>1 {print $1, $2-1, $3, $5}' abricate_output.tab > abricate_output.bed
```

Argumentos

- BEGIN{OFS="\t"}: Define el separador de salida como tabulación (formato requerido por .bed).
- NR>1: Omite la primera línea (cabecera del archivo .tab).
- \$1, \$2-1, \$3, \$5: Columnas correspondientes al inicio, fin y nombre del gen en el archivo .tab de ABRicate.

- abricate_output.bed: Archivo de salida en formato .bed con los resultados de ABRicate.

Se realiza el cálculo de la cobertura de los genes de resistencia a antibióticos identificados, es decir, el número de lecturas alineadas que se solapan con cada gen detectado por ABRicate. Para ello se emplea el siguiente comando de Bedtools:

```
bedtools intersect -a abricate_output.bed -b salida.bed -c > _cobertura.tsv
```

Argumentos

- -a: Archivo en formato .bed que contiene las coordenadas genómicas de los genes de resistencia identificados por ABRicate .
- -b: Archivo en formato .bed que contiene las posiciones de las lecturas alineadas contra el ensamblaje metagenómico obtenido con Bowtie2.
- -c: Indica que se debe contar cuántas regiones del archivo -b (lecturas alineadas) se solapan con cada intervalo del archivo -a (genes de resistencia).
- >: Archivo de salida en formato .tsv que contiene, para cada gen de resistencia, el número de lecturas que lo cubren.

9. ÍNDICE DE SIGLAS Y ACRÓNIMOS

A

ADN: Ácido desoxirribonucleico

AGP: *American Gut Project*

ANCOM: Análisis de composiciones de microbiomas (por sus siglas en inglés)

ARGs: Genes de resistencia a antibióticos (por sus siglas en inglés)

ARN: Ácido ribonucleico

ARNr: ARN ribosómico

ASV: Variante de secuencia de amplicón (por sus siglas en inglés)

E

EII: Enfermedad Inflamatoria Intestinal

ENA: European Nucleotide Archive

I

ITS: Espaciador transcrito interno (por sus siglas en inglés)

K

Kb: Kilobase

M

MAFFT: Alineamiento múltiple mediante transformación rápida de Fourier (por sus siglas en inglés)

Mb: Megabases

MDR: Multiresistencia (por sus siglas en inglés)

MLS: Macrólidos, lincosamidas y estreptograminas

O

ONT: Oxford Nanopore Technologies

P

PacBio: Pacific Biosciences

pb: Pares de bases

PCoA: Análisis de Coordenadas Principales (por sus siglas en inglés)

PCR: Reacción en cadena de la polimerasa (por sus siglas en inglés)

PERMANOVA: Análisis de varianza multivariante permutacional (por sus siglas en inglés)

R

RPKM: Lecturas por kilobase de lecturas en la muestra-por millón de lecturas mapeadas (por sus siglas en inglés)

S

SIBO: Sobrecrecimiento bacteriano del intestino delgado (por sus siglas en inglés)

SII: Síndrome del Intestino Irritable

U

UFCs: Unidades Formadoras de Colonias