Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Ciencias Biológicas
Escuela Profesional de Genética y Biotecnología

Variación espacio-temporal de la comunidad bacteriana en áreas de cultivo de *Oncorhynchus mykiss* de la cuenca hidrográfica del Lago Titicaca, Perú, durante el 2018 y 2019 utilizando ADN metabarcoding

TESIS

Para optar el Título Profesional de Biólogo Genetista
Biotecnólogo

AUTOR

Bryan Joaquin LLATANCE SALAZAR

ASESOR

Mg. Giovanna Elizabeth SOTIL CAYCHO

Lima, Perú

2026



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Llatance, B. (2026). *Variación espacio-temporal de la comunidad bacteriana en áreas de cultivo de Oncorhynchus mykiss de la cuenca hidrográfica del Lago Titicaca, Perú, durante el 2018 y 2019 utilizando ADN metabarcoding*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, Escuela Profesional de Genética y Biotecnología]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Bryan Joaquin Llatance Salazar
Tipo de documento de identidad	DNI
Número de documento de identidad	75556623
URL de ORCID	https://orcid.org/0000-0003-0986-8220
Datos de asesor	
Nombres y apellidos	Giovanna Elizabeth Sotil Caycho
Tipo de documento de identidad	DNI
Número de documento de identidad	25836223
URL de ORCID	https://orcid.org/0000-0002-1583-8821
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Julio Solis Sarmiento
Tipo de documento	DNI
Número de documento de identidad	09671619
Miembro del jurado 1	
Nombres y apellidos	Carlos David Neyra Rivera
Tipo de documento	DNI
Número de documento de identidad	42528324
Miembro del jurado 2	
Nombres y apellidos	Abraham Omar Espinoza Culupu
Tipo de documento	DNI
Número de documento de identidad	42688779
Datos de investigación	
Línea de investigación	A.1.1.3. Diversidad Genética
Grupo de investigación	No aplica

Agencia de financiamiento	PpR 2025 - Proyecto 5 Caracterización molecular de Organismos Acuáticos - IMARPE-PRODUCE
Ubicación geográfica de la investigación	Instituto del Mar del Perú (IMARPE), Laboratorio de Genética Molecular País: Perú Departamento: Lima Provincia: Callao Distrito: Callao Calle: Esquina Gamarra y General Valle s/n. 07021 Latitud: -12.0666687 Longitud: -77.1580306
Año o rango de años en que se realizó la investigación	Febrero 2025 - Noviembre 2025
URL de disciplinas OCDE	Bioquímica, Biología molecular https://purl.org/pe-repo/ocde/ford#1.06.03 Bioinformática https://purl.org/pe-repo/ocde/ford#1.02.03

Universidad Nacional Mayor de San Marcos
(Universidad del Perú, Decana de América)

FACULTAD DE CIENCIAS BIOLÓGICAS

ACTA DE SESIÓN PARA OPTAR AL TÍTULO PROFESIONAL DE BIÓLOGO GENETISTA BIOTECNÓLOGO
(MODALIDAD: SUSTENTACIÓN PRESENCIAL DE TESIS)

Siendo las 10:06 horas del 20 de abril de 2026, en el Salón de Grados de la Facultad de Ciencias Biológicas y en presencia del Jurado formado por los profesores que suscriben, se inició la sesión para optar al **Título Profesional de Biólogo Genetista Biotecnólogo** de **BRYAN JOAQUIN LLATANCE SALAZAR.**

Luego de dar lectura y conformidad al expediente **N.º UNMSM-20250133179**, el titulando expuso su Tesis: **VARIACIÓN ESPACIO-TEMPORAL DE LA COMUNIDAD BACTERIANA EN ÁREAS DE CULTIVO DE *Oncorhynchus mykiss* DE LA CUENCA HIDROGRÁFICA DEL LAGO TITICACA, PERÚ, DURANTE EL 2018 Y 2019 UTILIZANDO ADN METABARCODING** y el Jurado efectuó las preguntas del caso calificando la exposición con la nota 19, calificativo: APROBADO CON MAXIMOS HONORES

Finalmente, el expediente será enviado a la Escuela Profesional de Genética y Biotecnología y al Consejo de Facultad para que se apruebe otorgar el **Título Profesional de Biólogo Genetista Biotecnólogo** a **BRYAN JOAQUIN LLATANCE SALAZAR** y se eleve lo actuado al Rectorado para conferir el respectivo Título, conforme a ley.

Siendo las 11:28 horas se levantó la sesión.

Ciudad Universitaria, 20 de abril de 2026.

Dr. JULIO SOLIS SARMIENTO
(PRESIDENTE)

Mg. GIOVANNA ELIZABETH SOTIL CAYCHO
(ASESORA)

Dr. CARLOS DAVID NEYRA RIVERA
(MIEMBRO)

Dr. ABRAHAM OMAR ESPINOZA CULUPU
(MIEMBRO)

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú, Decana de América

Anexo 1
CERTIFICADO DE SIMILITUD

Yo, Giovanna Elizabeth Sotil Caycho en mi condición de asesora acreditada con el **Dictamen de la Escuela Profesional N° 006-2026-EPGBT-FCB/UNMSM** de la tesis cuyo título es:

Variación espacio-temporal de la comunidad bacteriana en áreas de cultivo de *Oncorhynchus mykiss* de la cuenca hidrográfica del Lago Titicaca, Perú, durante el 2018 y 2019 utilizando ADN metabarcoding, presentado por el Bachiller Sr. Bryan Joaquin LLATANCE SALAZAR para optar al Título Profesional de Biólogo Genetista Biotecnólogo,

CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 9% de similitud, nivel PERMITIDO para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del Título Profesional de Biólogo Genetista Biotecnólogo.

Lima, 08 de abril de 2026

Giovanna Elizabeth Sotil Caycho
DNI 25836223

La vigencia del presente certificado corresponde a un período de noventa (90) días desde la fecha de su emisión; siempre que no haya observaciones en el texto durante el proceso de revisión por el jurado informante o surja alguna contingencia.

DEDICATORIA

A mis padres Abel y Lucila, a mis hermanos Valery y Jordan, por ser el pilar de mi vida y la razón por la que me esfuerzo cada día. Todo lo que he logrado se los debo a ellos.

A mis abuelos Timoteo y Guadalupe por todo su cariño y apoyo incondicional. A mi abuela Paulina, por cada momento que compartimos y a mi abuelo Ananias, aunque ya no se encuentre junto a mí, todavía guardo el recuerdo de la gran persona que es.

A mis tíos y primos, por su apoyo y consejos que me dieron a lo largo de mi vida. Por los buenos deseos y palabras de aliento que me sirvieron de inspiración.

A todas las personas con quienes compartí momentos únicos que atesoraré toda mi vida. Todos forman en esencia quien soy, porque llevo conmigo una parte de cada uno de ellos.

AGRADECIMIENTOS

A la Mg. Giovanna Sotil Caycho, jefa del Laboratorio de Genética Molecular del IMARPE, por su apoyo, enseñanzas y la confianza depositada durante el desarrollo de esta tesis; por el gran compromiso y dedicación que demuestra en su labor, y por impulsarme a mejorar en cada aspecto de mí en lo académico y personal.

Al Laboratorio Continental de Puno del IMARPE, por el valioso apoyo logístico de coordinación con los piscicultores para la obtención de las muestras de agua durante sus campañas de muestreo, así como por facilitar sus instalaciones para el procesamiento de las muestras.

A los miembros del jurado revisor de tesis, Dr. Julio Solís, Dr. Carlos Neyra y Dr. Abraham Espinoza, por su compromiso, rigurosidad y valiosas observaciones, las cuales contribuyeron de manera significativa a fortalecer y mejorar el manuscrito de tesis.

A mis amigos y compañeros de laboratorio, Deivis y Paul, a quienes tengo un gran aprecio y admiración, por los momentos de arduo trabajo, los buenos ánimos, la paciencia y la calidez que fueron un impulso cotidiano en el laboratorio. A Valery, Luana, Andrea, Paola y Ana con quienes pasé gratos momentos en el laboratorio, por su apoyo constante y ayuda desinteresada.

A la Dra. Raquel Siccha, por permitirme dar mis primeros pasos en el Laboratorio de Genómica y Bioinformática para la Biodiversidad de la UNMSM, donde descubrí una gran pasión por la investigación y adquirí valiosas experiencias para mi desarrollo profesional.

A los profesores de la FCB, Wilser García, Jorge Ramírez, Claudia Barletta y Martha Valdivia, por sus enseñanzas, consejos y experiencia compartida a lo largo de mi formación.

Finalmente, un agradecimiento especial a mis mejores amigos, Cesar e Ines, por su apoyo constante durante tantos años en la universidad y colegio, y por todas las alegrías que pasamos juntos. También a mi amigos y compañeros de la Base 20, Jose, Marko, Fabricio, Sebastian y Daniel, por las buenas pláticas, las risas y el apoyo en momentos difíciles.

LISTA DE ABREVIATURAS

ADN	Ácido desoxirribonucleico
ARNr	Ácido ribonucleico ribosomal
ASV	*Amplicon sequence variant* (Variante de secuencia de amplicón)
FAO	*Food and Agriculture Organization* (Organización de las Naciones Unidas para la Alimentación y la Agricultura)
FAPROTAX	*Functional Annotation of Prokaryotic Taxa* (Anotación Funcional de Taxones Procariotas)
IMARPE	Instituto del Mar del Perú
MAFFT	*Multiple Alignment using Fast Fourier Transform* (Alineamiento múltiple mediante transformada rápida de Fourier)
MINAM	Ministerio del Ambiente
OTU	*Operational Taxonomic Unit* (Unidad taxonómica operacional)
pb	pares de bases
PCoA	*Principal coordinates analysis* (Análisis de coordenadas principales)
PCR	*Polymerase Chain Reaction* (Reacción en cadena de la polimerasa)
QIIME2	*Quantitative Insights Into Microbial Ecology 2* (Perspectivas cuantitativas sobre la ecología microbiana 2)
SERNANP	Servicio Nacional de Áreas Naturales Protegidas por el Estado
UPGMA	*Unweighted pair group method with arithmetic mean* (Método de agrupamiento de pares no ponderados con media aritmética)

ÍNDICE GENERAL

DEDICATORIA i
AGRADECIMIENTOS ii
ÍNDICE GENERAL iv
ÍNDICE TABLAS vi
ÍNDICE DE FIGURAS viii
RESUMEN x
ABSTRACT xi
INTRODUCCIÓN 1
MARCO TEÓRICO 3
2.1. Diversidad bacteriana en ecosistemas lacustres 3
2.2. ADN metabarcoding para el estudio de la diversidad bacteriana 5
2.3. El Lago Titicaca: Ecología y biodiversidad 9
2.4. Impacto de la acuicultura en la microbiota del Lago Titicaca 10
HIPÓTESIS 13
OBJETIVOS 13
4.1. Objetivo general 13
4.2. Objetivos específicos 13
MATERIALES Y MÉTODOS 14
5.1. Material biológico 14
5.2. Extracción de ADN 16
5.3. Preparación de librerías y secuenciación 17
5.4. Análisis bioinformático 18
5.4.1. Flujo de trabajo 18
5.4.2. Eliminación de adaptadores 20
5.4.3. Filtrado de las secuencias e inferencia de ASVs 22
5.4.4. Anotación taxonómica 23
5.4.5. Análisis de diversidad 24
5.4.6. Análisis de diversidad funcional 26

RESULTADOS 28
6.1. Caracterización de los datos de secuenciación 28
6.2. Abundancia de ASVs y OTUs 31
6.2.1. Estructura de las comunidades bacterianas 33
6.3. Análisis de la diversidad 42
6.3.1. Alfa-rarefacción 42
6.3.2. Diversidad alfa 42
6.3.3. Diversidad beta 45
6.4. Predicción funcional 47
DISCUSIÓN 54
CONCLUSIONES 67
RECOMENDACIONES 68
REFERENCIAS BIBLIOGRÁFICAS 69
ANEXOS 88

ÍNDICE TABLAS

Tabla 1. Coordenadas de los puntos de muestreo en la cuenca hidrográfica del Lago Titicaca, indicando el número (n) de muestras en total (entre 2018 y 2019) obtenidos por localidad para este estudio. LS = Latitud Sur; LO = Longitud Oeste 15

Tabla 2. Primers con adaptadores incluidos, utilizados en la amplificación de la región variable V3-V4 del gen 16S. En negritas se indica la secuencia de adaptadores. 17

Tabla 3. Conteo de lecturas antes y después del filtrado. 28

Tabla 4. Estadísticas del proceso de filtrado de lecturas por muestra (103,186 lecturas filtrados) (I: época lluviosa, II: época seca) 30

Tabla 5. Número de ASVs por muestra obtenido en cinco localidades de la cuenca del Lago Titicaca (I: época lluviosa, II: época seca)..... 32

Tabla 6. Número de niveles taxonómicos encontrados por localidad en el Lago Titicaca y su cuenca hidrográfica. 34

Tabla 7. Órdenes y familias únicos por muestras de agua recolectadas en cinco localidades de la cuenca del Lago Titicaca, entre 2018 y 2019. 39

Tabla 8. Géneros únicos por localidad registrados en las 5 localidades evaluadas en la cuenca del lago Titicaca, entre 2018 y 2019. 40

Tabla 9. Índices de diversidad alfa por localidad, año y época (lluviosa (abril), seca (agosto)). 44

Tabla 10. Porcentaje de funciones bacterianas predichas presente en las muestras analizadas en cinco localidades de la cuenca del Lago Titicaca. 48

Tabla 11. Familias y géneros con funciones ecológicas en las localidades de Arapa y Lagunillas (Cuenca del Lago Titicaca). 50

Tabla 12. Familias y géneros con funciones ecológicas en las localidades de Yunguyo, Ichu y Cachipucara..... 52

ÍNDICE DE FIGURAS

Figura 1. Lugares de toma de muestras de agua en piscigranjas de trucha arcoíris ubicados en la cuenca hidrográfica del Lago Titicaca. 14

Figura 2. Descripción general del flujo de trabajo en QIIME2 (Bolyen et al., 2019). 19

Figura 3. a) Gráfico de cajas y bigotes de calidad de las lecturas sentido (QIIME2). b) Gráfico de cajas y bigotes de calidad de las lecturas antisentido (QIIME2). 21

Figura 4. Histograma de calidad de las lecturas sentido y antisentido (MultiQC). 22

Figura 5. Diagrama de Venn del número de ASVs registrados por localidad, compartidos y únicos para las 5 localidades muestreadas en la cuenca del Lago Titicaca. En paréntesis se indica el número total de ASVs por localidad. 33

Figura 6. Gráfico de barras de órdenes con un mínimo de 5% de abundancia. 37

Figura 7. Gráfico de barras de familias con un mínimo de 5% de abundancia. 38

Figura 8. Diagrama de Venn de los órdenes por localidad registrados en las 5 localidades evaluadas en la cuenca del lago Titicaca, entre 2018 y 2019. 41

Figura 9. Diagrama de Venn de las familias por localidad registrados en las 5 localidades evaluadas en la cuenca del lago Titicaca, entre 2018 y 2019. 41

Figura 10. Rarefacción de las muestras con umbral de rarefacción 5000 lecturas. 42

Figura 11. Gráfico de cajas y bigotes del índice de diversidad de Shannon por localidad muestreada en la cuenca del lago Titicaca entre 2018 y 2019. 45

Figura 12. PCoA por localidad en 2018 y 2019 en época lluviosa y seca. 46

Figura 13. Árbol de agrupamiento por UPGMA del nivel taxonómico familia. 47

Figura 14. Mapa de calor (heatmap) de la predicción funcional de cada muestra en el 2018 y 2019 por época lluviosa (verde) y seca (marrón) 49

Figura S1. Esquema de las regiones hipervariables del gen ARNr 16S de procariotas (Fuente: Yang et al., 2016). 88

Figura S2. Gráfico de barras del número de lecturas de los datos de secuenciación 89

Figura S3. Filograma de la taxonomía de todas las muestras (MEGAN6) 90

Figura S4. Gráfico de burbujas de los géneros bacterianos en todas las muestras (MEGAN6) .. 91

Figura S5. Código en R para generar gráfico de barras 94

Figura S6. Código en R para generar Diagrama de Venn 95

Figura S7. Código en R para generar gráfico de cajas y bigotes 96

Figura S8. Código en R para generar Análisis de Coordenadas Principales 97

Figura S9. Código en R para generar mapa de calor 98

RESUMEN

La acuicultura de trucha arcoíris (*Oncorhynchus mykiss*) en el Lago Titicaca es crucial para las economías locales, pero su expansión en jaulas flotantes ha generado preocupaciones sobre su impacto ambiental, particularmente por la alteración de la diversidad microbiana. El presente estudio tuvo como objetivo caracterizar la variación espacio-temporal de las comunidades bacterianas en cinco áreas de cultivo de la cuenca del Lago Titicaca (Lago Arapa, Laguna Lagunillas, Ichu, Cachipucara y Yunguyo), durante las épocas lluviosa y seca de 2018 y 2019, mediante ADN metabarcoding de la región V3-V4 del gen ARNr 16S. Tras la secuenciación Illumina MiSeq y el procesamiento de los datos con QIIME2 y DADA2, se identificaron 875 ASVs (*Amplicon Sequence Variants*) de alta confiabilidad, con una dominancia de los filos Bacteroidota, Acidobacteriota, Cyanobacteria, Verrucomicrobiota, Actinobacteriota y Proteobacteria. A nivel espacial, Arapa y Lagunillas presentaron los mayores valores de diversidad alfa (Shannon, Simpson y Faith's PD) y mayor número de ASVs exclusivos, evidenciando alta riqueza microbiana, mientras que Ichu, Cachipucara y Yunguyo mostraron comunidades más homogéneas, siendo Cachipucara la de menor diversidad. A nivel temporal, la estacionalidad fue un factor determinante, con mayor riqueza de ASVs y diversidad de Shannon durante la época seca, especialmente en Arapa y Lagunillas. En conjunto, los análisis de diversidad beta (PCoA y UPGMA) evidenciaron patrones espacio-temporales asociados a la estacionalidad y la conectividad hidrológica. Asimismo, se identificaron perfiles taxonómicos diferenciados entre localidades, destacando bacterias con potencial patogénico en Lagunillas, taxones vinculados a aportes orgánicos ganaderos en Ichu, alta riqueza en Arapa y comunidades asociadas a estrés ambiental en Yunguyo y Cachipucara. La predicción funcional mediante FAPROTAX mostró un bajo porcentaje de asignación (33.6%), reflejando el carácter subexplorado del lago; no obstante, se identificaron funciones asociadas a eutrofización en Arapa y Yunguyo, presencia de potenciales patógenos en Ichu, Lagunillas y Yunguyo, y menor diversidad funcional en Cachipucara. En conjunto, estos resultados evidencian que las comunidades bacterianas responden a la interacción de factores espaciales, temporales y antrópicos, resaltando el impacto de la acuicultura y el valor del ADN metabarcoding para el monitoreo y manejo sostenible de ecosistemas altoandinos.

Palabras clave: ASVs, diversidad alfa, diversidad beta, predicción funcional, acuicultura.

ABSTRACT

Rainbow trout (*Oncorhynchus mykiss*) aquaculture in Lake Titicaca is crucial for local economies; however, its expansion in floating cages has raised concerns about its environmental impact, particularly regarding alterations in microbial diversity. The present study aimed to characterize the spatiotemporal variation of bacterial communities in five aquaculture areas within the Lake Titicaca basin (Lake Arapa, Laguna Lagunillas, Ichu, Cachipucara, and Yunguyo) during the rainy and dry seasons of 2018 and 2019, using DNA metabarcoding of the V3–V4 region of the 16S rRNA gene. Following Illumina MiSeq sequencing and bioinformatic processing with QIIME2 and DADA2, a total of 875 high-confidence ASVs (Amplicon Sequence Variants) were identified, with dominance of the phyla Bacteroidota, Acidobacteriota, Cyanobacteria, Verrucomicrobiota, Actinobacteriota, and Proteobacteria. At the spatial level, Arapa and Lagunillas exhibited the highest alpha diversity values (Shannon, Simpson, and Faith's PD) and the greatest number of exclusive ASVs, indicating high microbial richness, whereas Ichu, Cachipucara, and Yunguyo showed more homogeneous communities, with Cachipucara presenting the lowest diversity. Temporally, seasonality was a key determinant, with higher ASV richness and Shannon diversity during the dry season, particularly in Arapa and Lagunillas. Beta diversity analyses (PCoA and UPGMA) revealed spatiotemporal patterns associated with seasonality and hydrological connectivity. In addition, distinct taxonomic profiles were identified among localities, including bacteria with pathogenic potential in Lagunillas, taxa linked to livestock-derived organic inputs in Ichu, high richness in Arapa, and communities associated with environmental stress in Yunguyo and Cachipucara. Functional prediction using FAPROTAX showed a low assignment rate (33.6%), highlighting the underexplored nature of the system; however, functions related to eutrophication were identified in Arapa and Yunguyo, potential pathogenic functions in Ichu, Lagunillas, and Yunguyo, and lower functional diversity in Cachipucara. Overall, these results demonstrate that bacterial communities respond to the interaction of spatial, temporal, and anthropogenic factors, highlighting the impact of aquaculture and the value of DNA metabarcoding as an effective tool for monitoring and sustainable management of high Andean aquatic ecosystems.

Keywords: ASVs, alpha diversity, beta diversity, functional prediction, aquaculture.

INTRODUCCIÓN

El Lago Titicaca es uno de los ecosistemas más importantes en Sudamérica por su alta biodiversidad y relevancia económica, al sustentar actividades como la pesca y la acuicultura (Ministerio del Ambiente (MINAM), 2014; Lima et al., 2024). Sin embargo, en las últimas décadas ha enfrentado una creciente presión ambiental debido a la expansión del cultivo en jaulas flotantes de *Oncorhynchus mykiss* (trucha arcoíris). Este sistema aunque representa una fuente de ingresos para las comunidades locales (Montesinos, 2017), también incrementa la carga de materia orgánica y nutrientes, favoreciendo procesos de eutrofización que alteran la diversidad microbiana (Amaru et al., 2024; Suzuki et al., 2024; Zhang et al., 2023).

Las comunidades microbianas cumplen funciones clave en los ecosistemas acuáticos, como la transformación de nutrientes y la degradación de materia orgánica, además de actuar como bioindicadores del estado ecológico frente a impactos antrópicos como la acuicultura (Garner et al., 2023; Shen et al., 2024). La intensificación de esta actividad puede modificar la estructura bacteriana al aumentar compuestos nitrogenados y materia orgánica, favoreciendo bacterias heterótrofas y potencialmente patógenas (Milyanchuk et al., 2024; Zhang et al., 2023). A pesar de la relevancia de las comunidades microbianas, los estudios sobre su diversidad en el Lago Titicaca aún son limitados, lo que dificulta comprender completamente los efectos de las actividades antropogénicas sobre la estructura bacteriana (Maquera et al., 2024).

Asimismo, la variabilidad estacional en el Lago Titicaca influye en la dinámica bacteriana. Durante la época de lluvias, el escurrimiento y la remoción de sedimentos alteran la composición microbiana, mientras que en la estación seca la menor renovación del agua y la acumulación de

materia orgánica favorecen procesos de eutrofización y degradación anaeróbica (Shen et al., 2024; Milyanchuk et al., 2024). Estas dinámicas resaltan la necesidad de evaluar los cambios espacio-temporales en la diversidad bacteriana para una gestión sostenible de la acuicultura en el lago.

Uno de los desafíos para evaluar el impacto de la acuicultura en ecosistemas acuáticos es la falta de métodos eficientes para caracterizar la diversidad microbiana con alta resolución. En este contexto, el ADN metabarcoding basado en regiones variables del gen ARNr 16S permite identificar bacterias sin necesidad de cultivo, a partir de muestras ambientales, facilitando el análisis de la composición y diversidad microbiana (López-Aladid et al., 2023; Stoeck et al., 2018; Suzuki et al., 2024). Este enfoque consiste en la amplificación y secuenciación de dichas regiones, seguida de un análisis bioinformático que incluye el filtrado de lecturas (*reads*), la inferencia de ASVs y la asignación taxonómica en OTUs (*Operational Taxonomic Units*) mediante bases de datos de referencia. Además, permite evaluar variaciones espacio-temporales en las comunidades bacterianas y su relación con factores ambientales como la eutrofización, la contaminación por materia orgánica y las condiciones físico-químicas del agua (Rieder et al., 2023; Shen et al., 2024).

En este sentido, el presente estudio tuvo como objetivo caracterizar la variación espacio-temporal de la diversidad bacteriana en cinco áreas de cultivo de *O. mykiss* en la cuenca hidrográfica del Lago Titicaca. Así, utilizando ADN metabarcoding basado en las regiones V3-V4 del gen 16S, se analizaron muestras de agua de piscigranjas obtenidas en las épocas seca y lluviosa, entre 2018 y 2019. La información generada puede contribuir en comprender la dinámica microbiana en el lago. Además, puede ser aplicada en el desarrollo de estrategias de monitoreo ambiental que ayuden a minimizar los impactos negativos de la acuicultura y a promover un manejo más sostenible de este ecosistema.

MARCO TEÓRICO

2.1. Diversidad bacteriana en ecosistemas lacustres

La diversidad bacteriana en los ecosistemas acuáticos es fundamental para el funcionamiento de los mismos, ya que desempeñan roles esenciales en varios procesos biogeoquímicos, tales como la descomposición de materia orgánica, el ciclo del nitrógeno, la mineralización de nutrientes y la regulación del pH del agua (Logue et al., 2008). Estas comunidades microbianas incluyen una amplia variedad de taxones, que contribuyen a la productividad primaria y mantienen el equilibrio ecológico de los ecosistemas acuáticos, lo que las convierte en indicadores clave de la salud ambiental de estos sistemas (Garner et al., 2023). En lagos y otros cuerpos de agua dulce, las comunidades bacterianas suelen estar dominadas por grupos como Proteobacteria, Cyanobacteria, Actinobacteria, Bacteroidota, y Verrucomicrobia (Logue et al., 2008). Estos filos tienen funciones específicas y complementarias que permiten la descomposición de compuestos orgánicos y la disponibilidad de nutrientes esenciales para otros organismos acuáticos. Las Cyanobacteria, por ejemplo, son conocidas por su capacidad de fijar nitrógeno atmosférico, un proceso crucial en lagos con limitaciones de nutrientes (Garner et al., 2023). Por otro lado, las Actinobacteria y Bacteroidota desempeñan roles clave en la descomposición de materia orgánica y la mineralización de compuestos orgánicos complejos.

Los factores fisicoquímicos del agua, como la temperatura, el pH, el oxígeno disuelto, y la concentración de nutrientes (p. ej., nitrógeno y fósforo) (Shen et al., 2024), influyen en la estructura y dinámica de las comunidades bacterianas en lagos y otros cuerpos de agua. Estos determinan la disponibilidad de hábitats para distintos grupos bacterianos y afectan su abundancia y diversidad.

En ambientes eutróficos, debido a la acuicultura, se observa una mayor abundancia de bacterias heterótrofas que descomponen los nutrientes orgánicos (Milyanchuk et al., 2024).

La variabilidad estacional también juega un papel crucial en la dinámica de las comunidades bacterianas. En cuerpos de agua como el Lago Titicaca, las estaciones de lluvias y sequías inducen fluctuaciones significativas en las condiciones ambientales, como la turbidez del agua y la disponibilidad de nutrientes (Pouilly et al., 2014). Durante la temporada de lluvias, el escurrimiento de sedimentos y la remoción de microorganismos acumulados en el fondo del lago alteran la composición bacteriana en la columna de agua, lo que puede generar un cambio en las interacciones microbianas y en la estructura de las comunidades (Shen et al., 2024).

Además, la introducción de especies invasoras, sumado a las actividades humanas como la acuicultura intensiva, tienen un impacto directo sobre la microbiota. En el Lago Titicaca, la acuicultura de trucha arcoíris ha aumentado la carga de nutrientes en el agua, lo que favorece la proliferación de microorganismos patógenos y altera la diversidad bacteriana al modificar los ciclos de nutrientes y el equilibrio de los grupos bacterianos (Amaru et al., 2024; Suzuki et al., 2024).

Por otro lado, diversos estudios han demostrado que las comunidades bacterianas pueden reflejar los efectos de la contaminación por metales pesados, el exceso de nutrientes y el cambio climático (Custodio et al., 2022). Los estudios de Shafi y colaboradores (2017) en ecosistemas de agua dulce determinaron que la diversidad y dinámica del bacterioplancton están fuertemente condicionadas por factores ambientales y antrópicos, como la disponibilidad de nutrientes, la hidrología, la temperatura y la contaminación, lo que puede generar cambios en la estructura

comunitaria y favorecer la dominancia de especies tolerantes. Por ejemplo, Garner y colaboradores (2023) encontraron que en lagos afectados por la minería y la contaminación orgánica los taxones bacterianos pueden variar en función de los niveles de contaminantes, y algunos grupos microbianos pueden emplearse como bioindicadores de la calidad del agua y del estado ecológico del sistema. En particular, se ha evidenciado que los metales pesados pueden modificar significativamente la diversidad y composición microbiana, afectando negativamente la diversidad alfa y alterando funciones clave del ecosistema, como los ciclos del carbono y del nitrógeno, lo que ha permitido proponer taxones sensibles como indicadores de contaminación ambiental (Li et al., 2020). Asimismo, Stabili y colaboradores (2021) evaluaron en sistemas acuícolas que la calidad microbiológica del agua mediante enfoques integrados, incluyendo metabarcoding del gen ARNr 16S, ha permitido identificar comunidades bacterianas asociadas a contaminación y presencia de patógenos, destacando su utilidad en el monitoreo ambiental y en la gestión sostenible de estos ecosistemas.

2.2. ADN metabarcoding para el estudio de la diversidad bacteriana

El ADN metabarcoding es una técnica molecular de alta resolución que ha transformado la manera en que se estudian las comunidades microbianas. Esta metodología, basada en la secuenciación masiva de regiones específicas permite una caracterización detallada y precisa de la diversidad bacteriana sin necesidad de su cultivo. A diferencia de los métodos tradicionales, que dependen del aislamiento y cultivo de microorganismos, el ADN metabarcoding permite identificar no solo las especies bacterianas cultivables, sino también aquellas que no se pueden cultivar en condiciones de laboratorio, proporcionando una visión mucho más completa y representativa de la microbiota del ecosistema (Klindworth et al., 2013). Este enfoque se basa en

la obtención de secuencias de ciertas regiones del ADN bacteriano a partir de muestras ambientales, las cuales son posteriormente procesadas mediante herramientas bioinformáticas para inferir unidades biológicas representativas de la diversidad microbiana.

El gen ARNr 16S es un marcador utilizado para la identificación de comunidades bacterianas, debido a su presencia en casi todas las bacterias y arqueas, su conservación en muchas regiones y variabilidad en otras regiones hipervariables (Lin & Ju, 2023). Estas características lo convierten en una herramienta idónea para la identificación taxonómica a nivel de familia y género (López-Aladid et al., 2023). Este gen, de aproximadamente 1500 pb, está conformado por regiones altamente conservadas intercaladas con nueve regiones hipervariables (V1–V9) (Figura S1), lo que permite el diseño de primers universales dirigidos a las zonas conservadas, mientras que las regiones variables aportan la resolución necesaria para discriminar entre distintos taxones bacterianos (Ziesemer et al., 2015). Esta combinación de conservación y variabilidad ha permitido que el gen 16S ARNr sea ampliamente utilizado en estudios metataxonómicos, posibilitando una clasificación taxonómica cercana al nivel de especie en muchos casos.

Entre las regiones más utilizadas en estudios de ADN metabarcoding se encuentran V1–V3, V3–V4 y V4, debido a su adecuado balance entre longitud de secuencia, resolución taxonómica y compatibilidad con plataformas de secuenciación (Lee et al., 2023; Lin & Ju, 2023). La elección de la región depende del diseño experimental, ya que cada región presenta sesgos en la amplificación y diferencias en su capacidad de discriminación taxonómica. Por ejemplo, la región V4 es ampliamente utilizada por su alta reproducibilidad y cobertura, mientras que la región V3 presenta alta variabilidad y buena resolución taxonómica, aunque puede estar sujeta a sesgos asociados a la longitud del amplicón y eficiencia de amplificación (Ziesemer et al., 2015). En este

contexto, la combinación de las regiones V3–V4 representa una estrategia ampliamente adoptada, ya que proporciona un equilibrio entre resolución taxonómica y cobertura de diversidad bacteriana en muestras ambientales complejas, como las del Lago Titicaca (Zinger et al., 2011). Así, la secuenciación de estas regiones hipervariables ha permitido generar perfiles taxonómicos completos y detallados de comunidades bacterianas en diversos ecosistemas acuáticos (Requena-Castro et al., 2025; Stoeck et al., 2018).

El ADN metabarcoding implica una serie de etapas que permiten caracterizar la diversidad microbiana (Łach et al., 2024; Regueira-Iglesias et al., 2023). Inicialmente, se realiza la amplificación por PCR de regiones específicas del gen ARNr 16S, empleando cebadores dirigidos a regiones hipervariables, como V3–V4. Los productos amplificados son posteriormente utilizados para la construcción de librerías de secuenciación, etapa crítica que permite adaptar los fragmentos de ADN a la plataforma de secuenciación. En protocolos basados en Illumina, los cebadores incluyen secuencias "*overhang*" que permiten la posterior adición de adaptadores de secuenciación y códigos de identificación (índices) mediante una PCR adicional de ciclos limitados, facilitando la multiplexación de múltiples muestras en una misma corrida (Illumina, 2013).

Posteriormente, las librerías son purificadas, cuantificadas y normalizadas antes de ser agrupadas (pooling) para su secuenciación. En la plataforma Illumina MiSeq, la secuenciación se basa en la tecnología de "*sequencing by synthesis*", generando lecturas pareadas (*paired-end*), comúnmente de 2 × 300 pb, lo que permite solapar las secuencias y reconstruir el amplicón completo de la región V3–V4 (~460 pb) con alta precisión. Esta plataforma puede generar millones de lecturas por corrida, alcanzando una alta cobertura por muestra y permitiendo detectar incluso taxones de baja abundancia, lo que la convierte en una herramienta adecuada para estudios de

diversidad microbiana (Illumina, 2013; Illumina, 2024). A continuación, los datos obtenidos son procesados mediante análisis bioinformáticos, que incluyen el filtrado de calidad para eliminar secuencias erróneas o de baja calidad, la eliminación de quimeras y la corrección de errores de secuenciación. Posteriormente, las secuencias son agrupadas o inferidas en unidades representativas de la diversidad microbiana. En enfoques tradicionales, estas se agrupan en OTUs según un umbral de similitud, mientras que métodos más recientes permiten la inferencia de ASVs, las cuales representan secuencias únicas con resolución a nivel de un nucleótido (Fasolo et al., 2024). Finalmente, estas unidades son comparadas con bases de datos de referencia (p. ej. Silva Database, Greengenes, NCBI) para realizar la asignación taxonómica, permitiendo determinar la composición bacteriana de las muestras analizadas.

El uso de ADN metabarcoding ofrece varias ventajas clave sobre los métodos tradicionales de estudio de la microbiota, como el cultivo bacteriano. Mientras que los métodos de cultivo se limitan a las bacterias que pueden crecer en condiciones de laboratorio, el ADN metabarcoding permite identificar todas las bacterias presentes en una muestra ambiental, incluidas aquellas que no pueden ser cultivadas (Bukin et al., 2019).

Esta capacidad ha permitido caracterizar con mayor precisión la diversidad bacteriana y su variación espacial en ecosistemas lacustres, evidenciando que la composición microbiana puede diferir significativamente incluso dentro de un mismo lago, en función de condiciones ambientales locales (Schallenberg et al., 2021). Estudios en múltiples lagos como los de Sadeghi y colaboradores (2021) han demostrado que la estructura de las comunidades bacterianas está determinada tanto por factores abióticos (como nutrientes y condiciones fisicoquímicas) como por interacciones bióticas y procesos espaciales, reflejando la complejidad de los ensamblajes

microbianos en estos sistemas. Mou y colaboradores (2013) observaron mediante ADN metabarcoding que las comunidades bacterianas pueden variar entre fracciones de bacterioplancton, así como en respuesta a gradientes ambientales, lo que resalta la heterogeneidad de estos ecosistemas. De igual forma, estudios recientes como los de Skopina y colaboradores (2024) han evidenciado que la diversidad, estructura y potencial funcional de las comunidades bacterianas pueden estar asociadas a procesos como la eutrofización, la entrada de materia orgánica y la degradación de contaminantes, constituyéndose en indicadores sensibles del estado ecológico del sistema. En conjunto, estos enfoques confirman que el ADN metabarcoding es una herramienta fundamental para caracterizar la diversidad bacteriana en ecosistemas acuáticos, como el Lago Titicaca, donde la microbiota está fuertemente influenciada por factores ambientales y actividades antrópicas, como la estacionalidad y la acuicultura.

2.3. El Lago Titicaca: Ecología y biodiversidad

El Lago Titicaca, ubicado en el Altiplano andino a una altitud de 3,812 m.s.n.m., es reconocido como el lago navegable más alto del mundo y el cuerpo de agua dulce más grande de Sudamérica. Abarca una superficie aproximada de 8,400 km², de los cuales el 56% se encuentra en territorio peruano y el 44% en Bolivia. Esta característica geográfica lo convierte en un ecosistema único, que no solo sustenta una rica biodiversidad, sino que también tiene un profundo valor cultural y económico para las comunidades locales (MINAM, 2014).

El Lago Titicaca y sus alrededores albergan una gran cantidad de especies endémicas, tanto en fauna como en flora. Destacan diversas especies de peces del género Orestias y aves acuáticas migratorias, que dependen de este ecosistema para su supervivencia. Además, los totorales,

representados por *Schoenoplectus tatora*, juegan un papel crucial en la estabilización de las orillas y la filtración de contaminantes, contribuyendo a la regulación ecológica del lago (Villwock, 1993; Ramos et al., 2019). La biodiversidad del lago no solo es importante desde una perspectiva ecológica, sino también desde el punto de vista socioeconómico, ya que sustenta las actividades de pesca y acuicultura en la región (Servicio Nacional de Áreas Protegidas por el Estado (SERNANP), 2019).

La crianza de trucha arcoíris (*O. mykiss*) en jaulas flotantes ha incrementado considerablemente en los últimos años, generando preocupación sobre sus efectos en la calidad del agua y en la biodiversidad microbiana del lago (Montesinos, 2017; Amaru et al., 2024). Es conocido que las prácticas acuícolas pueden incrementar la eutrofización, un proceso en el que el aumento de nutrientes favorece el crecimiento descontrolado de microorganismos, alterando las condiciones ecológicas y favoreciendo la proliferación de bacterias potencialmente patógenas (Milyanchuk et al., 2024).

2.4. Impacto de la acuicultura en la microbiota del Lago Titicaca

La acuicultura de *O. mykiss* en el Lago Titicaca, especialmente en áreas como la Bahía de Puno, la Bahía Interior de Puno, Chucuito, Pomata y Juli, ha experimentado un crecimiento significativo en las últimas décadas (Ocola Villasante, 2022). Estas zonas concentran una alta densidad de piscigranjas, lo que ha generado un aumento en la carga de nutrientes (como nitrógeno y fósforo) y materia orgánica en el ecosistema lacustre, provenientes de los desechos de los peces y alimentos no consumidos. Este enriquecimiento excesivo favorece la proliferación de bacterias heterótrofas y patógenas, alterando la composición y funcionalidad de las comunidades bacterianas

locales (Orna & Mamani, 2023; Pouilly et al., 2014). Esto ha resultado en una disminución de la oxigenación del agua y en la proliferación de microorganismos anaeróbicos. Además, se ha observado un aumento en la presencia de bacterias potencialmente patógenas, como *Aeromonas* y *Yersinia ruckeri*, que afectan tanto a la trucha como a especies nativas del lago (Pouilly et al., 2014).

Estudios previos han evidenciado que las presiones antrópicas, incluyendo la acuicultura y las descargas de origen doméstico, constituyen fuentes importantes de contaminación en el Lago Titicaca, incrementando significativamente la carga de nutrientes y materia orgánica en el sistema (Siguayro et al., 2022). Este enriquecimiento ha sido asociado a condiciones eutróficas e incluso hipertróficas en diversas zonas del lago, como la Bahía Interior de Puno y Yunguyo, lo que repercute directamente en la calidad del agua y en la dinámica de las comunidades biológicas. En estudios de sistemas acuáticos, Findlay y colaboradores (2009) determinaron que el incremento sostenido de nutrientes derivado de la acuicultura puede generar efectos acumulativos, como aumentos en la biomasa fitoplanctónica y alteraciones en la estructura bacteriana, evidenciando cambios progresivos en el funcionamiento del ecosistema.

En el Lago Titicaca, estudios realizados en zonas como Chucuito, Pomata y Juli han mostrado que la composición bacteriana está estrechamente relacionada con factores ambientales como la temperatura, la turbidez del agua y la disponibilidad de nutrientes, los cuales son influenciados tanto por la estacionalidad como por las actividades acuícolas (Chura & Mollocondo, 2009; Orna & Mamani, 2023). Durante la temporada de lluvias, el escurrimiento superficial transporta nutrientes acumulados hacia el lago, generando cambios en la estructura de las comunidades bacterianas. A nivel microbiano, investigaciones como las de Qin y colaboradores

(2020) han demostrado que la acuicultura puede modificar significativamente la diversidad y composición bacteriana, tanto en la columna de agua como en sedimentos, donde factores como el fósforo total juegan un rol clave en el ensamblaje de las comunidades. Asimismo, se ha observado que los sistemas acuícolas tienden a favorecer comunidades dominadas por bacterias heterótrofas y potencialmente patógenas, asociadas a la acumulación de materia orgánica y a cambios en la calidad del agua (Clols-Fuentes et al., 2024). En conjunto, estos estudios evidencian que la interacción entre factores ambientales, estacionales y antrópicos modula la estructura y función de las comunidades microbianas, resaltando la importancia de enfoques espacio-temporales para comprender las dinámicas bacterianas en el Lago Titicaca.

HIPÓTESIS

Existen diferencias en la diversidad temporal (época lluviosa y seca) de comunidades bacterianas asociadas a la distribución de los cuerpos de agua usados para el cultivo de trucha arcoíris, en la cuenca hidrográfica del Lago Titicaca, lo que permite identificar grupos funcionales bacterianos como potenciales bioindicadores de la calidad del agua.

OBJETIVOS

4.1. Objetivo general

Caracterizar la diversidad espacio-temporal de las comunidades bacterianas en cinco áreas de cultivo de *O. mykiss* en la cuenca hidrográfica del Lago Titicaca, Perú, basado en las regiones hipervariables V3-V4 del gen ARNr 16S.

4.2. Objetivos específicos

- Identificar los OTUs predominantes en las áreas de cultivo de *O. mykiss* en la cuenca hidrográfica del Lago Titicaca (Cachipucara, Yunguyo, Ichu, Laguna Lagunillas y Lago Arapa) utilizando la técnica de ADN metabarcoding basada en las regiones V3-V4 del gen ARNr 16S.
- Comparar la variación espacio-temporal de la composición y diversidad de las comunidades bacterianas en las cinco áreas de cultivo de *O. mykiss* durante las épocas lluviosas y secas, de los años 2018 y 2019.
- Evaluar la diversidad funcional de bacterias y su relación con la eutrofización y otros indicadores de calidad del agua en áreas de cultivo intensivo de *O. mykiss*.

MATERIALES Y MÉTODOS

5.1. Material biológico

Se seleccionaron piscigranjas de truchas arcoíris en cinco localidades (Figura 1), ubicadas en la zona norte (Ichu), centro (Cachipucara) y sur (Yunguyo) del lago Titicaca, y dos zonas de la cuenca hidrográfica (laguna Lagunillas y lago Arapa). Se tomaron muestras de agua en época lluviosa (abril) y seca (agosto), en los años 2018 y 2019 (Tabla 1), en colaboración con el Laboratorio continental de Puno, del Instituto del Mar del Perú, y como parte de las campañas periodicas de muestreo que realiza.



Figura 1. Lugares de toma de muestras de agua en piscigranjas de trucha arcoíris ubicados en la cuenca hidrográfica del Lago Titicaca.

Tabla 1. Coordenadas de los puntos de muestreo en la cuenca hidrográfica del Lago Titicaca, indicando el número (n) de muestras en total (entre 2018 y 2019) obtenidos por localidad para este estudio. LS = Latitud Sur; LO = Longitud Oeste

Zona	Localidad	n	Año	Época		Coordenadas
Cuenca del Lago Titicaca	Arapa	4	2018	I	Lluviosa	15°10´16.636" LS 70°6¨4.869" LO
				II	Seca	
			2019	I	Lluviosa	
				II	Seca	
	Lagunillas	4	2018	I	Lluviosa	15°45´19.730" LS 70°40¨19.490" LO
				II	Seca	
			2019	I	Lluviosa	
				II	Seca	
Lago Titicaca	Ichu	2	2019	I	Lluviosa	15°51´32.180" LS 69°55¨42.628" LO
				II	Seca	
	Cachipucara	2	2019	I	Lluviosa	16°3´21.822" LS 69°24¨48.867" LO
				II	Seca	
	Yunguyo	3	2018	I	Lluviosa	16°14´46.973" LS 69°8¨38.328" LO
			2019	I	Lluviosa	
				II	Seca	

Fuente: (Ministerio de la producción (PRODUCE), 2022)

Dentro de cada piscigranja por localidad, se colectó 1 L de agua en botellas de vidrio estériles, a 1 m de profundidad. Las botellas fueron previamente lavadas con lejía al 10%, enjuagadas con agua destilada y autoclavadas. Luego, las botellas fueron rotuladas y almacenadas en frío en un *cooler*, hasta su transporte al laboratorio continental del IMARPE de Puno.

5.2. Extracción de ADN

El agua fue filtrada al vacío utilizando filtros Sterivex, y posteriormente almacenados a -80°C en el Laboratorio de Genética Molecular del IMARPE. A partir de dichos filtros, se realizaron las extracciones de ADN utilizando el kit PowerWater Sterivex (Qiagen), siguiendo las instrucciones del fabricante. La elución del ADN se realizó utilizando agua libre de nucleasas (NFW). Se incluyeron 3 controles negativos, que consistió en realizar el proceso de extracción utilizando un filtro blanco sin muestra de agua, para garantizar la integridad del procedimiento. La concentración de ADN fue cuantificada utilizando un fluorómetro Qubit 4 (Invitrogen), mientras que la calidad se evaluó con un espectrofotómetro de microvolúmenes Nanodrop OneC (Thermo Fisher Scientific). Finalmente, todas las muestras de ADN fueron normalizadas a una concentración de 30 ng/µL empleando agua NFW, asegurando la uniformidad para los análisis posteriores.

5.3. Preparación de librerías y secuenciación

Se amplificó la región variable V3-V4 del gen ARNr 16S para la caracterización de la diversidad bacteriana, utilizando los primers de Lee et al. (2023) indicados en la Tabla 2. Las reacciones de PCR se realizaron utilizando el kit KAPA HiFi HotStart Ready Mix (Roche), considerando 0.2 µM de cada primer y 1 ng/µL de ADN, en un volumen total de 25 µL de reacción.

Tabla 2. Primers con adaptadores incluidos, utilizados en la amplificación de la región variable V3-V4 del gen 16S. En negritas se indica la secuencia de adaptadores.

Región	Primer	Secuencia (5'- 3')
V3-V4	V3F	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**CCTACGGGNGGCWGCAG
	V4R	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**GACTACHVGGGTATCTAATCC

Las reacciones de amplificación se realizaron en un termociclador ProFlex (Applied Biosystems), con el ciclo de PCR: pre denaturación de 95°C durante 3 min, seguido de 25 ciclos de denaturación a 95°C durante 30 s, 55°C durante 30 s de hibridación y 72°C durante 30 s de extensión, finalizando con 72°C durante 5 min de extensión final. Los productos de amplificación obtenidos fueron evaluados mediante electroforesis en geles de agarosa 1.5 % con tampón TBE (Tris–borato–EDTA) para verificar el tamaño esperado de los amplificados (550 pb).

Posteriormente, los amplicones fueron purificados utilizando perlas magnéticas AMPure XP beads (Beckman Coulter™), y se realizó una reacción de PCR index para el multiplexado, considerando 5 µL de la primera PCR purificada y los index Nextera XT i5 e i7, con el kit KAPA HotStart ReadyMix. La PCR index se realizó utilizando el ciclo de 95°C por 3 min, seguido de 8 ciclos de 95°C por 30 s, 55°C por 30 s y 72°C por 30 s, para finalizar con 72°C por 5 min. Estas

reacciones fueron purificadas utilizando perlas magnéticas y cuantificadas por fluorometría. Finalmente, se normalizaron las concentraciones de las librerías y se juntaron para la secuenciación con TruSeq *paired-end* en la plataforma MiSeq Illumina.

5.4. Análisis bioinformático

Como resultado de la secuenciación *paired-end* se obtuvieron 2 archivos con lecturas de las 18 muestras secuenciadas (incluyendo las 3 muestras control). Uno de esos archivos correspondió a las lecturas sentido y el otro a lecturas antisentido. La obtención de archivos separados por cada muestra indica que las secuencias crudas fueron previamente separadas y etiquetadas por su identificador o SampleID (demultiplexado), por lo cual durante el flujo de trabajo se obvió el paso de demultiplexado.

5.4.1. Flujo de trabajo

Se utilizó como base el flujo de trabajo de la página oficial de QIIME2 (Bolyen et al., 2019), en donde se muestran los pasos que se deben seguir desde la importación de los archivos de secuencias crudas, resultados del secuenciamiento, hasta los análisis de diversidad y taxonomía (Figura 2).

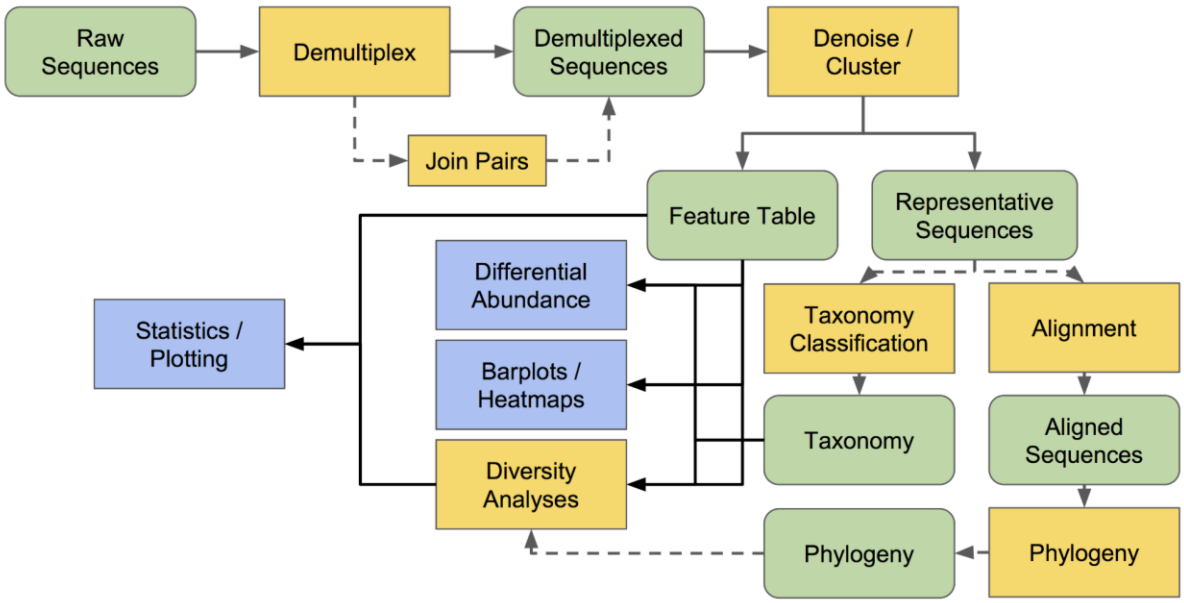


Figura 2. Descripción general del flujo de trabajo en QIIME2 (Bolyen et al., 2019).

Para la visualización de los artefactos de QIIME se requiere de usar un comando para que el artefacto (.qza) se convierta en visualizable (.qzv). Para esto se requiere usar diferentes comandos dependiendo del tipo de artefacto que se ha generado (secuencias, tablas, gráficos de diversidad, árboles taxonómicos, entre otros). En el glosario de la página oficial de QIIME2 (Bolyen et al., 2019) se encuentra la información para convertir los artefactos en visualizables mediante el comando *qiime demux summarize* (Bolyen et al., 2019; Hamday et al., 2008; Hamday & Knight, 2009) y usando el comando *qiime tools view* (Bolyen et al., 2019) se abrirá una ventana del visualizador de QIIME2 con la información del archivo visualizable (.qzv).

5.4.2. Eliminación de adaptadores

Las lecturas obtenidas de la secuenciación contienen en sus extremos las secuencias de los *primers,* además de las secuencias de los adaptadores que son necesarios en la secuenciación, ya que fueron secuenciados junto con los amplicones generados por PCR. Por este motivo, es necesario eliminarlos utilizando el comando *qiime cutadapt trim-paired* (Bolyen et al., 2019; Martin, 2011), que toma como entrada el artefacto sequences.qza junto con las secuencias de los primers sentido y antisentido empleados en las PCRs (Tabla 2).

La ejecución del comando genera el artefacto trimmed_sequences.qza, que contiene las lecturas de las muestras sin las secuencias de los adaptadores. Al convertirlo en un archivo visualizable (trimmed_sequences.qzv), se pueden observar los mismos datos que en el artefacto sequences.qza sobre el número de lecturas por muestra. Sin embargo, este artefacto visualizable permite identificar visualmente, a través de la gráfica interactiva de gráfico de cajas y bigotes, los puntos de truncado en los que la calidad de las lecturas sentido (Figura 3) y antisentido (Figura 4) cae por debajo de Q30 aproximadamente en las posiciones 235 y 205 respectivamente. Se puede revisar la calidad de las lecturas sentido y antisentido de todas las muestras generadas con los programas FastQC (Andrews, S., 2010) y MultiQC (Ewels et al., 2016) (Figura 16).

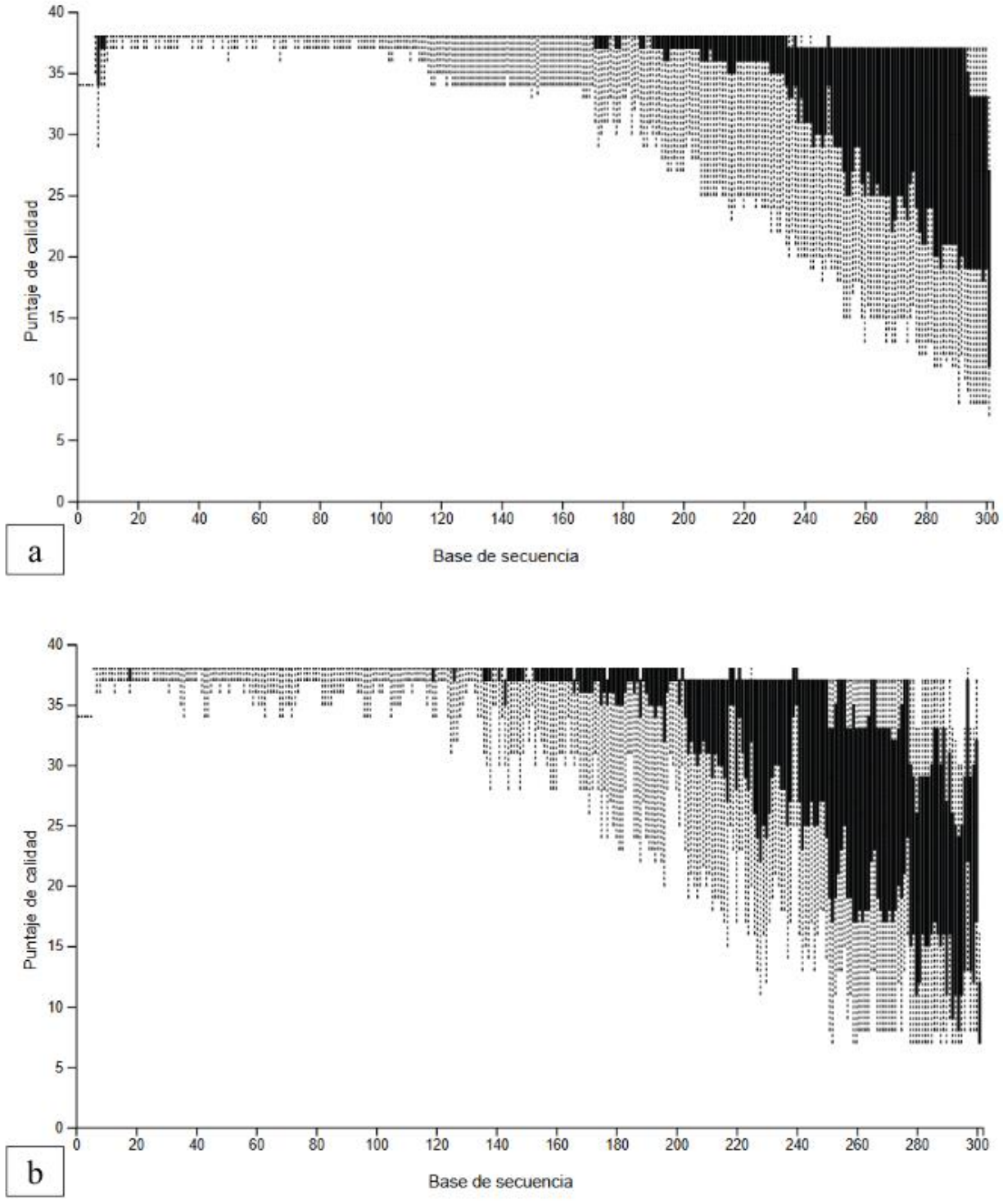


Figura 3. a) Gráfico de cajas y bigotes de la calidad de las lecturas sentido (QIIME2). b) Gráfico de cajas y bigotes de calidad de las lecturas antisentido (QIIME2).

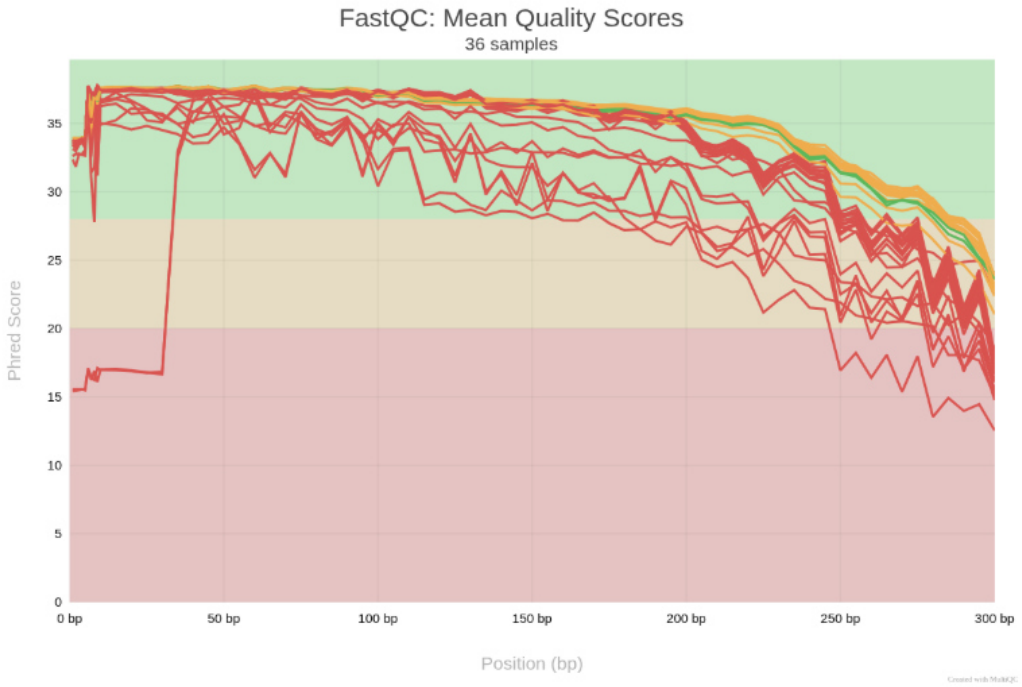


Figura 4. Histograma de calidad de las lecturas sentido y antisentido (MultiQC).

5.4.3. Filtrado de las secuencias e inferencia de ASVs

Para este proceso, empleamos el comando *qiime dada2 denoise-paired* (Bolyen et al., 2019; Callaham et al., 2016), que utiliza el *plugin* DADA2. Este algoritmo aplica un modelo de corrección de errores a las secuencias crudas (R1 y R2), basado en un enfoque probabilístico para la inferencia de variantes de secuencia. El proceso de filtrado y corrección de secuencias se desarrolla en las siguientes etapas:

- Inspección y filtrado de calidad: Se analizan los perfiles de calidad de las secuencias (*Phred score* o *Q-score*) y se eliminan aquellas que no cumplen con los parámetros establecidos, como una puntuación de calidad mínima y una longitud adecuada. Además, si una secuencia

contiene bases ambiguas (N) se elimina, ya que esto indica incertidumbre en la llamada de bases.

- Modelado de tasas de error y *denoising*: Se estima un modelo de tasas de error (*learn error rates*), que posteriormente se aplica a las secuencias filtradas para corregir errores derivados del proceso de secuenciación.

- Deduplicación de secuencias (*dereplication*): Las secuencias idénticas se combinan en una única entrada, conservando su abundancia relativa en la muestra. Reduciendo el tamaño de los archivos y el consumo de memoria.

- Inferencia de ASVs: Se identifican las secuencias únicas presentes en la muestra, tratándolas con una resolución de un solo nucleótido.

- Emparejamiento de lecturas: Se sobrelapan las lecturas sentido (R1) y antisentido (R2) para reconstruir las secuencias originales de los amplicones.

- Eliminación de quimeras: Se detectan y descartan las secuencias quiméricas, que son secuencias falsas formadas por la combinación errónea de fragmentos de ADN durante la amplificación por PCR.

5.4.4. Anotación taxonómica

Para la anotación taxonómica en QIIME2, se utilizó la base de datos ribosomal SILVA 138 (Quast et al., 2012) en la versión 1.4.2 (Bokulich et al., 2018; Robeson et al., 2020). Este archivo es un clasificador taxonómico pre-entrenado que se usará en el comando *qiime feature-classifier*

(Bolyen et al., 2019; Bokulich et al., 2018) para la anotación taxonómica mediante el método *classify-sklearn* (Bolyen et al., 2019; Pedregosa et al., 2011).

5.4.5. Análisis de diversidad

Antes de realizar los análisis de diversidad alfa y beta, se construyó un árbol filogenético a partir de las secuencias representativas (ASVs). Para ello, se realizó un alineamiento de secuencias múltiples (MSA) utilizando el comando *qiime phylogeny* (Bolyen et al., 2019; Price et al., 2010). El proceso comienza con la alineación de los ASVs mediante MAFFT (Alineamiento múltiple mediante transformada rápida de Fourier). Posteriormente, se filtró para eliminar (enmascarar) aquellas columnas de la alineación que no aportan información filogenética o que presentan alineamientos ambiguos. La alineación resultante se empleó para inferir un árbol filogenético (Figura S3), el cual fue enraizado en su punto medio debido a la ausencia de un *outgroup*. Los archivos generados en cada etapa del análisis fueron almacenados, incluyendo tanto la alineación enmascarada como la no enmascarada obtenida con MAFFT, así como las filogenias enraizadas y no enraizadas derivadas de los métodos de inferencia filogenética. La taxonomía asignada sin procesar se puede visualizar en el programa MEGAN6, mediante un gráfico de burbujas (Figura S4).

5.4.5.1. Diversidad alfa

La diversidad alfa, que mide la riqueza de especies, se evaluó mediante análisis de alfa-rarefacción, usando el comando *qiime diversity alpha-rarefaction* (Bolyen et al., 2019), y métricas filogenéticas principales usando el comando *qiime diversity core-metrics-phylogenetic* (Bolyen et al., 2019), calculando los índices de Shannon (H'), Simpson (S) y diversidad filogenética (PD).

Con la ejecución del comando *qiime diversity core-metrics-phylogenetic* se generaron los índices de diversidad alfa y beta (archivos .qza y .qzv).

El proceso de rarefacción consistió en seleccionar un umbral específico de lecturas (punto de rarefacción), que debe ser igual o menor que el número de lecturas de la muestra más pequeña. Luego, se descartaron aleatoriamente las lecturas de las muestras más grandes hasta que todas las muestras restantes tuviesen la misma cantidad de lecturas. Este enfoque permitió calcular métricas de diversidad de manera estandarizada, facilitando la comparación justa entre ecosistemas con diferentes tamaños de muestra (Willis, 2019).

5.4.5.2. Diversidad beta

Los análisis de la diversidad beta se generaron del mismo comando utilizado para los análisis de la diversidad alfa, mediante las métricas filogenéticas principales (usando el comando *qiime diversity core-metrics-phylogenetic*).

La diversidad beta, que mide la disimilitud entre localidades, fueron evaluadas a través de análisis de coordenadas principales (PCoA) y el método de agrupamiento UPGMA (Método de agrupamiento de pares no ponderados con media aritmética). Se utilizó el índice de Jaccard para la gráfica de PCoA, mientras que para el método de agrupamiento UPGMA se requirió convertir la tabla de abundancias de OTUs con su nivel taxonómico y exportarla al formato BIOM (Maldonado, 2021, "Obtener archivos BIOM (MEGAN) y tsv (Excel) que incluyen abundancia y taxonomía") para poder visualizar los datos en el programa MEGAN (Huson et al., 2016).

5.4.6. Análisis de diversidad funcional

La predicción funcional, basada en la tabla de abundancia de OTUs y su taxonomía asociada, se realizó utilizando el programa FAPROTAX versión 1.2.10 (Louca et al., 2016). Esta herramienta está específicamente diseñada para amplicones del gen ARNr 16S (procariontes) y se basa exclusivamente en la información taxonómica como "clave" para compararla con una base de datos de bacterias que contiene registros de funciones metabólicas conocidas. Este enfoque facilita la interpretación, ya que permitió relacionar las comunidades bacterianas con procesos biológicos previamente establecidos. FAPROTAX utiliza la tabla de conteo de OTUs para correlacionar la abundancia de cada categoría funcional presente en cada muestra, permitiendo así la estimación de las funciones ecológicas predominantes.

El programa FAPROTAX requiere una tabla de OTUs o ASVs, de preferencia que sea una tabla rarefaccionada, para asignar la predicción de funciones biológicas. Esta tabla debe estar en un formato de columnas delimitadas por tabulaciones (.tsv) y requiere la información de taxonomía la cual debe estar en la última columna, además el encabezado no debe tener el símbolo #.

Para asignar y agrupar funciones metabólicas a partir de tablas de abundancia de OTUs se usó el script de FAPROTAX *collapse_table.py* (Louca et al., 2016), el cual colapsa los datos taxonómicos en funciones ecológicas basadas en la base de datos FAPROTAX, generando una tabla que se visualizó en QIIME2 y gŕaficas generadas en el programa RStudio v4.5.2 (Posit Team, 2025) usando los paquetes *dplyr* (Wickham et al. 2026), *ggplot2* (Wickham H., 2016), *tidyverse* (Wickham et al. 2019), *readr* (Wickham et al. 2025), *ggthemes* (Arnold J., 2025), *RColorBrewer* (Neuwirth E., 2022), *ggvenn* (Yan L., 2025), *forcats* (Wickham H., 2025) y *conflicted* (Wickham

H., 2023). Los códigos usados para la generación de las gráficas se encuentran en la sección de Anexos (Figura S5, S6, S7, S8 y S9).

RESULTADOS

6.1. Caracterización de los datos de secuenciación

A partir del proceso de secuenciación Illumina MiSeq, se obtuvieron un total de 2,844,826 lecturas (sentido y antisentido) de la región V3-V4 del gen ARNr 16S (Figura S2), de longitud promedio 300 pb, para el total de 18 muestras de agua secuenciadas (Tabla 3). Luego del proceso de filtrado de calidad (>Q30), *denoising*, *dereplication*, eliminación de quimeras y emparejamiento, se obtuvieron un total de 103,186 lecturas (Tabla 4) con un promedio de 5,732.6 lecturas por muestra (Tabla 3).

Tabla 3. Conteo de lecturas antes y después del filtrado.

Lecturas	Total	Longitud promedio (pb)	Media pre-filtrado	Media post-filtrado
sentido	1,422,413	301	79,022.9	5,732.6
antisentido	1,422,413	300		

El método DADA2 permitió realizar el filtrado por calidad, reducción de ruido (*denoising*), eliminación de quimeras, deduplicación e inferencia de ASVs, obteniendo un conjunto de secuencias depuradas y de alta precisión para los análisis posteriores. El valor mínimo de calidad utilizado en el filtrado fue Q30, estableciéndose los puntos de truncado en 235 pb para las lecturas sentido y 205 pb para las antisentido.

Como resultado de la ejecución del comando *qiime dada2 denoise-paired*, se generó una matriz de características (feature_table.qza) con la abundancia de cada ASV por muestra (Tabla

5), junto con las estadísticas del proceso (denoising_stats.qza) (Tabla 4) y un archivo de secuencias representativas (representative-sequences.qza), el cual sirvió como base para la posterior asignación taxonómica. La clasificación taxonómica realizada mediante el método *classify-sklearn*, asignó los taxones a las secuencias representativas utilizando una base de datos de referencia. Estas secuencias fueron agrupadas en grupos taxonómicos denominados OTUs y visualizadas en gráficos de barras mediante el comando *qiime taxa barplot*. Además, los ASVs de las 3 muestras control fueron eliminados de todas las muestras para evitar que posibles contaminaciones durante el proceso de secuenciación afectarán los resultados de las muestras experimentales.

Tabla 4. Estadísticas del proceso de filtrado de lecturas por muestra (103,186 lecturas filtrados) (I: época lluviosa, II: época seca)

Código de muestra			Lecturas iniciales	Lecturas que pasaron el filtrado		N de Lecturas sin ruido (*denoised*)	Lecturas quiméricos		Lecturas no quiméricos	
Localidad	Año	Época	N	N	%		N°	%	N°	%
Arapa	2018	I	90,042	7,927	8.8	7726	7529	8.36	7511	8.34
		II	82,876	6,654	8.03	6434	6199	7.48	6165	7.44
	2019	I	98,187	6,298	6.41	6123	5991	6.1	5979	6.09
		II	81,699	8,972	10.98	8784	8580	10.5	8551	10.47
Lagunillas	2018	I	97,239	7,249	7.45	7047	6775	6.97	6730	6.92
		II	98,805	10,482	10.61	10277	10059	10.18	10004	10.12
	2019	I	61,402	84	0.14	60	54	0.09	54	0.09
		II	85,598	6,409	7.49	6229	6041	7.06	6032	7.05
Ichu	2019	I	84,591	8,388	9.92	8176	7890	9.33	7789	9.21
		II	105,825	9,317	8.8	9048	8770	8.29	8658	8.18
Cachipucara	2019	I	160,829	1,455	0.9	1311	1233	0.77	1233	0.77
		II	96,727	6,947	7.18	6759	6517	6.74	6502	6.72
Yunguyo	2018	I	97,612	7,870	8.06	7625	7361	7.54	7062	7.23
	2019	I	99,162	8,594	8.67	8243	7911	7.98	7857	7.92
		II	80,090	6,509	8.13	6349	6147	7.68	6130	7.65
Control	2018	I	183	4	2.19	2	0	0	0	0
		II	1,422	27	1.9	24	24	1.69	20	1.41
	2019	I y II	124	0	0	0	0	0	0	0

6.2. Abundancia de ASVs y OTUs

Luego del proceso de filtrado de las lecturas, se obtuvieron 875 ASVs efectivos, derivados de secuencias válidas que superaron los filtros de calidad y la eliminación de quimeras. Estos ASVs fueron empleados para la asignación taxonómica y la obtención de OTUs, utilizados posteriormente en los análisis de diversidad y predicción funcional. En la Tabla 5 se presenta el número de ASVs obtenidos por muestra, destacando que la muestra Arapa_Sec_2018_II (Arapa época seca) registró el valor más alto con 227 ASVs, mientras que Lagunillas_Lluv_2019_I (Lagunillas época lluviosa) mostró el valor más bajo con 11 ASVs. El promedio de ASVs en las muestras experimentales (excluyendo los controles) fue de 163. Por otro lado, las muestras control correspondientes a los períodos 2018-I (época lluviosa), 2018-II (época seca) y 2019-I-II (estaciones lluviosa y seca) presentaron en conjunto sólo 4 ASVs.

A partir de la agrupación de todos los ASVs encontrados en las cinco localidades de estudio (Arapa, Lagunillas, Ichu, Cachipucara y Yunguyo), se identificó el número de ASVs compartidos y exclusivos entre localidades, tal como se observa en el diagrama de Venn (Figura 5). La localidad de Arapa presentó el mayor número de ASVs únicos, con 172 ASVs no encontrados en las demás localidades. En contraste, Cachipucara registró el menor número de ASVs exclusivos, con 20 ASVs particulares. Asimismo, se observó que 46 ASVs fueron compartidos entre las cinco localidades. En cuanto al número de ASVs compartidos entre pares de localidades, las mayores similitudes se dieron entre Arapa y Lagunillas, con 42 ASVs comunes, seguidas por Ichu y Lagunillas, con 24 ASVs compartidos.

Tabla 5. Número de ASVs por muestra obtenido en cinco localidades de la cuenca del Lago Titicaca (I: época lluviosa, II: época seca).

Zona	Código de muestra			N° de ASVs
	Localidad	Año	Época	
Cuenca del Lago Titicaca	Arapa	2018	I	219
			II	227
		2019	I	170
			II	170
	Lagunillas	2018	I	198
			II	211
		2019	I	11
			II	153
Lago Titicaca	Ichu	2019	I	180
		2019	II	184
	Cachipucara	2019	I	59
		2019	II	155
	Yunguyo	2018	I	175
		2019	I	197
		2019	II	140
	Control	2018	I	0
		2018	II	4
		2019	I y II	0

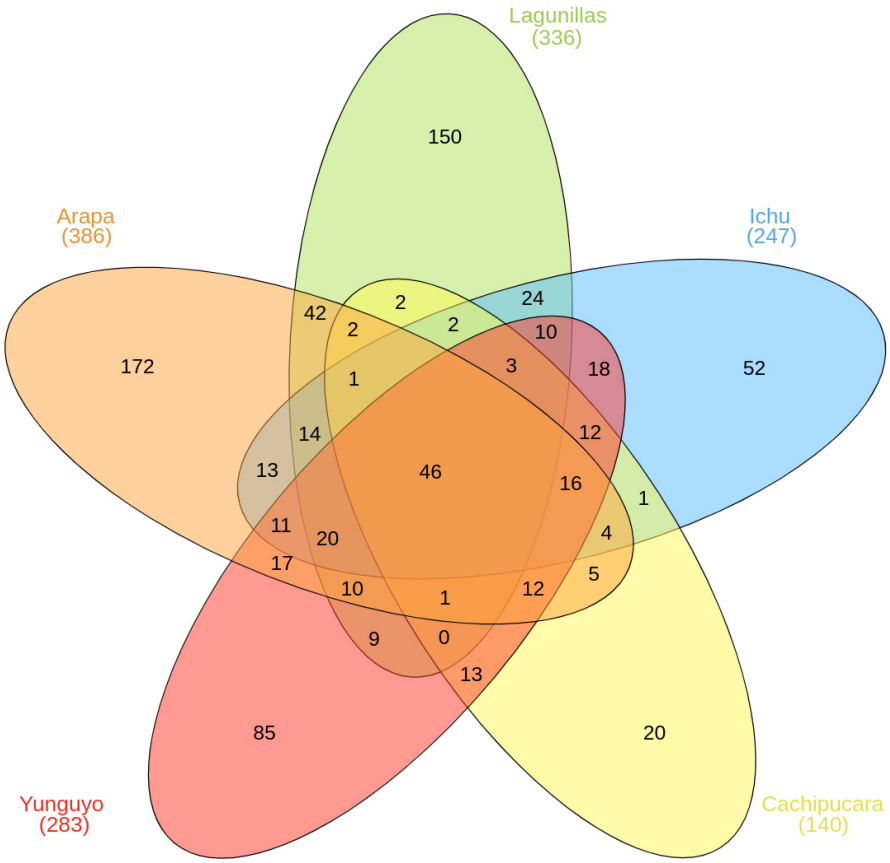


Figura 5. Diagrama de Venn del número de ASVs registrados por localidad, compartidos y únicos para las 5 localidades muestreadas en la cuenca del Lago Titicaca. En paréntesis se indica el número total de ASVs por localidad.

6.2.1. Estructura de las comunidades bacterianas

Luego de la asignación taxonómica de los ASVs, estos se agruparon por localidad para obtener el número de taxones encontrados en cada nivel taxonómico (Tabla 6). Las localidades de Arapa (124 géneros y 21 filos) y Lagunillas (110 géneros y 18 filos) presentaron las mayores cantidades de taxones, mientras que Cachipucara fue la localidad con menor número de taxones

encontrados (64 géneros y 14 filos). Es importante señalar que, aunque se partió de un total de 18 muestras, incluyendo 3 muestras control, en las gráficas 6 y 7 solo se incluyen 13 muestras debido a que los ASVs de las muestras control fueron filtradas de todas las demás muestras. En cuanto a las muestras experimentales correspondientes a Lagunillas (época lluviosa 2019) y Cachipucara (época lluviosa 2019) no alcanzaron el 5% de la abundancia total de secuencias, por lo que no se incluyeron en los gráficos.

Tabla 6. Número de niveles taxonómicos encontrados por localidad en el Lago Titicaca y su cuenca hidrográfica.

Zona	Localidad	Niveles taxonómicos					
		Dominio	Filo	Clase	Orden	Familia	Género
Cuenca del Lago Titicaca	Arapa	1	21	32	64	96	124
	Lagunillas	1	18	26	50	85	110
Lago Titicaca	Ichu	1	16	22	46	67	86
	Cachipucara	1	14	28	36	55	64
	Yunguyo	1	17	24	51	82	102

Se identificaron 8 órdenes más representativos en todas las muestras (Figura 6), definidos como aquellos con abundancias relativas superiores al 5%. Este umbral se estableció con el fin de destacar los taxones dominantes dentro de la comunidad bacteriana y minimizar la influencia de grupos de baja frecuencia o presencia esporádica. Los órdenes predominantes fueron: Synechococcales, Holophagales, Chitinophagales, Frankiales, Microtrichales, Flavobacteriales,

SAR11 clade y Chthoniobacterales. A nivel de familia, se identificaron nueve grupos representativos en todas las muestras (Figura 7): Cyanobiaceae, Flavobacteriaceae, Holophagaceae, Chitinophagaceae, Sporichthyaceae, Ilumatobacteraceae, Crocinitomicaceae, Clade III y Terrimicrobiaceae. Por otro lado, se obtuvieron taxones únicos por localidad, órdenes y familias (Figura 8 y 9), siendo la localidad Arapa la más diferente con 14 órdenes únicos y 20 familias únicas encontradas. Las Tablas 7 y 8 presentan las familias, órdenes y géneros únicos encontrados en cada una de las localidades de cultivo de *O. mykiss* en la cuenca hidrográfica del lago. En Arapa, se identificaron varias familias exclusivas, incluyendo Holosporaceae, Margulisbacteria, Hydrogenophilaceae y Nitrospiraceae. En Lagunillas, las familias exclusivas fueron Leptotrichiaceae, Piscirickettsiaceae, WCHB1-41, Lentimicrobiaceae, y Devosiaceae. Ichu presentó las familias Absconditabacteraceae, Coxiellaceae, y Oligoflexaceae. En Cachipucara, se encontraron familias como Cytophagaceae, FukuN18, y Psychrobacteraceae. Finalmente, en Yunguyo, se identificaron Halieaceae, Rhizobiaceae, Vibrionaceae, Defluviitaleaceae, y Clostridiaceae.

En Arapa, los órdenes exclusivos incluyeron Berkelbacteria, Candidatus Komeilibacteria, Elsterales, Holosporales, y JTB23. En Lagunillas, se encontraron los órdenes Piscirickettsiales, Chlamydiales, BD7-11, y Leptotrichiales. Ichu presentó Absconditabacteriales (SR1), Coxiellales, y Oligoflexales. En Cachipucara, el orden Cytophagales fue exclusivo, mientras que en Yunguyo, se detectaron Clostridiales, Lachnospirales, Omnitrophales, y Vibrionales.

A nivel de la presencia de géneros bacterianos, en Arapa se destacaron Babeliales, Bauldia, Candidatus Nitrotoga, Nitrospira, Sphingobium, y Thauera. Lagunillas presentó los géneros como Leptotrichiaceae, Piscirickettsiaceae, AKYH767, Lentimicrobium, Candidatus Endoecteinascidia,

y Crocinitomix. En Ichu, los géneros exclusivos fueron Coxiella, Oligoflexus, Absconditabacteriales (SR1), y Oligoflexus. Cachipucara mostró géneros como Psychrobacter, FukuN18, IS-44, y FukuN18 freshwater group. Finalmente, en Yunguyo, se identificaron géneros como Clostridium, Blfdi19, Omnitrophus y Rhizobiaceae.

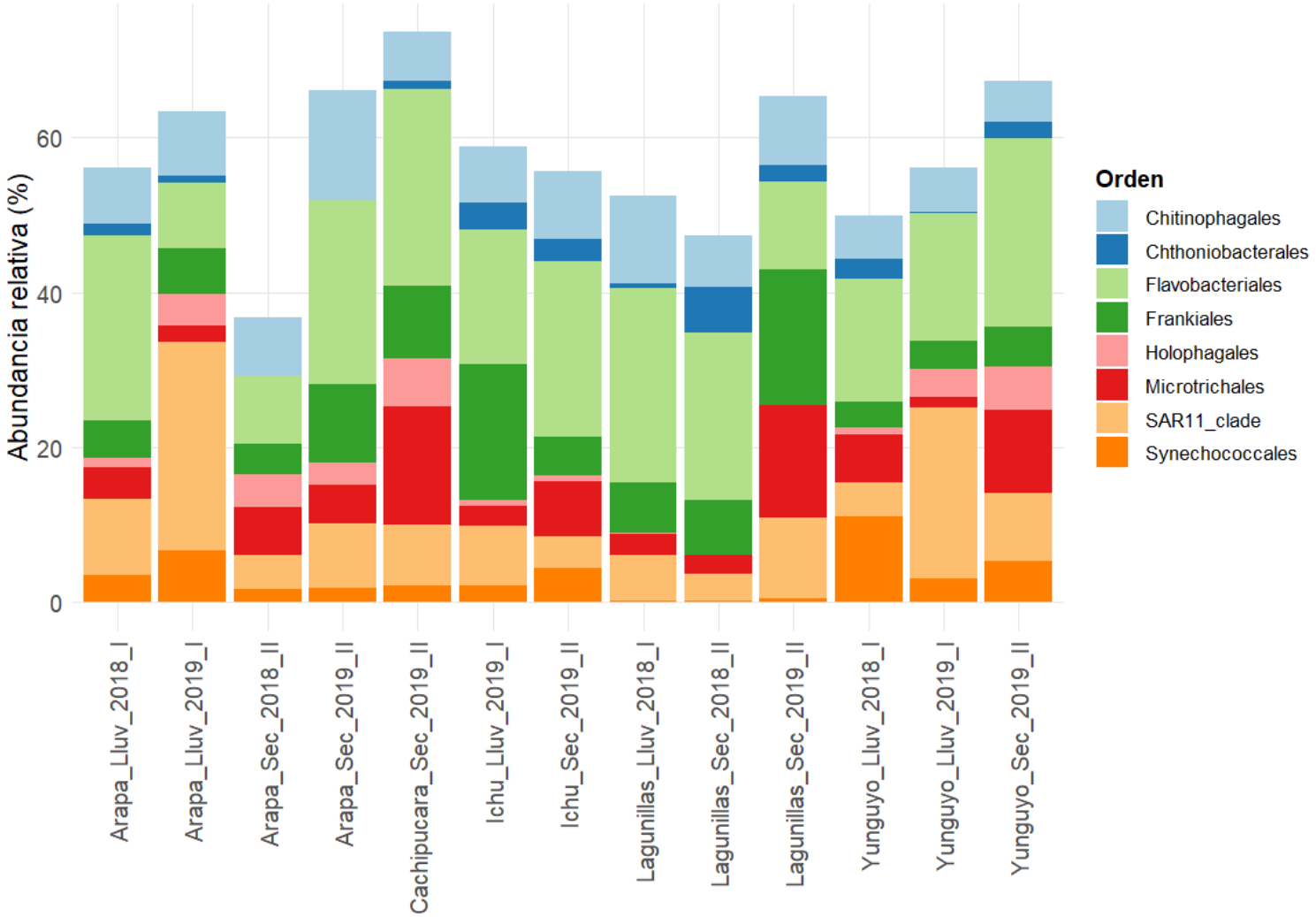


Figura 6. Gráfico de barras de órdenes con un mínimo de 5% de abundancia.

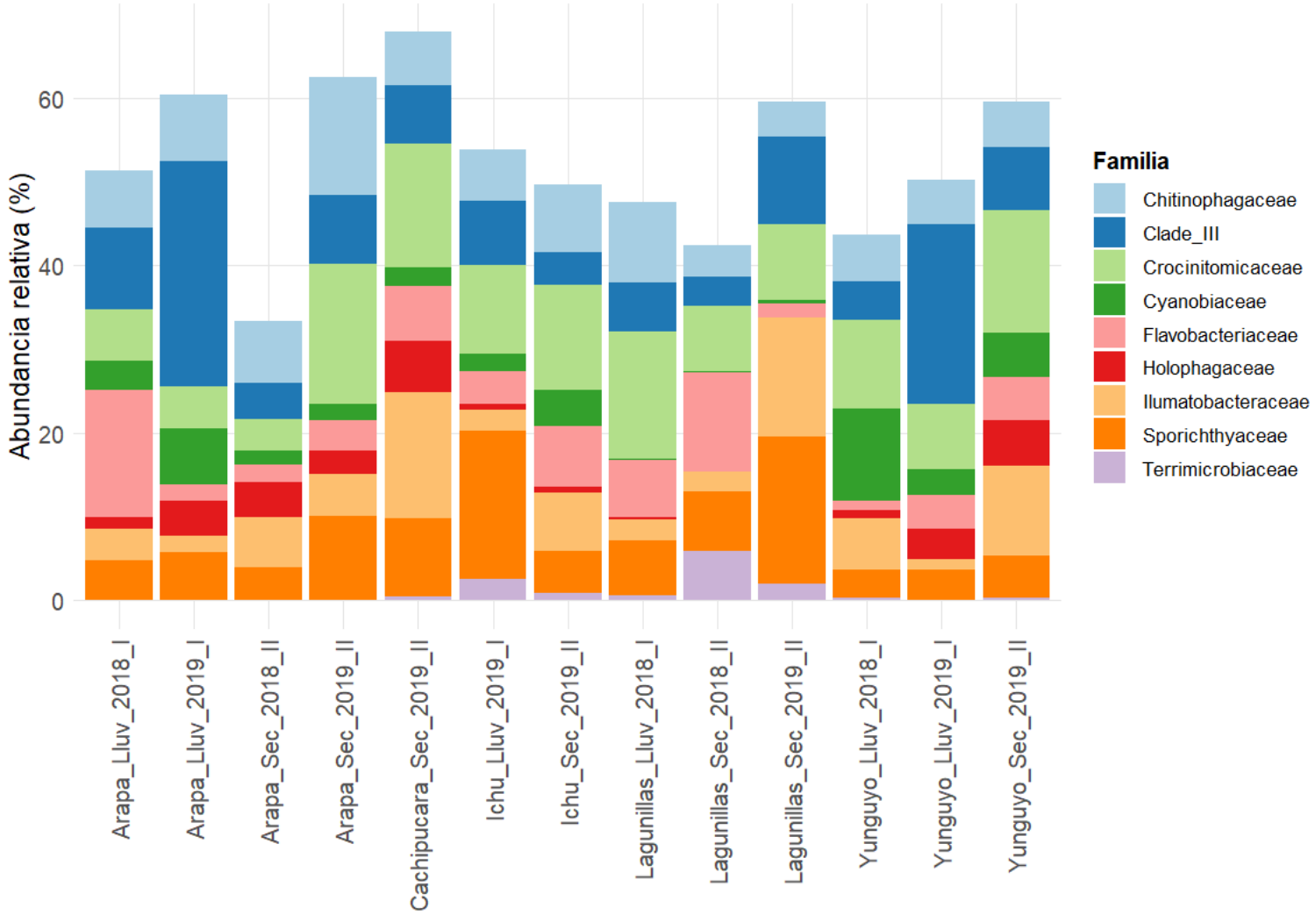


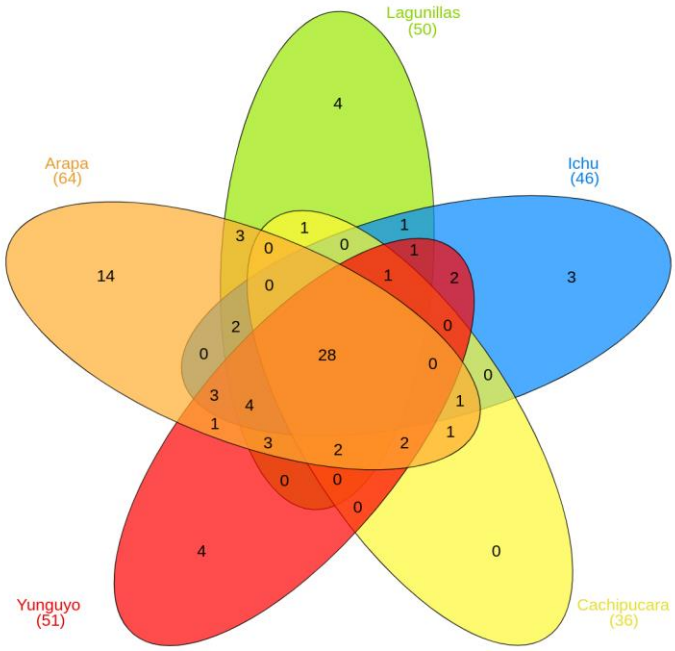
Figura 7. Gráfico de barras de familias con un mínimo de 5% de abundancia.

Tabla 7. Órdenes y familias únicos por muestras de agua recolectadas en cinco localidades de la cuenca del Lago Titicaca, entre 2018 y 2019.

Localidad	Orden	Familia
Arapa	AT-s3-44 Berkelbacteria Candidatus Komeilibacteria Candidatus Nomurabacteria Candidatus Uhrbacteria Elsterales Holosporales JTB23 JG30-KF-CM66 Margulisbacteria Nitrospirales Polyangiales vadinHA49 WWE3	AT-s3-44 Babeliales Berkelbacteria Candidatus Komeilibacteria Candidatus Nomurabacteria Candidatus Uhrbacteria Gallionellaceae Hafniaceae Holosporaceae Hydrogenophilaceae Hyphomonadaceae JTB23 JG30-KF-CM66 Margulisbacteria Methylococcaceae Nitrospiraceae Sandaracinaceae SM2D12 vadinHA49 WWE3
Lagunillas	BD7-11 Chlamydiales Piscirickettsiales WCHB1-41	AKYH767 BD7-11 Devosiaceae Lentimicrobiaceae Leptotrichiaceae Piscirickettsiaceae WCHB1-41
Ichu	Absconditabacteriales (SR1) Coxiellales Oligoflexales	Absconditabacteriales (SR1) Coxiellaceae Oligoflexaceae
Cachipucara	–	Cytophagaceae
Yunguyo	Blfdi19 Clostridiales Lachnospirales Omnitrophales	Blfdi19 Clostridiaceae Defluviitaleaceae Halieaceae Microscillaceae Nostocaceae Omnitrophaceae Rhizobiaceae Rhodanobacteraceae Sutterellaceae Vibrionaceae

Tabla 8. Géneros únicos por localidad registrados en las 5 localidades evaluadas en la cuenca del lago Titicaca, entre 2018 y 2019.

Localidad	Género	
Arapa	AT-s3-44	JTB23
	Babeliales	Margulisbacteria
	Bauldia	MND1
	Berkelbacteria	Nitrospira
	Candidatus Komeilibacteria	Nitrosomonas
	Candidatus Nitrotoga	Nordella
	Candidatus Nomurabacteria	Roseimaritima
	Candidatus Uhrbacteria	Sandaracinus
	Emticicia	SM2D12
	Hafnia-Obesumbacterium	Sphingobium
	Hirschia	Thiobacillus
	Holospora	Tibeticola
	IMCC26134	vadinHA49
	JG30-KF-CM66	WWE3
Lagunillas	AKYH767	Lentimicrobium
	Alsobacter	Methylobacter
	BD7-11	Novosphingobium
	Candidatus Anammoximicrobium	Rhizobacter
	Candidatus Endoecteinascidia	Sphingoaurantiacus
	Candidatus Planktoluna	Sphingomonas
	Crocinitomix	Verticiella
	Hypnocyclicus	WCHB1-41
Ichu	Absconditabacteriales_(SR1)	Oligoflexus
	Coxiella	Roseococcus
	ML602J-51	
Cachipucara	Cytophaga	IS-44
	FukuN18 freshwater group	Psychrobacter
Yunguyo	AAP99	Defluviitaleaceae_UCG-011
	Actibacter	Enhydrobacter
	Aphanizomenon NIES81	Marinagarivorans
	Aquimonas	Nodularia_PCC-9350
	Arcicella	OLB12
	Blfdi19	OM60(NOR5)_clade
	Candidatus Omnitrophus	Photobacterium
	Clostridium sensu stricto 8	
	Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium	


Lagunillas
(50)
Arapa
(64)
Ichu
(46)
Yunguyo
(51)
Cachipucara
(36)
4
14
3
28
4
0
0
1
1
0
1
1
2
0
0
2
3
0
0
3
4
1
1
3
2
2
1
0
0
0
0
0
1
1
0
0
0
0
0

6.3. Análisis de la diversidad

6.3.1. Alfa-rarefacción

Para determinar el punto de rarefacción, se evaluó el número de lecturas por muestra (Tabla 5). Con base en esta información, se estableció un umbral de 5,000 lecturas (Figura 10), considerando que la muestra Arapa (lluviosa, 2019), con 5,979 lecturas, presentó el menor número de lecturas entre las muestras viables incluidas en el análisis. Se excluyeron del proceso las 3 muestras control (2018-I, 2018-II y 2019-I-II), así como las muestras Cachipucara (lluviosa, 2019) (1,455 lecturas) y Lagunillas (lluviosa, 2019) (84 lecturas), debido al bajo número de lecturas que las hace inadecuadas para la normalización por rarefacción.

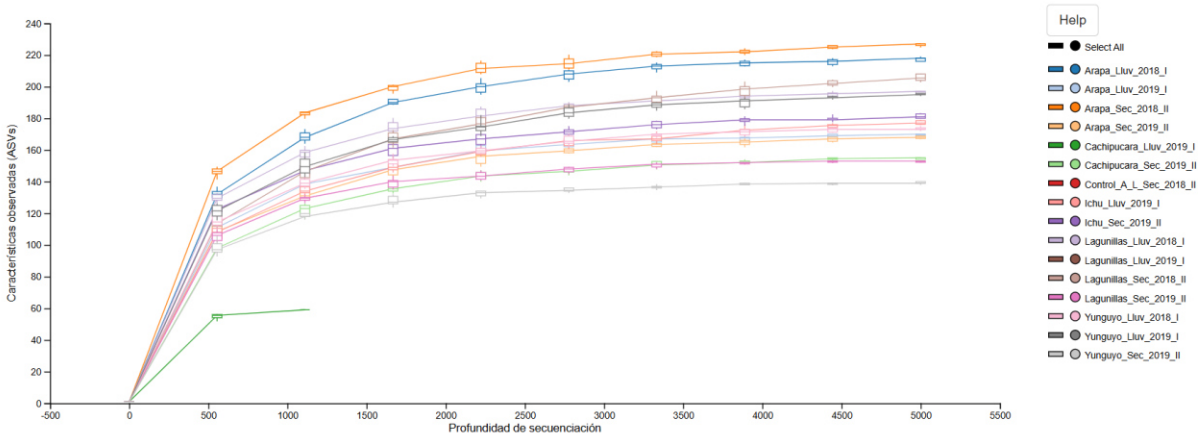


Figura 10. Rarefacción de las muestras con umbral de rarefacción 5000 lecturas.

6.3.2. Diversidad alfa

La diversidad alfa se evaluó mediante los índices de Shannon, Simpson y Faith_pd (*Faith's phylogenetic diversity*), cuyos valores se calcularon para cada muestra (Tabla 9). En general, los

resultados mostraron una alta diversidad bacteriana en la mayoría de las localidades, con valores de Shannon superiores a 5.5 y de Simpson mayores a 0.96.

La mayor diversidad se registró en la localidad de Arapa, con valores de Shannon entre 5.5 y 6.7, Simpson entre 0.93 y 0.98, y Faith_pd entre 14.4 y 19.3, además de alcanzar el número más alto de ASVs (216). Por el contrario, Cachipucara presentó la menor diversidad, con 51 ASVs durante la época lluviosa de 2019 y los valores más bajos en el índice de Shannon (5.75) y Faith_pd (14.6). De acuerdo con el gráfico de cajas y bigotes del índice de Shannon (Figura 11), las localidades Arapa y Lagunillas presentaron la mayor variabilidad en la diversidad bacteriana respecto a las localidades del lago Titicaca con un rango más amplio de valores, mientras que Ichu mostró una distribución más uniforme y valores intermedios de diversidad. En contraste, Yunguyo y Cachipucara presentaron los valores más bajos de diversidad y menor dispersión en el índice de Shannon.

Tabla 9. Índices de diversidad alfa por localidad, año y época (lluviosa (abril), seca (agosto)).

Zona	Localidad	Año	Época	Número de ASVs	Índice de diversidad		
					Shannon	Simpson	Faith_pd
Cuenca del Lago Titicaca	Arapa	2018	Lluviosa	206	6.2212	0.9670	19.2447
			Seca	216	6.7094	0.9844	19.3611
		2019	Lluviosa	161	5.5929	0.9347	15.0559
			Seca	163	5.9619	0.9709	14.4871
	Lagunillas	2018	Lluviosa	177	6.3966	0.9803	15.6172
			Seca	195	6.6031	0.9828	17.7346
		2019	Lluviosa	10	*	*	*
			Seca	147	5.9111	0.9700	13.6337
Lago Titicaca	Ichu	2019	Lluviosa	170	6.3169	0.9790	15.1604
			Seca	171	6.3677	0.9816	14.2054
	Cachipucara	2019	Lluviosa	51	*	*	*
			Seca	140	5.7585	0.9690	14.6218
	Yunguyo	2018	Lluviosa	162	6.0079	0.9714	15.4490
			Seca	-	**	**	**
		2019	Lluviosa	172	6.0027	0.9610	15.9682
			Seca	125	5.8270	0.9734	13.4624

*: No pasaron la rarefacción por tener un bajo número de *lecturas* luego del filtrado
**: No presente en los datos de secuenciación (no se obtuvo muestra)

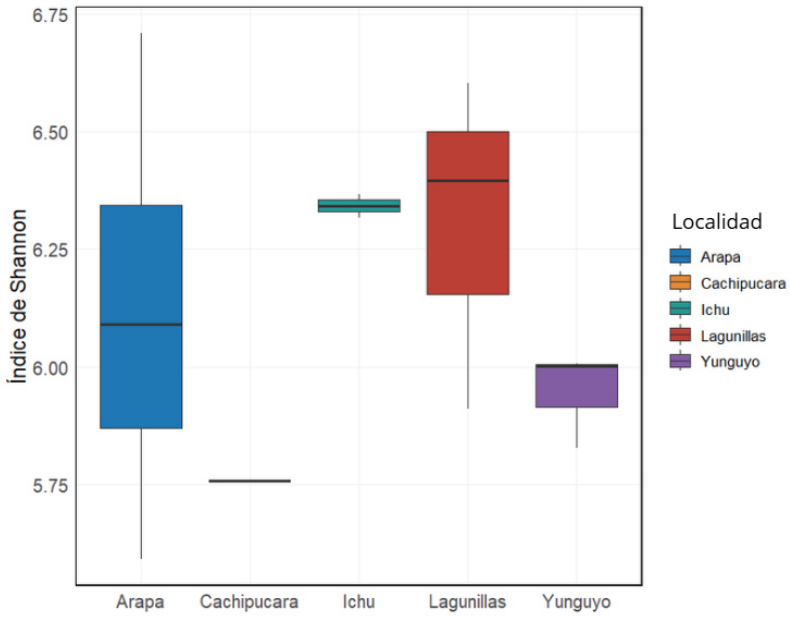


Figura 11. Gráfico de cajas y bigotes del índice de diversidad de Shannon por localidad muestreada en la cuenca del lago Titicaca entre 2018 y 2019.

6.3.3. Diversidad beta

Usando el programa RStudio, se graficó un Análisis de Coordenadas Principales (PCoA) a partir de una matriz de distancias basada en el índice de Jaccard de todas las muestras (Figura 12), agrupándolas por su localidad y época. Las dos primeras componentes principales explicaron el 35.37% de la variación total, donde la PC1 representó el 18.15% y la PC2 el 17.22%. En el gráfico se observa una separación parcial entre las muestras por localidad, donde las localidades de Arapa y Lagunillas se han agrupado por separado de las localidades dentro del Lago Titicaca.

Adicionalmente, se construyó un árbol de agrupamiento jerárquico UPGMA a nivel taxonómico de familia utilizando el programa MEGAN6 (Figura 13). El árbol mostró una tendencia en la organización de las muestras según la época, agrupando las correspondientes a la

época lluviosa principalmente en un mismo clado, mientras que las de la época seca conformaron un grupo independiente. La muestra de Lagunillas (época seca, 2018) fue la única que se agrupó más estrechamente con las muestras de la época lluviosa.

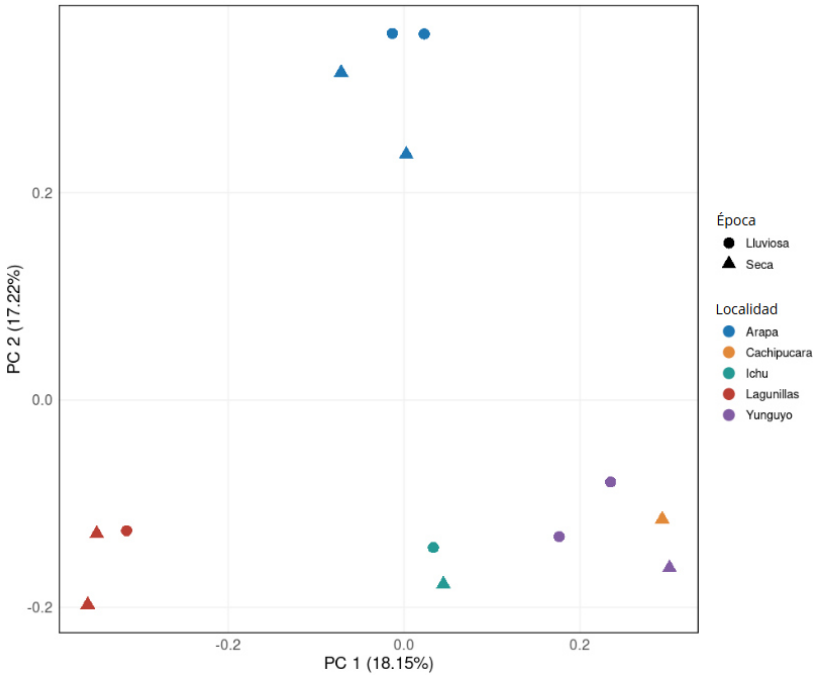


Figura 12. PCoA por localidad en 2018 y 2019 en época lluviosa y seca. *(PC: Coordenada Principal)*

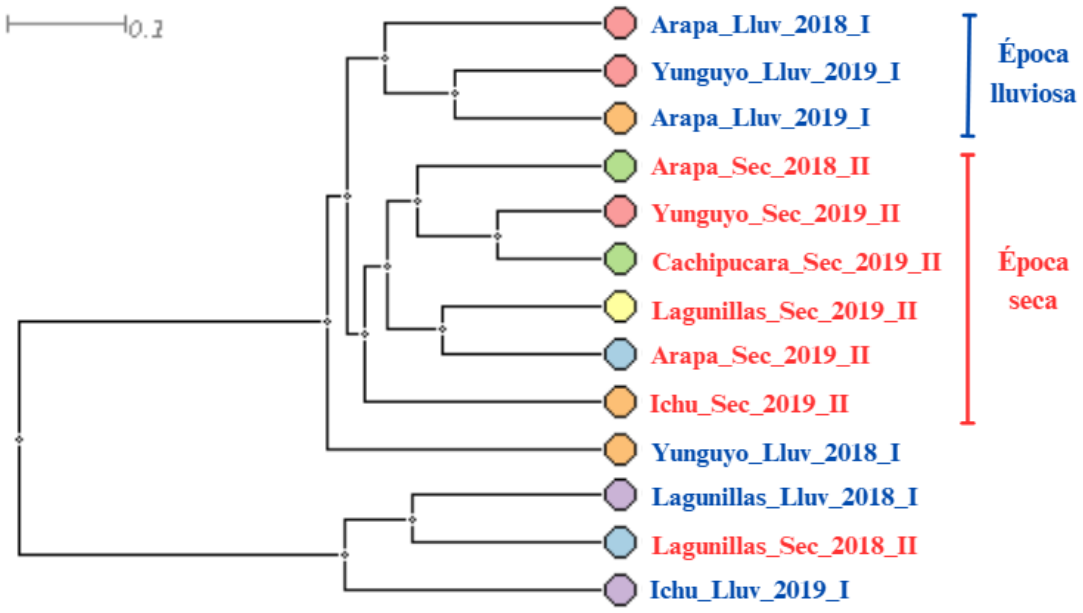


Figura 13. Árbol de agrupamiento por UPGMA del nivel taxonómico familia.

6.4. Predicción funcional

Empleando el programa R, se graficaron los datos de predicción funcional (Tabla 10, 11 y 12) en un mapa de calor (*heatmap*) en base a la abundancia de las funciones predichas en cada muestra (Figura 14). Siendo las funciones más abundantes las de quimioheterotrofia (30.97%), quimioheterotrofia aeróbica (22.25%), fototrofia (8.18%), fotoautotrofia oxigenada (7.38%), fotoautotrofia (7.38%), cianobacterias fotosintéticas (7.38%). Otras funciones que tuvieron abundancias mayores al 1% fueron: degradación de compuestos aromáticos (2.81%), metilotrofía (2.54%), fermentación (2.48%), oxidación del metanol (2.04%), parásitos intracelulares (1.11%) y reducción de nitratos (1.02%).

Tabla 10. Porcentaje de funciones bacterianas predichas presente en las muestras analizadas en cinco localidades de la cuenca del Lago Titicaca.

Función	Frecuencia (%)
Quimioheterotrofia	30.97
Quimioheterotrofia aeróbica	22.25
Fototrofia	8.18
Fotoautotrofia oxigenada	7.38
Fotoautotrofia	7.38
Cianobacterias fotosintéticas	7.38
Degradación de compuestos aromáticos	2.81
Metilotrofía	2.54
Fermentación	2.48
Oxidación del metanol	2.04
Parásitos intracelulares	1.11
Reducción de nitratos	1.02
Fotoheterotrofia	0.81
Degradación de hidrocarburos	0.50
Metanotrofía	0.50
Oxidación oscura del hidrógeno	0.39
Nitrificación	0.34
Parásitos o simbiontes animales	0.33
Asociado a humanos	0.33
Patógenos humanos	0.33
Oxidación aeróbica del amoníaco	0.29
Ureólisis	0.28
Depredador o ectoparásito	0.11
Respiración con nitrato	0.09
Respiración con nitrógeno	0.09
Oxidación aeróbica de nitrito	0.04
Celulólisis	0.01
Oxidación oscura de compuestos de azufre	0.01
Oxidación oscura de sulfuros	0.01

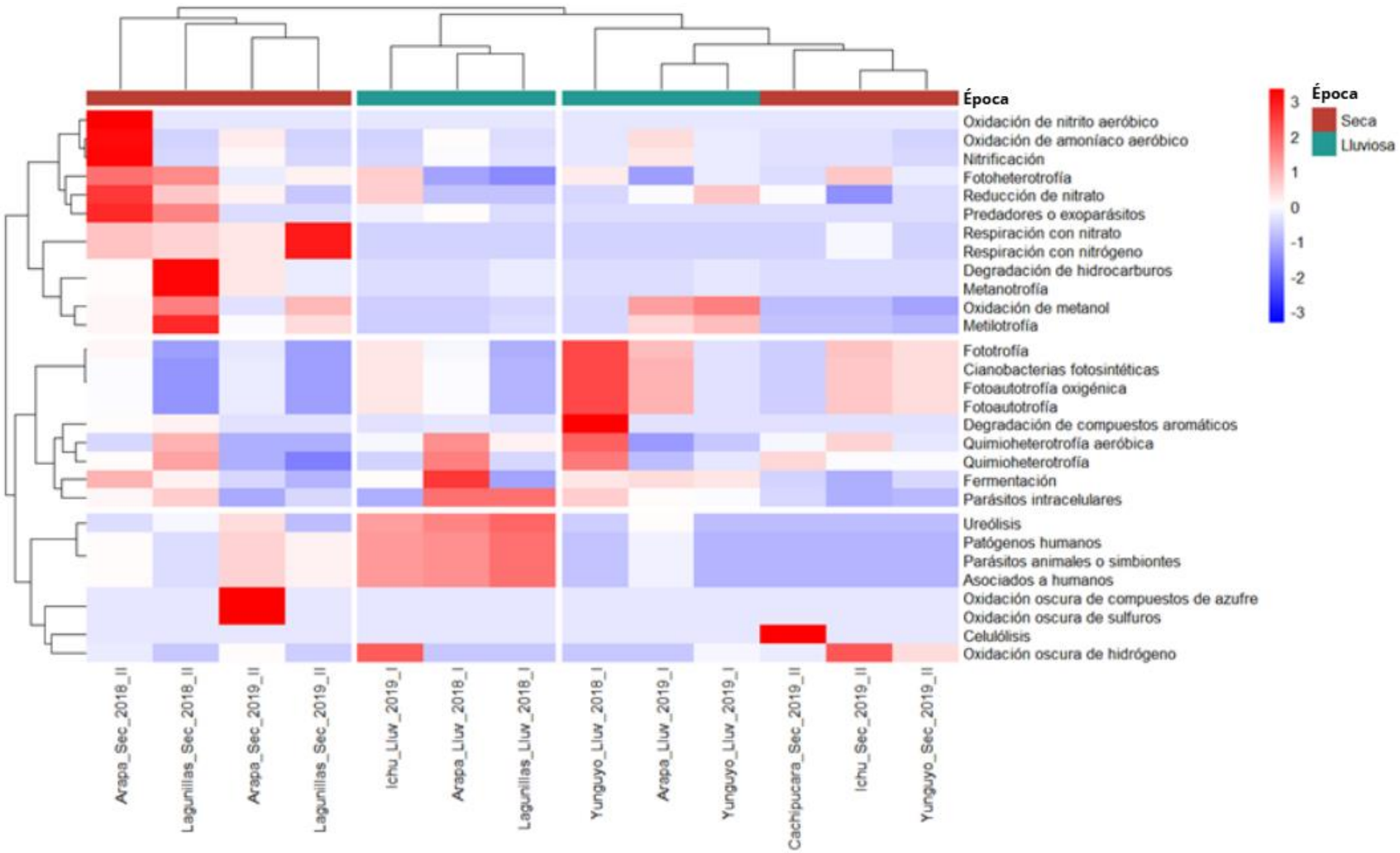


Figura 14. Mapa de calor de la predicción funcional de cada muestra en el 2018 y 2019 por época lluviosa (verde) y seca (marrón).

Tabla 11. Familias y géneros con funciones ecológicas en las localidades de Arapa y Lagunillas (Cuenca del Lago Titicaca).

Función	Familia	Género	Arapa				Lagunillas		
			2018_I	2018_II	2019_I	2019_II	2018_I	2018_II	2019_II
Oxidación aeróbica del amoníaco	Nitrosomonadaceae	966-1	2	2	2	2	2	0	0
	Nitrosomonadaceae	IS-44	0	0	0	0	0	0	0
	Nitrosomonadaceae	MND1	0	1	0	0	0	0	0
	Nitrosomonadaceae	Nitrosomonas	0	1	0	0	0	0	0
Oxidación aeróbica de nitrito	Gallionellaceae	Candidatus Nitrotoga	0	1	0	0	0	0	0
Degradación de compuestos aromáticos	Moraxellaceae	Acinetobacter	15	60	0	0	0	60	15
Patógenos humanos	Acetobacteraceae	Roseomonas	2	2	2	2	2	0	0
	Coxiellaceae	Coxiella	0	0	0	0	0	0	0
	Xanthomonadaceae	Stenotrophomonas	0	8	0	4	0	4	8
Parásitos intracelulares	AB1	AB1	0	0	0	0	0	2	0
	Candidatus Hepatincola	Candidatus Hepatincola	3	3	0	3	9	3	3
	Coxiellaceae	Coxiella	0	0	0	0	0	0	0
	Fokiniaceae	MD3-55	2	0	0	0	2	0	0
	Holosporaceae	Holospora	1	0	0	0	0	0	0
	Holosporaceae	uncultured	0	0	0	0	0	0	2
	Legionellaceae	Legionella	0	7	7	7	7	7	7
	Rickettsiaceae	Candidatus Megaira	3	3	3	0	0	3	0
	Rickettsiaceae	Rickettsia	3	0	3	0	6	0	0
	Rickettsiaceae	uncultured	0	0	0	0	0	0	0
	Rickettsiaceae	NA	0	0	0	0	0	1	0
	SM2D12	SM2D12	1	0	1	0	0	0	0
	Simkaniaceae	uncultured	0	1	0	0	0	0	0
Reducción de nitratos	Aeromonadaceae	Aeromonas	2	0	0	2	2	2	0

	Opitutaceae	Opitutus	5	15	5	5	10	5	0
	Shewanellaceae	Shewanella	0	0	0	0	0	0	1
	Spirosomaceae	Emticicia	0	0	0	1	0	0	0
	Xanthomonadaceae	Stenotrophomonas	0	8	0	4	0	4	8
Respiración con nitratos	Shewanellaceae	Shewanella	0	0	0	0	0	0	1
	Xanthomonadaceae	Stenotrophomonas	0	8	0	4	0	4	8
Nitrificación	Gallionellaceae	Candidatus Nitrotoga	0	1	0	0	0	0	0
	Nitrosomonadaceae	966-1	2	2	2	2	2	0	0
	Nitrosomonadaceae	IS-44	0	0	0	0	0	0	0
	Nitrosomonadaceae	MND1	0	1	0	0	0	0	0
	Nitrosomonadaceae	Nitrosomonas	0	1	0	0	0	0	0
Cianobacterias fotosintéticas	Cyanobiaceae	Cyanobium PCC-6307	36	27	36	36	9	9	18
	Leptolyngbyaceae	Chamaesiphon PCC-7430	1	1	1	1	0	0	0
	Leptolyngbyaceae	NA	0	0	0	0	0	0	0
	Microcystaceae	Microcystis PCC-7914	1	0	1	0	1	0	0
	Microcystaceae	Snowella 0TU37S04	1	0	0	0	0	0	0
	Nostocaceae	Aphanizomenon NIES81	0	0	0	0	0	0	0
	Nostocaceae	Nodularia PCC-9350	0	0	0	0	0	0	0
	Phormidiaceae	Tychonema CCAP 1459-11B	0	2	2	2	0	0	0
Depredador o ectoparásito	Bdellovibrionaceae	OM27 clade	12	12	0	0	0	12	0

2018-I y 2019-I: corresponden a la época lluviosa
2018-II y 2019-II: corresponden a la época seca

Tabla 12. Familias y géneros con funciones ecológicas en las localidades de Yunguyo, Ichu y Cachipucara.

Función	Familia	Género	Yunguyo			Ichu		Cachipucara
			2018_I	2019_I	2019_II	2019_I	2019_II	2019_II
Oxidación aeróbica del amoníaco	Nitrosomonadaceae	966-1	2	2	0	0	2	0
	Nitrosomonadaceae	IS-44	0	0	0	0	0	1
	Nitrosomonadaceae	MND1	0	0	0	0	0	0
	Nitrosomonadaceae	Nitrosomonas	0	0	0	0	0	0
Oxidación aeróbica de nitrito	Gallionellaceae	Candidatus Nitrotoga	0	0	0	0	0	0
Degradación de compuestos aromáticos	Moraxellaceae	Acinetobacter	180	15	0	0	0	0
Patógenos humanos	Acetobacteraceae	Roseomonas	2	0	0	2	0	0
	Coxiellaceae	Coxiella	0	0	0	1	0	0
	Xanthomonadaceae	Stenotrophomonas	0	0	0	0	0	0
Parásitos intracelulares	AB1	AB1	0	0	0	2	2	0
	Candidatus Hepatincola	Candidatus Hepatincola	0	3	3	0	3	3
	Coxiellaceae	Coxiella	0	0	0	1	0	0
	Fokiniaceae	MD3-55	0	0	2	0	0	2
	Holosporaceae	Holospora	0	0	0	0	0	0
	Holosporaceae	uncultured	2	2	0	2	0	0
	Legionellaceae	Legionella	7	0	0	0	0	0
	Rickettsiaceae	Candidatus Megaira	3	3	0	0	0	3
	Rickettsiaceae	Rickettsia	0	0	0	0	0	0
	Rickettsiaceae	uncultured	0	4	0	2	0	0
	Rickettsiaceae	NA	0	0	0	0	0	0
	SM2D12	SM2D12	0	0	0	0	0	0
	Simkaniaceae	uncultured	0	0	0	0	0	0
Reducción de nitratos	Aeromonadaceae	Aeromonas	2	4	2	0	0	0

	Opitutaceae	Opitutus	10	10	5	10	0	5
	Shewanellaceae	Shewanella	0	0	0	0	1	0
	Spirosomaceae	Emticicia	0	0	0	0	0	0
	Xanthomonadaceae	Stenotrophomonas	0	0	0	0	0	0
Respiración con nitratos	Shewanellaceae	Shewanella	0	0	0	0	1	0
	Xanthomonadaceae	Stenotrophomonas	0	0	0	0	0	0
Nitrificación	Gallionellaceae	Candidatus Nitrotoga	0	0	0	0	0	0
	Nitrosomonadaceae	966-1	2	2	0	0	2	0
	Nitrosomonadaceae	IS-44	0	0	0	0	0	1
	Nitrosomonadaceae	MND1	0	0	0	0	0	0
	Nitrosomonadaceae	Nitrosomonas	0	0	0	0	0	0
Cianobacterias fotosintéticas	Cyanobiaceae	Cyanobium PCC-6307	27	27	45	27	54	36
	Leptolyngbyaceae	Chamaesiphon PCC-7430	0	0	0	0	1	0
	Leptolyngbyaceae	NA	0	0	0	0	0	1
	Microcystaceae	Microcystis PCC-7914	0	0	0	0	0	0
	Microcystaceae	Snowella 0TU37S04	0	1	1	1	1	1
	Nostocaceae	Aphanizomenon NIES81	1	0	0	0	0	0
	Nostocaceae	Nodularia PCC-9350	1	0	0	0	0	0
	Phormidiaceae	Tychonema CCAP 1459-11B	0	2	0	0	0	0
Depredador o ectoparásito	Bdellovibrionaceae	OM27 clade	0	0	0	6	0	0

2018-I y 2019-I: corresponden a la época lluviosa
2018-II y 2019-II: corresponden a la época seca

DISCUSIÓN

La diversidad bacteriana en las áreas de cultivo de *O. mykiss* en la cuenca hidrográfica del Lago Titicaca fue caracterizada a través del análisis de las comunidades bacterianas en las cinco localidades (Cachipucara, Yunguyo, Ichu, Laguna Lagunillas y Lago Arapa). Se identificaron un total de 875 ASVs y 186 géneros a lo largo de todas las localidades, basado en el análisis de la región hipervariable V3-V4 del gen ARNr 16S.

Los filos bacterianos más abundantes (>5%), identificados en este estudio como Bacteroidota, Acidobacteriota, Cyanobacteria, Verrucomicrobiota, Actinobacteriota y Proteobacteria, son ampliamente reconocidos como componentes dominantes en comunidades microbianas de ambientes de agua dulce (Logue, Bürgmann, & Robinson, 2008). Esta composición filogenética refleja patrones típicos observados en sistemas lacustres y sugiere la existencia de funciones ecológicas compartidas, como la degradación de materia orgánica, la fijación de nitrógeno y la fotosíntesis. Estudios en ecosistemas fríos y estacionalmente congelados han evidenciado que la estructura de las comunidades bacterianas varía notablemente entre hábitats (hielo, agua y sedimento), destacando la dominancia de filos como Proteobacteria, Bacteroidota, Actinobacteria y Cyanobacteria en la columna de agua, mientras que Acidobacteriota y Chloroflexi predominan en los sedimentos, posiblemente debido a su capacidad para tolerar condiciones de baja oxigenación y alta carga orgánica (Zhang et al., 2024). Estas diferencias espaciales y funcionales subrayan la importancia de considerar la heterogeneidad ambiental en la evaluación de la diversidad bacteriana en ecosistemas acuáticos de alta montaña como el Lago Titicaca y sus afluentes.

La utilización de la región hipervariable V3-V4 del gen ARNr 16S, combinada con la base de datos SILVA, resulta ser una estrategia efectiva para la caracterización de la diversidad bacteriana en los sistemas acuáticos del Lago Titicaca. Esta región proporciona una resolución taxonómica suficiente para distinguir comunidades microbianas complejas, especialmente en ecosistemas de agua dulce, sin comprometer la cobertura de diversidad (Bukin et al., 2019). Por otro lado, la base de datos SILVA proporcionó una anotación taxonómica robusta al incluir secuencias de clados ambientales no cultivados (Glöckner et al., 2017), característica crítica para ambientes como el Lago Titicaca que son poco estudiados.

Las diferencias observadas en la diversidad alfa parecen estar relacionadas no solo con el impacto de la actividad acuícola, sino también con la distribución espacial de las localidades y la variación temporal entre las estaciones lluviosa y seca. Estas variaciones reflejan la heterogeneidad ambiental de la cuenca del Lago Titicaca y la influencia de los distintos factores ecológicos y productivos presentes en cada zona de muestreo, los cuales pueden afectar la composición y estructura de las comunidades bacterianas.

Las diferencias notables entre localidades y estaciones en diversidad alfa, reflejaron variaciones en la riqueza (número de ASVs), equitatividad (índices de Shannon y Simpson) y diversidad filogenética (índice Faith_pd) de las comunidades microbianas. Estos patrones sugieren una respuesta directa de las comunidades bacterianas a los factores ambientales locales y al grado de presión antrópica en cada punto de muestreo.

En la laguna Lagunillas, parte de la cuenca del lago Titicaca, los valores elevados y consistentes de los índices de Shannon (>6.3) y Simpson (>0.97) indicaron la presencia de

comunidades bacterianas diversas, complejas y equitativas. Además, esta localidad presentó uno de altos valores de ASVs únicos (150), lo que refuerza su papel como uno de los sitios con mayor riqueza microbiana dentro del área de estudio. Desde el año 2007, la intensificación de la actividad acuícola en la laguna ha generado una mayor densidad de productores y una carga significativa de nutrientes, asociada al incremento de fosfatos (>20 mg/L), nitrógeno amoniacal (>600 mg/L) y coliformes, que evidencian un proceso de enriquecimiento del agua y la influencia directa de la acuicultura intensiva sobre la calidad del ecosistema (Cornejo Sánchez, 2022). Sin embargo, algunos sectores de la laguna mantienen condiciones fisicoquímicas dentro de los estándares nacionales para la vida acuática, con niveles moderados de oxígeno disuelto y actividad fotosintética estable (Chura-Cruz et al., 2014). Esta coexistencia de una alta carga de nutrientes con condiciones ambientales aún favorables podría explicar los elevados valores de diversidad alfa observados en Lagunillas, donde la presencia simultánea de materia orgánica y parámetros fisicoquímicos adecuados favorece la persistencia de comunidades bacterianas diversas y funcionalmente activas.

De igual forma, el Lago Arapa forma parte de la cuenca del Lago Titicaca, y durante la época de lluvias, se conecta con el Lago Titicaca a través del río Ramis, que actúa como un afluente importante (Ocola & Laqui, 2017). Esto podría vincularse con que si bien se registra una diversidad alfa elevada, también presentó una variabilidad entre estaciones y años evaluados. Además, se resalta que esta localidad presentó el mayor número de ASVs únicos, en comparación con las otras zonas evaluadas en este estudio. Este patrón puede explicarse por la influencia de descargas domésticas provenientes de la planta de tratamiento de aguas residuales del distrito de Chupa, cuya calidad de efluentes ha sido reportada como deficiente, con concentraciones extremadamente altas

de DQO (690 mg/L) y coliformes termotolerantes (1,500,000 NMP/100 mL), generando una presión ambiental intermitente que puede favorecer comunidades dinámicas pero inestables (Larico-Robles, 2024).

Otro aspecto interesante de resaltar es que, las zonas dentro del lago Titicaca evaluadas en este estudio, mostraron valores menores de ASVs únicos. En particular, en la localidad de Cachipucara, se obtuvieron los valores más bajos de diversidad alfa (Shannon ≈ 5.75 y Simpson ≈ 0.96) del estudio. Si bien no se tienen reportes previos respecto a la comunidad bacteriana ni el estado ecológico del lugar, el Ministerio del Ambiente (MINAM, 2014) indicó que el distrito de Pilcuyo, cercano a Cachipucara, generó más de 7 toneladas diarias de residuos sólidos municipales. Esta carga, si no es gestionada adecuadamente, podría estar impactando en la calidad del agua y promoviendo la dominancia de unos pocos taxones adaptados. Por otro lado, en Ichu los índices de diversidad (Shannon ≈ 6.3 y Simpson ≈ 0.98) registrados sugeriría la presencia de una comunidad microbiana compleja y balanceada, a pesar de la proximidad a zonas urbanas como la ciudad de Puno. Siguayro y colaboradores (2022) reportaron que la bahía interior de Puno presenta condiciones hipertróficas con exceso de nutrientes, coliformes termotolerantes y clorofila-a. Sin embargo, la alta diversidad observada en Ichu podría vincularse a una heterogeneidad ambiental o a la influencia combinada de fuentes de nutrientes y zonas de mayor circulación de agua, lo que estaría favoreciendo la coexistencia de múltiples grupos bacterianos. Finalmente, Yunguyo, localizada en el extremo sur del lago, mostró alta riqueza (hasta 172 ASVs) pero valores moderados de equitatividad (Shannon ≈ 6.00). En esta localidad se han documentado condiciones de hipertrofia relacionadas con descargas urbanas y actividad acuícola, incluyendo niveles elevados de fósforo, nitrógeno y coliformes (Siguayro et al., 2022). Esto puede explicar la

dominancia de ciertos taxones en la comunidad, lo cual reduce la distribución equitativa de abundancias y, por tanto, los valores de diversidad alfa.

En el Lago Arapa, se identificaron varios órdenes bacterianos únicos con implicaciones ecológicas y evolutivas relevantes, que podrían estar relacionados con la actividad acuícola intensiva en el lago. El orden Elsterales se asocia con ambientes copiotróficos, donde la alta disponibilidad de nutrientes favorece su proliferación, particularmente en zonas eutrofizadas (Shen et al., 2022). Asimismo, el orden Nitrospirales, clave en el proceso de nitrificación, sugiere una relación con la acumulación de compuestos nitrogenados derivados de los desechos acuícolas, lo que puede alterar la dinámica de las comunidades microbianas en el fondo del lago. Este hallazgo, junto con la presencia de Clostridiales y Myxococcales, está vinculado a la posibilidad de estados alternativos en la microbiota, lo que indica posibles cambios en la funcionalidad del ecosistema debido al estrés ambiental y la contaminación por nutrientes (Shang et al., 2021). Por otro lado, Candidatus Komeilibacteria, un taxón asociado a procesos de reducción de sulfato, refleja condiciones anóxicas en los sedimentos del lago, relacionadas con la degradación anaeróbica de materia orgánica (Li et al., 2024), lo cual podría estar relacionado a la acumulación de residuos de la acuicultura. Por último, Margulisbacteria, un grupo bacteriano no cultivable estrechamente relacionado con los ancestros de las cianobacterias, desempeña un papel crucial en la descomposición de materia orgánica y en el ciclo de elementos esenciales en ambientes anóxicos y estuarinos (Wang et al., 2024), sugiriendo su importancia ecológica en la transformación de compuestos en el lago.

En la Laguna Lagunillas, se identificaron órdenes bacterianos únicos que reflejan las condiciones ecológicas particulares de esta zona, destacando la presencia de taxones posiblemente

asociados a patogenicidad. El orden BD7-11 encontrado en la zona, ha sido relacionado con la degradación de polímeros naturales y microplásticos en ambientes acuáticos. En un estudio realizado en lagos boreales, se observó que ciertos grupos, como *Singulisphaera sp.*, *I-8* (Phycisphaeraceae), y un *BD7-11 no cultivable*, se correlacionaron con valores elevados de actividad indicativos de su rol en la degradación de compuestos como polietileno (PE) y polipropileno (PP) (Vesamäki et al., 2022). Otro orden exclusivo es Piscirickettsiales, que incluye patógenos conocidos, como *Piscirickettsia salmonis*, causante de enfermedades en salmónidos. Su presencia podría estar asociada a la influencia de actividades acuícolas en la propagación de microorganismos potencialmente patógenos en el ecosistema lacustre, y el hallazgo de secuencias de una posible segunda especie de vida libre podría ofrecer más información sobre la evolución de la patogenicidad en este género (Wan et al., 2016). Además, se detectó el orden Chlamydiales, que incluye organismos como *Criblamydia sequanensis*, identificado como patógeno emergente en humanos y animales. La aparición de este taxón en la laguna Lagunillas podría estar vinculada a la proliferación de microorganismos patógenos en ambientes acuáticos influenciados por factores antropogénicos (Thomas et al., 2006). Asimismo, el orden WCHB1-41 identificado en esta localidad, juega un papel clave en la degradación de materia orgánica, especialmente en ambientes anaeróbicos. Se ha demostrado que las bacterias de este género, al descomponer ácidos grasos de cadena larga, contribuyen a la metanación del sustrato y ayudan en la purificación del agua, manteniendo la salud del ecosistema acuático. Su alta abundancia en ambientes con mayor proporción de descarga de aguas residuales sugiere que el cambio a condiciones anaeróbicas podría favorecer la proliferación de este grupo (Zhou et al., 2024).

En la localidad de Ichu, se identificaron órdenes bacterianos exclusivos que presentan características interesantes tanto desde el punto de vista ecológico como funcional, vinculado además con influencia de ganado. El orden Absconditabacteriales (SR1) ha sido relacionado con microorganismos en el microbioma ruminal de vacas lecheras con salud deteriorada, particularmente durante la temporada de lluvias. Este orden ha mostrado una mayor abundancia en condiciones de deterioro de la salud de los animales, lo que sugiere un potencial como posible biomarcador en problemas de salud, como la mastitis en vacas lecheras. A pesar de que no se ha cultivado en laboratorio, su presencia ha sido detectada en otros hábitats como termitas y en el tracto digestivo de mamíferos, con una menor abundancia en individuos sanos y un incremento en aquellos con enfermedades orales (Maynez-Perez et al., 2024; Zhong et al., 2020). El orden Coxiellales, por su parte, incluye especies patógenas, como *Coxiella burnetii*, que es responsable de la fiebre Q en humanos. En estudios sobre microbiomas marinos, se ha observado que los miembros del microbioma de ciertos organismos, como los choanoflagelados, presentan una relación con las bacterias del orden Coxiellales. Estos hallazgos sugieren que las bacterias de este orden podrían desempeñar un papel en la interacción con organismos bacterívoros y patógenos (Needham et al., 2022). Asimismo, el orden Oligoflexales mostró una fuerte correlación con la abundancia de simbiontes algales en anémonas del género Aiptasia. Esta asociación sugiere que las bacterias de Oligoflexales podrían ser simbiontes primarios de las algas dinoflageladas (Aguirre et al., 2023).

En la localidad de Cachipucara, se identificaron 3 géneros bacterianos únicos que presentan características ecológicas relevantes que podrían ser considerados como indicadores potenciales de condiciones ambientales específicas, potencialmente asociadas a entornos impactados por

actividades antropogénicas o condiciones extremas. El grupo de agua dulce FukuN18 ha sido propuesto como un bioindicador relevante en ecosistemas fluviales afectados por contaminación orgánica y nutrientes, especialmente en zonas de alta perturbación antrópica. Yang y colaboradores (2019) demostraron su asociación con áreas altamente degradadas del río Songhua, donde la abundancia de este grupo se correlacionó con niveles elevados de nitrato y otras variables fisicoquímicas vinculadas al deterioro ambiental. Por otro lado, el género Psychrobacter está compuesto por bacterias psicrotróficas y halotolerantes que habitan ambientes fríos como suelos de permafrost, hielo marino y aguas polares. Estudios recientes han revelado que estas bacterias evolucionaron a partir de ancestros asociados a hospedadores, conservando algunos rasgos genómicos y fenotípicos ancestrales a pesar de su transición hacia un estilo de vida libre (Welter et al., 2021). Asimismo, el género IS-44, perteneciente a la familia Nitrosomonadaceae, ha sido identificado en comunidades microbianas asociadas a eventos estacionales de proliferación de arqueas oxidantes de amonio en ecosistemas estuarinos (Rasmussen & Francis, 2022). Aunque aún no se ha confirmado su papel como bacterias oxidantes de amonio, su co-ocurrencia con otros taxones nitrificantes sugiere una posible función en procesos de transformación del nitrógeno.

En la localidad de Yunguyo se identificaron dos órdenes bacterianos únicos con posibles implicancias ecológicas relevantes vinculadas a ambientes acuáticos con alta carga orgánica y procesos biogeoquímicos intensivos. El orden Clostridiales incluye bacterias anaerobias obligadas comúnmente asociadas a la degradación de materia orgánica en ambientes ricos en nutrientes, tales como sedimentos eutróficos y aguas residuales. En estudios realizados en ambientes lacustres eutrofizados, Clostridiales ha sido reportado en estudios previos con una alta eficacia en la remoción de sulfatos y compuestos orgánicos complejos, especialmente bajo condiciones anóxicas

(Custodio et al., 2021). Por otro lado, el orden Omnitrophales identificado en ambientes con condiciones hidrogeoquímicas particulares, ha sido relacionado con procesos del ciclo del carbono y transformaciones redox complejas. En un estudio reciente, se observó que la presencia de Omnitrophales estaba positivamente correlacionada con parámetros como el bicarbonato, el carbono orgánico disuelto, el sulfato y el pH, lo que sugiere una participación activa en procesos de oxidación-reducción y mineralización de compuestos orgánicos en ambientes acuáticos impactados (Li et al., 2024).

En conjunto, las localidades evaluadas evidencian patrones diferenciados en la composición y estructura de las comunidades bacterianas, los cuales estarían posiblemente asociados a la interacción entre factores ambientales y presiones antrópicas locales. Las diferencias espaciales detectadas en este estudio reflejan no solo variaciones en la diversidad bacteriana, sino también cambios potenciales en la funcionalidad microbiana, lo que resalta la importancia de considerar un enfoque integrador para comprender la dinámica de las comunidades bacterianas en sistemas acuáticos influenciados por la acuicultura. Así por ejemplo, en Ichu, la presencia de órdenes bacterianos asociados a microbiota ruminal sugeriría una posible influencia de actividades ganaderas; mientras que en la laguna Lagunillas, la detección de taxones potencialmente patógenos y degradadores de compuestos complejos reflejaría condiciones eutrofizadas vinculadas a la acuicultura intensiva; En el lago Arapa, la predominancia de grupos relacionados con la acumulación y transformación de nutrientes, podría indicar un sistema dinámico influenciado por descargas orgánicas y conectividad hidrológica. Por su parte, Cachipucara presenta géneros asociados a perturbación antrópica y procesos del ciclo del nitrógeno, sugiriendo impactos derivados de la carga de residuos, mientras que en Yunguyo se identifican órdenes vinculados a la

degradación de materia orgánica bajo condiciones de alta carga nutricional. Patrones espaciales y estacionales diferenciados en la estructura de las comunidades bacterianas pudieron ser detectados a través de los análisis de diversidad beta realizados mediante PCoA (Figura 12) y UPGMA (Figura 13). Además, los resultados de este estudio reflejaron el grado de conectividad hidrológica entre las localidades, en especial entre aquellas dentro del lago principal frente a las que se ubican en subcuencas o intercuencas.

Se evidenció una agrupación clara entre las localidades ubicadas dentro del Lago Titicaca (Ichu, Cachipucara y Yunguyo). Estas localidades, a pesar de tener condiciones locales distintas, comparten una mayor conectividad hidrológica directa dentro del cuerpo lacustre principal, lo que podría facilitar el intercambio microbiano, sedimentario y de nutrientes entre zonas. Esta conectividad podría estar modulada por factores como el régimen monomíctico del lago, caracterizado por un ciclo anual de mezcla vertical, y el intercambio constante de masas de agua, especialmente durante la época de lluvias (Pouilly et al., 2014; MINAM, 2014). En contraste, las localidades de la cuenca hidrográfica externa al Lago Titicaca (Arapa y Lagunillas) formaron clústeres separados. Esta separación espacial se explica por su ubicación en una intercuenca (lago Arapa) y una subcuenca (laguna Lagunillas), respectivamente, limitando su interacción hidrológica directa con el cuerpo principal del lago Titicaca. Arapa, al encontrarse conectado al sistema del lago Titicaca a través del río Ramis, recibe influencias tanto de escorrentías agrícolas y urbanas como de descargas domésticas (Ocola & Laqui, 2017; Larico-Robles, 2024), lo cual podría haber generado una comunidad bacteriana diferencial. Por su parte, laguna Lagunillas, al estar represada y localizada en una subcuenca relativamente aislada (Ocola & Laqui, 2017),

presenta una dinámica hidrológica más independiente y menor intercambio de agua con el Lago Titicaca, favoreciendo una estructura microbiana más diferenciada.

Según la composición taxonómica a nivel de familia, se reveló un patrón distinto según la época del año (lluviosa o seca). Esta organización estacional estaría sugiriendo que las comunidades bacterianas responden de manera coherente a los cambios hidrológicos y ambientales propios de cada época del año. Durante la época lluviosa, el aumento en el escurrimiento superficial y la carga de nutrientes y materia orgánica altera significativamente las condiciones físico-químicas, favoreciendo el establecimiento de grupos bacterianos adaptados a condiciones más variables y dinámicas (Pouilly et al., 2014; Siguayro et al., 2022). Cabe resaltar que una excepción notable en este patrón fue la muestra de Lagunillas en 2018, que no se agrupó con las demás muestras de su misma época. Este comportamiento atípico podría estar relacionado a la relativa desconexión hidrológica de la laguna con respecto al sistema principal del Lago Titicaca, dada su ubicación en una subcuenca y su condición de cuerpo represado (Chura-Cruz et al., 2014). Esta independencia reduce el efecto homogenizante de la estacionalidad, lo que podría haber permitido que se mantenga una comunidad bacteriana más estable o distinta del patrón general observado en otras localidades durante la temporada de lluvias.

La predicción funcional de las comunidades bacterianas realizada mediante el enfoque FAPROTAX permitió inferir los posibles roles ecológicos de las comunidades microbianas a partir del análisis de ASVs obtenidos por metabarcoding del gen ARNr 16S. No obstante, cabe destacar que esta estrategia se basa en las asociaciones taxonómicas previamente reportadas y no garantiza la viabilidad de los microorganismos detectados, ya que el ADN ambiental puede provenir de células vivas o muertas (Laroche et al., 2021). Considerando esta limitación, se enfocó el análisis

principalmente en los ASVs más abundantes, bajo el supuesto de que estos representan a los microorganismos activos o dominantes en el ecosistema en el momento de la recolección de muestras. Así, de los 783 OTUs anotados taxonómicamente, solo 33.6% (263 OTUs) pudieron ser asignados a grupos funcionales. Este bajo porcentaje evidencia una importante proporción de taxones desconocidos o escasamente caracterizados, lo cual es característico de ecosistemas de alta montaña donde la diversidad microbiana permanece subexplorada (Logue, Bürgmann, & Robinson, 2008; Louca et al., 2016). La escasa representación funcional disponible subraya la necesidad de futuros estudios de metagenómica que permitan una caracterización más completa del potencial metabólico de las comunidades bacterianas del Lago Titicaca. Por otro lado, se observaron diferencias notables entre localidades, reflejando las condiciones ecológicas particulares de cada una. En Arapa, se identificó una elevada asignación de OTUs a funciones como quimioheterotrofía aeróbica, fermentación y nitrificación, lo que indica un ecosistema donde los procesos de reciclaje de materia orgánica y transformación de nitrógeno son activos, probablemente influenciados por la actividad acuícola, pero sin evidencia clara de procesos anóxicos severos. En Lagunillas, se detectó una diversidad funcional aún mayor, con una presencia importante de funciones asociadas a fotoautotrofía oxigénica, reducción de nitrato y oxidación de metano. Estos resultados sugieren un ecosistema en mejor estado ecológico, con una estructura funcional compleja y resiliente, similar a lo descrito para lagos de bajo impacto antrópico en otras regiones (Picazo et al., 2021).

Por el contrario, en Cachipucara, la asignación de OTUs funcionales fue menor predominando procesos anaeróbicos como la oxidación oscura de compuestos de azufre y la fermentación. Esta señal sugiere condiciones sedimentarias más anóxicas, probablemente

derivadas de una mayor carga de materia orgánica y baja oxigenación, características típicas de ambientes eutrofizados o impactados por actividades humanas intensivas (Sun et al., 2013). En Yunguyo, los perfiles funcionales estuvieron dominados por funciones asociadas a patogenicidad humana, simbiosis animal y degradación de hidrocarburos, reflejando un estado ecológico más alterado, posiblemente por descargas domésticas o el manejo intensivo de actividades acuícolas, como ha sido reportado en otros sistemas contaminados (Custodio et al., 2021; Wang et al., 2023).

La representación funcional también presentó un patrón estacional evidente (Figura 14). Durante la época seca, las funciones predominantes estuvieron asociadas a procesos aeróbicos, como la nitrificación, la oxidación de amonio y la oxidación de nitrito, procesos típicos de ambientes bien oxigenados y de mayor estabilidad hidrológica (Newton et al., 2011). En contraste, durante la época lluviosa, las funciones anaeróbicas, como la fermentación y la degradación de compuestos aromáticos, fueron más prevalentes, probablemente debido al incremento de materia orgánica y al desarrollo de condiciones más reducidas provocado por el mayor aporte de escorrentía (Niño-García et al., 2017). Si bien los enfoques predictivos basados en el ADN metabarcoding representan una herramienta valiosa y costo-eficiente para el monitoreo ambiental (Djemiel et al., 2022; Laroche et al., 2021), las predicciones deben interpretarse con cautela. Las funciones asignadas no implican necesariamente que estas actividades metabólicas estén siendo expresadas activamente en el ambiente, debido a la posible redundancia funcional entre linajes microbianos y a la limitación inherente de los métodos de inferencia a partir del ARNr 16S (Louca et al., 2016). A pesar de estas limitaciones, los resultados obtenidos permiten inferir tendencias funcionales relevantes que reflejan los impactos espaciales y estacionales en las comunidades bacterianas del Lago Titicaca.

CONCLUSIONES

- Las comunidades bacterianas en las áreas de cultivo de *O. mykiss* en la cuenca del Lago Titicaca presentaron una alta diversidad filogenética, dominada por los filos Bacteroidota, Acidobacteriota, Cyanobacteria, Verrucomicrobiota, Actinobacteriota y Proteobacteria, reflejando una estructura microbiana compleja.

- Los perfiles taxonómicos permitieron identificar patrones a nivel espacial, como asociados al potencial patogénico en Lagunillas, aportes orgánicos ganaderos en Ichu, una alta riqueza y número de ASVs exclusivos en Arapa, y asociados a estrés ambiental y baja diversidad en Yunguyo y Cachipucara.

- A nivel temporal, la estacionalidad actuó como un factor determinante de la diversidad, registrándose una mayor riqueza de ASVs y valores del índice de Shannon durante la época seca (especialmente en Arapa y Lagunillas), mientras que mediante la diversidad beta se evidencian patrones espacio-temporales asociados a la estacionalidad y la conectividad hidrológica.

- El bajo porcentaje (33.6%) de asignación funcional resalta el carácter subexplorado del lago, destacando una asignación asociada a procesos de eutrofización en Yunguyo y Arapa, funciones asociadas a patógenos en Ichu Lagunillas y Yunguyo, y una menor diversidad funcional en Cachipucara.

RECOMENDACIONES

- Para profundizar y ampliar los resultados obtenidos en este estudio, se recomienda realizar un análisis continuo de las comunidades bacterianas en la cuenca del Lago Titicaca, lo que permitiría observar la dinámica de las comunidades microbianas a lo largo de varios años. Esto es particularmente importante para evaluar cómo las variaciones estacionales y los cambios climáticos a largo plazo afectan la estructura bacteriana, y para establecer patrones más robustos que ayuden en la conservación y manejo de los recursos acuáticos.
- Sería beneficioso ampliar la cobertura de muestreo a otras zonas dentro del Lago Titicaca y sus alrededores, especialmente a áreas alejadas de la actividad humana. Esto permitiría obtener una representación más completa de la microbiota bacteriana y una comparación más robusta entre zonas más influenciadas por actividades antropogénicas, como la acuicultura, y aquellas más naturales o menos afectadas por estas prácticas.
- El análisis de la variabilidad estacional se podría complementar con datos meteorológicos más detallados que incluyan no solo las estaciones lluviosa y seca, sino también variables como temperatura, precipitación y humedad. Esto ayudaría a comprender mejor la relación entre los factores climáticos y la estructura bacteriana, y cómo estos influyen en las funciones microbianas en diferentes momentos del año.

REFERENCIAS BIBLIOGRÁFICAS

Aguirre, E. G., Fine, M. J., & Kenkel, C. D. (2023). Abundance of Oligoflexales bacteria is associated with algal symbiont density, independent of thermal stress in Aiptasia anemones. *Ecology and Evolution, 13*(12). https://doi.org/10.1002/ece3.10805

Amaru Chambilla, G. R., Amaro Chambilla, G., Yucra Quispe, M. Á., & Orna Rivas, E. F. (2024). Design of a floating cage prototype with organic waste collector (food remains and excreta), produced by trout farming in lake Titicaca, Perú. *Revista de Investigación e Innovación Agropecuaria y de Recursos Naturales, 11*(1), 39-46. http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S2409-16182024000100039

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data* [Online]. Disponible en: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Arnold J (2025). *ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'*. Paquete de R versión 5.2.0. https://doi.org/10.32614/CRAN.package.ggthemes

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., Huttley, G. A., & Gregory Caporaso, J. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome, 6*(1). https://doi.org/10.1186/s40168-018-0470-z

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M.,

Chase, J., & Cope, E. K. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology, 37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bukin, Y. S., Galachyants, Y. P., Morozov, I. V., Bukin, S. V., Zakharenko, A. S., & Zemskaya, T. I. (2019). The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Scientific Data, 6*(1). https://doi.org/10.1038/sdata.2019.7

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods, 13(7)*, 581–583. https://doi.org/10.1038/nmeth.3869

Chura, R., & Mollocondo, H. (2009). Desarrollo de la acuicultura en el Lago Titicaca (Perú). *AquaTIC, (31)*, 6-19. http://www.redalyc.org/articulo.oa?id=49422781002

Chura-Cruz, R., Robles Trujillo, J. H., Yujra Flores, E., Puntaca Quispe, L. A., Yana Tipo, W., Casas Yupanqui, C., & Siguayro, H. (2015). *Monitoreo ecológico y limnológico de la laguna Lagunillas, Lampa - Puno, 2014 (Informe técnico, Convenio IMARPE–PELT)*. https://doi.org/10.13140/RG.2.1.1592.6004

Clols-Fuentes, J., Nguinkal, J. A., Unger, P., Kreikemeyer, B., & Palm, H. W. (2024). Bacterial Communities From Two Freshwater Aquaculture Systems in Northern Germany. *Environmental Microbiology Reports*, *16*(6). https://doi.org/10.1111/1758-2229.70062

Cornejo-Sánchez, A. C. (2022). *Caracterización espacio temporal de la contaminación por acuicultura en las aguas superficiales de la laguna Lagunillas* [Tesis de licenciatura,

Universidad Nacional del Altiplano]. http://repositorio.unap.edu.pe/handle/20.500.14082/18337

Custodio, M., Peñaloza, R., Espinoza, C., Espinoza, W., & Mezarina, J. (2021). Treatment of dairy industry wastewater using bacterial biomass isolated from eutrophic lake sediments for the production of agricultural water. *Bioresource Technology Reports, 17*, 100891–100891. https://doi.org/10.1016/j.biteb.2021.100891

Custodio, M., Espinoza, C., Peñaloza, R., Tessy Peralta-Ortiz, Héctor Sánchez-Suárez, Ordinola-Zapata, A., & Enedia Vieyra-Peña. (2022). Microbial diversity in intensively farmed lake sediment contaminated by heavy metals and identification of microbial taxa bioindicators of environmental quality. *Scientific Reports, 12*(1). https://doi.org/10.1038/s41598-021-03949-7

Djemiel, C., Maron, P. A., Terrat, S., Dequiedt, S., Cottin, A., & Ranjard, L. (2022). Inferring microbiota functions from taxonomic genes: a review. *GigaScience, 11*. https://doi.org/10.1093/gigascience/giab090

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics, 32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Fasolo, A., Deb, S., Stevanato, P., Concheri, G., & Squartini, A. (2024). ASV vs OTUs clustering: Effects on alpha, beta, and gamma diversities in microbiome metabarcoding studies. *PLOS One, 19*(10). https://doi.org/10.1371/journal.pone.0309065

Findlay, D. L., Podemski, C. L., & Kasian, S. E. M. (2009). Aquaculture impacts on the algal and bacterial communities in a small boreal forest lake. *Canadian Journal of Fisheries and Aquatic Sciences*, *66*(11), 1936–1948. https://doi.org/10.1139/f09-121

Garner, R. E., Kraemer, S. A., Onana, V. E., Fradette, M., Varin, M.-P., Huot, Y., & Walsh, D. A. (2023). A genome catalogue of lake bacterial diversity and its drivers at continental scale. *Nature Microbiology, 8*(10), 1920–1934. https://doi.org/10.1038/s41564-023-01435-6

Glöckner, F. O., Yilmaz, P., Quast, C., Gerken, J., Beccati, A., Ciuprina, A., Bruns, G., Yarza, P., Peplies, J., Westram, R., & Ludwig, W. (2017). 25 years of serving the community with ribosomal RNA gene reference databases and tools. *Journal of Biotechnology, 261*, 169–176. https://doi.org/10.1016/j.jbiotec.2017.06.1198

Hamady, M., Walker, J. J., Harris, J. K., Gold, N. J., & Knight, R. (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nature Methods, 5*(3), 235–237. https://doi.org/10.1038/nmeth.1184

Hamady, M., & Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research, 19*(7), 1141–1152. https://doi.org/10.1101/gr.085464.108

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLOS Computational Biology, 12*(6), e1004957. https://doi.org/10.1371/journal.pcbi.1004957

Illumina. (2013). *16S Metagenomic Sequencing Library Preparation. Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System*. [Archivo PDF] https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf

Illumina. (2024). MiSeq™ System, Velocidad y simplicidad para la secuenciación selectiva y la secuenciación de genomas pequeños. [Archivo PDF] https://www.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/miseq-system-data-sheet-m-gl-00006/miseq-data-sheet-m-gl-00006-esp.pdf

Klindworth, A., Elmar Pruesse, Schweer, T., Jörg Peplies, Quast, C., Horn, M., & Glöckner, F. O. (2012). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research, 41*(1), e1–e1. https://doi.org/10.1093/nar/gks808

Łach, Ł., Khomutovska, N., Kwiatowski, J., & Jasser, I. (2024). Testing 16S Primers for Proper Identification of Cyanobacterial Communities in Small Water Bodies. *Water, 16*(10). https://doi.org/10.3390/w16101357

Larico-Robles, I. F. (2024). *Variación de la calidad del agua del lago Arapa-Chupa por vertimiento de aguas residuales domésticas, por la Municipalidad Distrital de Chupa* [Tesis de licenciatura, Universidad Andina Néstor Cáceres Velásquez]. Dspace. https://repositorio.uancv.edu.pe/items/825c839b-3405-46ea-b90f-86b35c486095

Laroche, O., Pochon, X., Wood, S. A., & Keeley, N. (2021). Beyond taxonomy: Validating functional inference approaches in the context of fish-farm impact assessments. *Molecular Ecology Resources, 21*(7), 2264–2277. https://doi.org/10.1111/1755-0998.13426

Lee, H. B., Jeong, D. H., Cho, B. C., & Park, J. S. (2023). Comparative analyses of eight primer sets commonly used to target the bacterial 16S rRNA gene for marine metabarcoding-based studies. *Frontiers in Marine Science, 10*. https://doi.org/10.3389/fmars.2023.1199116

Li, C., Quan, Q., Gan, Y., Dong, J., Fang, J., Wang, L., & Liu, J. (2020). Effects of heavy metals on microbial communities in sediments and establishment of bioindicators based on microbial taxa and function for environmental monitoring and management. *Science of the Total Environment*, *749*, 141555. https://doi.org/10.1016/j.scitotenv.2020.141555

Li, Y., Su, Z., Xiu, W., Huang, L., Huang, T., & Zheng, J. (2024). Exploring Methane Capture Potential in Alkaline Coal Mine Drainage: Insight from the Microbial Community Structure and Function Analysis. *Water, 16*(13), 1915. https://doi.org/10.3390/w16131915

Lima, N., Ruelland, D., Escobar, M., Rabatel, A., Lavado, W., & Condom, T. (2024). Integrated water balance reveals that Lake Titicaca is driven by extreme climate variability (Nos. EGU24-12006). EGU24. *Copernicus Meetings*. https://doi.org/10.5194/egusphere-egu24-12006

Lin, L., & Ju, F. (2023). Evaluation of different 16S rRNA gene hypervariable regions and reference databases for profiling engineered microbiota structure and functional guilds in

a swine wastewater treatment plant. *Interface Focus, 13*(4). https://doi.org/10.1098/rsfs.2023.0012

Logue, J. B., Bürgmann, H., & Robinson, C. T. (2008). Progress in the Ecological Genetics and Biodiversity of Freshwater Bacteria. *BioScience, 58*(2), 103–113. https://doi.org/10.1641/b580205

López-Aladid, R., Fernández-Barat, L., Alcaraz-Serrano, V., Bueno-Freire, L., Vázquez, N., Pastor-Ibáñez, R., Palomeque, A., Oscanoa, P., & Torres, A. (2023). Determining the most accurate 16S rRNA hypervariable region for taxonomic identification from respiratory samples. *Scientific Reports, 13*(1), 3974. https://doi.org/10.1038/s41598-023-30764-z

Louca, S., Parfrey, L. W., & Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science, 353*(6305), 1272–1277. https://doi.org/10.1126/science.aaf4507

Maldonado, J. (2021). *Qiime2-r16S: Guía de comandos (Dada2)*. Maldonado Lab. https://maldonadolab.com/qiime2-r16s-guia-de-comandos-dada2/

Mamani, R. M. (2018). Estudio histórico del impacto de actividades antropogénicas en los humedales y totorales de la zona sur de la bahía interior del Lago Titicaca. *Zenodo*. https://doi.org/10.5281/zenodo.1407806

Mantilla Mendoza, B. (2008). *Efectos de la acuicultura en jaulas flotantes en el ecosistema acuático de la bahía de Chucuito, Lago Titicaca* [Tesis de maestría - Universidad Nacional

del Altiplano]. Repositorio Institucional - UNAP. http://repositorio.unap.edu.pe/handle/20.500.14082/616

Maquera-Lupaca, R., Fernández-Asqui, G. A., Sardón-Ari, Z. Y., & Layme-Chura, S. (2024). Gobernanza ambiental para la mitigación de la contaminación del Lago Titicaca, Perú 2022. *Revista de Climatología, 24*, 938-944. https://rclimatol.eu/2024/02/04/gobernanza-ambiental-para-la-mitigacion-de-la-contaminacion-del-lago-titicaca-peru-una-revision-sistematica/

Maynez-Perez, A., Jahuey-Martínez, F. J., Martínez-Quintana, J. A., Hume, M. E., Anderson, R. C., Agustín Corral-Luna, Rodríguez-Almeida, F. A., Yamicela Castillo-Castillo, & Monserrath Felix-Portillo. (2024). The Rumen Microbiome Composition of Raramuri Criollo and European Cattle in an Extensive System. *Microorganisms, 12*(11), 2203–2203. https://doi.org/10.3390/microorganisms12112203

Milyanchuk, N., Kuchko, Y., & Ilmast, N. (2024). The impact of fish farming on the Tarasmozero Lake ecosystem, Onega Lake basin. *BIO Web of Conferences, 118*, 02002. https://doi.org/10.1051/bioconf/202411802002

Ministerio del Ambiente (MINAM). (2014). *Línea base ambiental de la cuenca del Lago Titicaca*. Repositorio Institucional - ANA. https://repositorio.ana.gob.pe/handle/20.500.12543/4435

Ministerio de la Producción (PRODUCE). (16 de enero de 2022). *Buscar concesiones y autorizaciones acuícolas*. Gop.pe. https://www.gob.pe/7700-buscar-concesiones-y-autorizaciones-acuicolas

Montesinos López, Jeansen Aníbal. (2017). *Diagnóstico situacional de la crianza de truchas arco iris (Oncorhynchus mykiss) en centros de cultivo del Lago Titicaca* [Tesis de maestría, Universidad Peruana Cayetano Heredia] Repositorio Institucional – Universidad Peruana Cayetano Heredia. https://hdl.handle.net/20.500.12866/3862

Mou, X., Jacob, J., Lu, X., Robbins, S., Sun, S., & Ortiz, J. D. (2013). Diversity and distribution of free-living and particle-associated bacterioplankton in Sandusky Bay and adjacent waters of Lake Erie Western Basin. *Journal of Great Lakes Research*, *39*(2), 352–357. https://doi.org/10.1016/j.jglr.2013.03.014

Needham, D. M., Poirier, C., Bachy, C., George, E. E., Wilken, S., Yung, C. C. M., Limardo, A. J., Morando, M., Sudek, L., Malmstrom, R. R., Keeling, P. J., Santoro, A. E., & Worden, A. Z. (2022). The microbiome of a bacterivorous marine choanoflagellate contains a resource-demanding obligate bacterial associate. *Nature Microbiology, 7*(9), 1466–1479. https://doi.org/10.1038/s41564-022-01174-0

Neuwirth, E. (2022). *RColorBrewer: ColorBrewer Palettes*. Paquete R versión 1.1-3. https://CRAN.R-project.org/package=RColorBrewer

Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D., & Bertilsson, S. (2011). A Guide to the Natural History of Freshwater Lake Bacteria. *Microbiology and Molecular Biology Reviews, 75*(1), 14–49. https://doi.org/10.1128/mmbr.00028-10

Niño-García, J. P., Ruiz-González, C., & A. del Giorgio, P. (2017). Exploring the Ecological Coherence between the Spatial and Temporal Patterns of Bacterioplankton in Boreal Lakes. *Frontiers in Microbiology, 8*. https://doi.org/10.3389/fmicb.2017.00636

Ocola, J. J., & Laqui, W. F. (2017). *Fuentes contaminantes en la cuenca del lago Titicaca: Un aporte al conocimiento de las causas que amenazan la calidad del agua del maravilloso lago Titicaca*. Autoridad Nacional del Agua. https://repositorio.ana.gob.pe/handle/20.500.12543/636?show=full

Ocola Villasante, D. (2022). Evaluación de las características físico químicas del agua y la composición del sustrato en zonas de crianza intensiva de trucha (Oncorhynchus mykiss) en el lago Titicaca, Pomata-Puno. *Ñawparisun - Revista de Investigación Científica, 4* (1), 19–29. https://doi.org/10.47190/nric.v4i1.2

Orna Rivas, E. F., & Larico Mamani, C. J. (2023). Impactos de la truchicultura en los nutrientes del agua y macroinvertebrados bentónicos en la bahía de Puno, lago Titicaca: Impacts of trout farming on water nutrients and benthic macroinvertebrates in the bay of puno, lake Titicaca. *Investigación Universitaria UNU, 13*(1), 949. https://doi.org/10.53470/riu.v13i1.109

Pedregosa, F., Gaël Varoquaux, Alexandre Gramfort, Michel, V., Thirion, B., Grisel, O., Mathieu Blondel, Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Édouard Duchesnay. (2026). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*(85), 2825–2830. https://jmlr.org/papers/v12/pedregosa11a.html

Picazo, A., Villaescusa, J. A., Rochera, C., Miralles-Lorenzo, J., Quesada, A., & Camacho, A. (2021). Functional Metabolic Diversity of Bacterioplankton in Maritime Antarctic Lakes. *Microorganisms, 9*(10), 2077–2077. https://doi.org/10.3390/microorganisms9102077

Posit team (2025*). RStudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA. http://www.posit.co/.

Pouilly, M., Lazzaro, X., Point, D., & Aguirre, M. (2014). *Línea base de conocimientos sobre los recursos hidrológicos e hidrobiológicos en el sistema TDPS con enfoque en la cuenca del Lago Titicaca*. IRD - UICN, Quito, Ecuador, 320 pp. https://portals.iucn.org/library/node/44673

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE, 5*(3), e9490. https://doi.org/10.1371/journal.pone.0009490

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research, 41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

Qin, M., Xu, H., Zeng, J., Zhao, D., Yu, Z., & Wu, Q. L. (2020). Composition and assembly of bacterial communities in surface and deeper sediments from aquaculture-influenced sites in Eastern Lake Taihu, China. *Aquatic Sciences*, *82*(4). https://doi.org/10.1007/s00027-020-00755-7

Ramos, V. E., Quispe, J. A., & Elias, R. K. (2019). Evaluación de la abundancia relativa de *Telmatobius culeus* en la zona litoral del lago Titicaca, Perú. *Revista Peruana de Biología, 26*(4). https://doi.org/10.15381/rpb.v26i4.17216

Rasmussen, A. N., & Francis, C. A. (2022). Genome-Resolved Metagenomic Insights into Massive Seasonal Ammonia-Oxidizing Archaea Blooms in San Francisco Bay. *MSystems, 7*(1). https://doi.org/10.1128/msystems.01270-21

Regueira-Iglesias, A., Balsa-Castro, C., Blanco-Pintos, T., & Tomás, I. (2023). Critical review of 16S rRNA gene sequencing workflow in microbiome studies: From primer selection to advanced data analysis. *Molecular Oral Microbiology, 38*(5), 347–399. https://doi.org/10.1111/omi.12434

Requena-Castro, R., Aguilera-Arreola, M. G., Martínez-Vázquez, A. V., Cruz-Pulido, W. L., Rivera, G., Fernández-Dávila, S., Flores-Magallón, R., Acosta-Cruz, E., & Bocanegra-García, V. (2025). Identification of Bacterial Communities in Surface Waters of Rio Bravo/Rio Grande Through 16S rRNA Gene Metabarcoding. *Water, 17*(11). https://doi.org/10.3390/w17111575

Rieder, J., Kapopoulou, A., Bank, C., & Adrian-Kalchhauser, I. (2023). Metagenomics and metabarcoding experimental choices and their impact on microbial community characterization in freshwater recirculating aquaculture systems. *Environmental Microbiome, 18*(1), 8. https://doi.org/10.1186/s40793-023-00459-z

Robeson, M. S., II, O'Rourke, D. R., Kaehler, B. D., Ziemski, M., Dillon, M. R., Foster, J. T., & Bokulich, N. A. (2020). RESCRIPt: Reproducible sequence taxonomy reference database

management. *PLOS Computational Biology, 17*(11), e1009581. https://doi.org/10.1371/journal.pcbi.1009581

Sadeghi, J., Chaganti, S. R., Shahraki, A. H., & Heath, D. D. (2021). Microbial community and abiotic effects on aquatic bacterial communities in north temperate lakes. *Science of the Total Environment*, *781*, 146771. https://doi.org/10.1016/j.scitotenv.2021.146771

Schallenberg, L. A., Pearman, J. K., Burns, C. W., & Wood, S. A. (2021). Spatial abundance and distribution of picocyanobacterial communities in two contrasting lakes revealed using environmental DNA metabarcoding. *FEMS Microbiology Ecology*, *97*(7). https://doi.org/10.1093/femsec/fiab075

Servicio Nacional de Áreas Protegidas por el Estado (SERNANP). (1 de enero 2019). *Reserva Nacional del Titicaca*. Gop.pe. https://www.gob.pe/institucion/sernanp/informes-publicaciones/1749491-reserva-nacional-del-titicaca

Shafi, S., Kamili, A. N., Shah, M. A., Parray, J. A., & Bandh, S. A. (2017). Aquatic bacterial diversity: Magnitude, dynamics, and controlling factors. *Microbial Pathogenesis*, *104*, 39–47. https://doi.org/10.1016/j.micpath.2017.01.016

Shang, J., Zhang, W., Chen, X., Li, Y., Niu, L., Wang, L., & Zhang, H. (2021). How environmental stress leads to alternative microbiota states in a river ecosystem: A new insight into river restoration. *Water Research, 203*. https://doi.org/10.1016/j.watres.2021.117538

Shen, M.-C., Zhang, Y.-Z., Bo, G.-D., Yang, B., Wang, P., Ding, Z.-Y., Wang, Z.-B., Yang, J.-M., Zhang, P., & Yuan, X.-L. (2022). Microbial Responses to the Reduction of Chemical

Fertilizers in the Rhizosphere Soil of Flue-Cured Tobacco. *Frontiers in Bioengineering and Biotechnology, 9*. https://doi.org/10.3389/fbioe.2021.812316

Shen, Z., Xie, G., Gong, Y., Shao, K., Gao, G., & Tang, X. (2024). Seasonal dynamics of environmental heterogeneity augment microbial interactions by regulating community structure in different trophic lakes. *Environmental Research, 263*, 120031. https://doi.org/10.1016/j.envres.2024.120031

Siguayro, H., Pasapera, J., Villanueva, C., Coila, Y., & Gamarra, C. (2022). Evaluación de fuentes contaminantes en el anillo circunlacustre del lago Titicaca (sector peruano), 2017. *Boletín Instituto del Mar del Perú, 37*(2), 361–386. https://doi.org/10.53554/boletin.v37i2.378

Skopina, M. Yu., Kuznetsova, E. V., Tikhonenkov, D. V., Zhang, H., Gerasimova, E. A., Tsvetkov, A. I., & Kosolapov, D. B. (2024). Distribution, community structure and metabolic potential of bacterioplankton in a small boreal lake: Microscopy-based survey and 16S rRNA-based DNA metabarcoding. *Limnologica*, *109*, 126213. https://doi.org/10.1016/j.limno.2024.126213

Stabili, L., Di Salvo, M., Alifano, P., & Talà, A. (2021). An Integrative, Multiparametric Approach for the Comprehensive Assessment of Microbial Quality and Pollution in Aquaculture Systems. *Microbial Ecology*, *83*(2), 271–283. https://doi.org/10.1007/s00248-021-01731-w

Stoeck, T., Frühe, L., Forster, D., Cordier, T., Martins, C. I. M., & Pawlowski, J. (2018). Environmental DNA metabarcoding of benthic bacterial communities indicates the benthic

footprint of salmon aquaculture. *Marine Pollution Bulletin, 127*, 139-149. https://doi.org/10.1016/j.marpolbul.2017.11.065

Sun, M. Y., Dafforn, K. A., Johnston, E. L., & Brown, M. V. (2013). Core sediment bacteria drive community response to anthropogenic contamination over multiple environmental gradients. *Environmental Microbiology, 15*(9), 2517–2531. https://doi.org/10.1111/1462-2920.12133

Suzuki, S., Otomo, Y., Dazai, A., Abe, T., & Kondoh, M. (2024). Assessing the impacts of aquaculture on local fish communities using environmental DNA metabarcoding analysis. *Environmental DNA, 6*(3), e551. https://doi.org/10.1002/edn3.551

Thomas, V., Casson, N., & Greub, G. (2006). *Criblamydia sequanensis*, a new intracellular Chlamydiales isolated from Seine river water using amoebal co-culture. *Environmental Microbiology, 8*(12), 2125–2135. https://doi.org/10.1111/j.1462-2920.2006.01094.x

Vásquez-Quispesivana, W., Talavera-Núñez, M., & Inga-Guevara, M. (2016). Evaluación del impacto en la calidad de agua debido a la producción semi intensiva de trucha (*Oncorhynchus mykiss*) en jaulas flotantes en la laguna Arapa - Puno. *Revista de la Sociedad Química del Perú, 82*(1), 15–28. http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1810-634X2016000100003&lng=es&tlng=es

Vesamäki, J. S., Riitta Nissinen, Kainz, M. J., Matthias Pilecky, Marja Tiirola, & Taipale, S. J. (2022). Decomposition rate and biochemical fate of carbon from natural polymers and

microplastics in boreal lakes. *Frontiers in Microbiology, 13*. https://doi.org/10.3389/fmicb.2022.1041242

Vettorazzo, S., Boscaini, A., Cerasino, L., & Salmaso, N. (2024). From small water bodies to lakes: Exploring the diversity of freshwater bacteria in an Alpine Biosphere Reserve. *The Science of the Total Environment*, 176495–176495. https://doi.org/10.1016/j.scitotenv.2024.176495

Villwock, W. (1993). Die Titicaca-See-Region auf dem Altiplano von Peru und Bolivien und die Folgen eingeführter Fische für Wildarten und ihren Lebensraum. *Naturwissenschaften, 80*(1), 1-8. https://doi.org/10.1007/BF01139750

Wan, X., Lee, A. J., Hou, S., Ushijima, B., Nguyen, Y. P., Thawley, J. A., & Donachie, S. P. (2016). Draft Genome Sequence of *Piscirickettsia litoralis*, Isolated from Seawater. *Genome Announcements, 4*(6). https://doi.org/10.1128/genomea.01252-16

Wang, S., Zhang, P., Zhang, D., & Chang, J. (2023). Evaluation and comparison of the benthic and microbial indices of biotic integrity for urban lakes based on environmental DNA and its management implications. *Journal of Environmental Management, 341*, 118026–118026. https://doi.org/10.1016/j.jenvman.2023.118026

Wang, Q., Yu, J., Li, X., Zhang, Y., Zhang, J., Wang, J., Mu, J., Yu, X., & Hui, R. (2024). Seasonal and anthropogenic influences on bacterioplankton communities: ecological impacts in the coastal waters of Qinhuangdao, Northern China. *Frontiers in Microbiology, 15*. https://doi.org/10.3389/fmicb.2024.1431548

Welter, D. K., Ruaud, A., Henseler, Z. M., De Jong, H. N., van Coeverden de Groot, P., Michaux, J., Gormezano, L., Waters, J. L., Youngblut, N. D., & Ley, R. E. (2021). Free-Living, Psychrotrophic Bacteria of the Genus Psychrobacter Are Descendants of Pathobionts. *MSystems, 6*(2). https://doi.org/10.1128/msystems.00258-21

Wickham H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., & Takahashi, K. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham H. (2023). *conflicted: An Alternative Conflict Resolution Strategy*. Paquete R versión 1.2.0. https://doi.org/10.32614/CRAN.package.conflicted

Wickham H (2025). *forcats: Tools for Working with Categorical Variables (Factors)*. Paquete R versión 1.0.1. https://forcats.tidyverse.org/.

Wickham, H., Hester, J., Bryan, J. (2025). *readr: Read Rectangular Text Data*. Paquete R versión 2.1.6. https://doi.org/10.32614/CRAN.package.readr

Wickham, H., François, R., Henry, L., Mülle,r K., Vaughan, D. (2026). *dplyr: A Grammar of Data Manipulation*. Paquete R versión 1.2.0. https://doi.org/10.32614/CRAN.package.dplyr

Willis, A. D. (2019). Rarefaction, alpha diversity, and statistics. *Frontiers in microbiology, 10*, 2407. https://doi.org/10.3389/fmicb.2019.02407

Yan, L. (2025). *ggvenn: Draw Venn Diagram by 'ggplot2'*. Paquete R versión 0.1.19 https://doi.org/10.32614/CRAN.package.ggvenn

Yang, B., Wang, Y., & Qian, P. Y. (2016). Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics*, *17*(1). https://doi.org/10.1186/s12859-016-0992-y

Yang, Y., Li, S., Gao, Y., Chen, Y., & Zhan, A. (2019). Environment-driven geographical distribution of bacterial communities and identification of indicator taxa in Songhua River. *Ecological Indicators, 101*, 62–70. https://doi.org/10.1016/j.ecolind.2018.12.047

Yapuchura-Saico, C. R., Mamani-Choque, S. E., Pari Quispe, D., & Flores Mamani, E. (2018). Importance of the curves of growth and efficiency in the food of trout arcoiris (*Oncorhynchus Mikyss*) in the cost of production. *Comuni@cción, 9*(1), 68-77. http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S2219-71682018000100007&lng=es&tlng=en.

Zhang, L., Fang, S., Hong, W., Shen, Z., Li, S., & Fang, W. (2023). Differences in pathogenic community assembly processes and their interactions with bacterial communities in river and lake ecosystems. *Environmental Research, 236*, 116847. https://doi.org/10.1016/j.envres.2023.116847

Zhang, Z., Lu, J., Zhang, S., Tian, Z., Feng, C., & Liu, Y. (2024). Analysis of bacterial community structure, functional variation, and assembly mechanisms in multi-media habitats of lakes during the frozen period. *Ecotoxicology and Environmental Safety, 284*, 116903. https://doi.org/10.1016/j.ecoenv.2024.116903

Zhong, Y., Xue, M.-Y., Sun, H.-Z., Valencak, T. G., Guan, L. L., & Liu, J. (2020). Rumen and Hindgut Bacteria Are Potential Indicators for Mastitis of Mid-Lactating Holstein Dairy Cows. *Microorganisms,* *8*(12), 2042–2042. https://doi.org/10.3390/microorganisms8122042

Zhou, Y., Cheng, L., Lian, Y., Feng, J., Zhou, M., Jing, D., Yin, W., Wang, H., & Liu, L. (2024). High-proportions of tailwater discharge alter microbial community composition and assembly in receiving sediments. *Scientific Reports, 14*(1). https://doi.org/10.1038/s41598-024-63859-2

Ziesemer, K. A., Mann, A. E., Sankaranarayanan, K., Schroeder, H., Ozga, A. T., Brandt, B. W., Zaura, E., Waters-Rist, A., Hoogland, M., Salazar-García, D. C., Aldenderfer, M., Speller, C., Hendy, J., Weston, D. A., MacDonald, S. J., Thomas, G. H., Collins, M. J., Lewis, C. M., Hofman, C., & Warinner, C. (2015). Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. *Scientific Reports*, *5*(1). https://doi.org/10.1038/srep16498

Zinger, L., Gobet, A., & Pommier, T. (2011). Two decades of describing the unseen majority of aquatic microbial diversity. *Molecular Ecology,* *21*(8), 1878–1896. https://doi.org/10.1111/j.1365-294x.2011.05362.x

ANEXOS

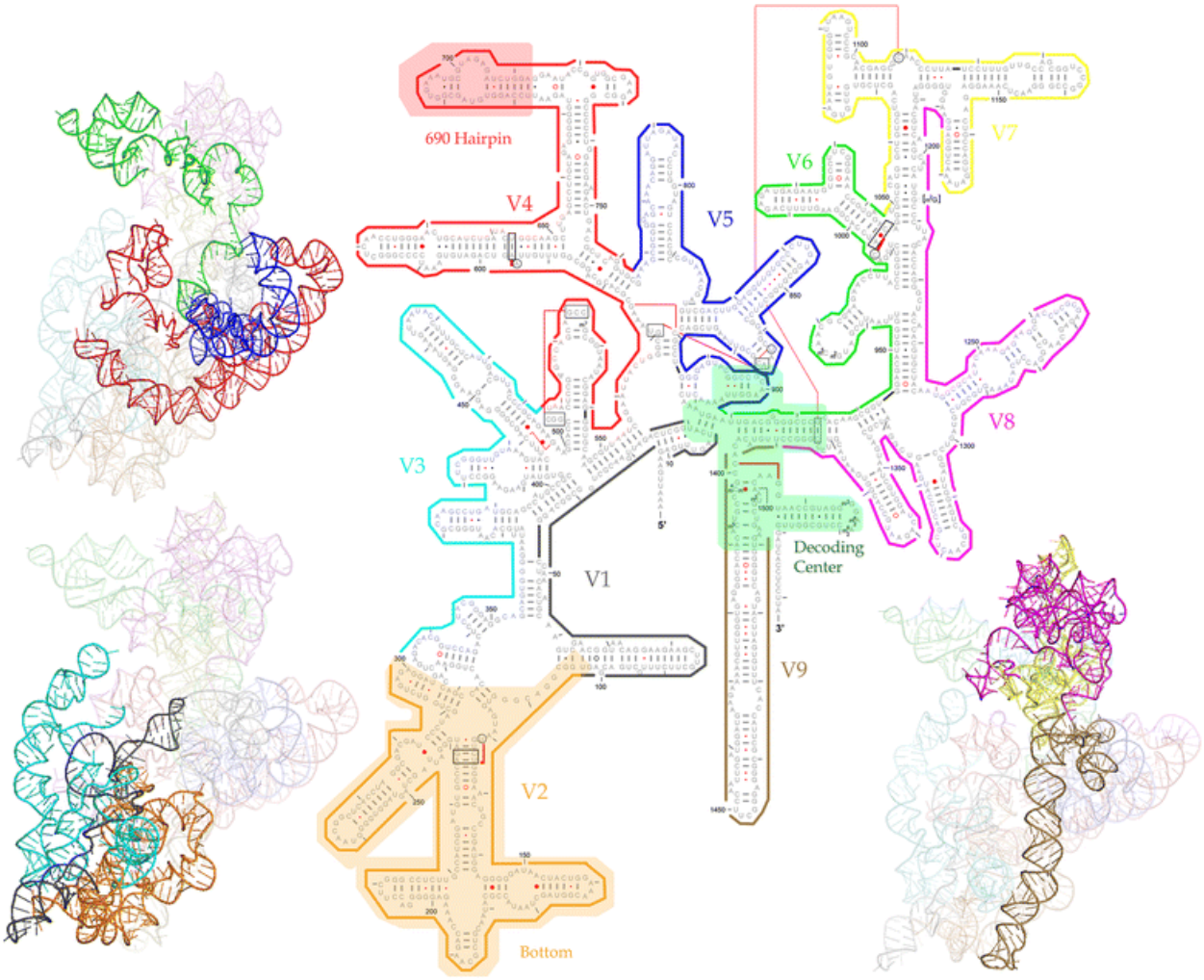


Figura S1. Esquema de las regiones hipervariables del gen ARNr 16S de procariotas (Fuente: Yang et al., 2016).

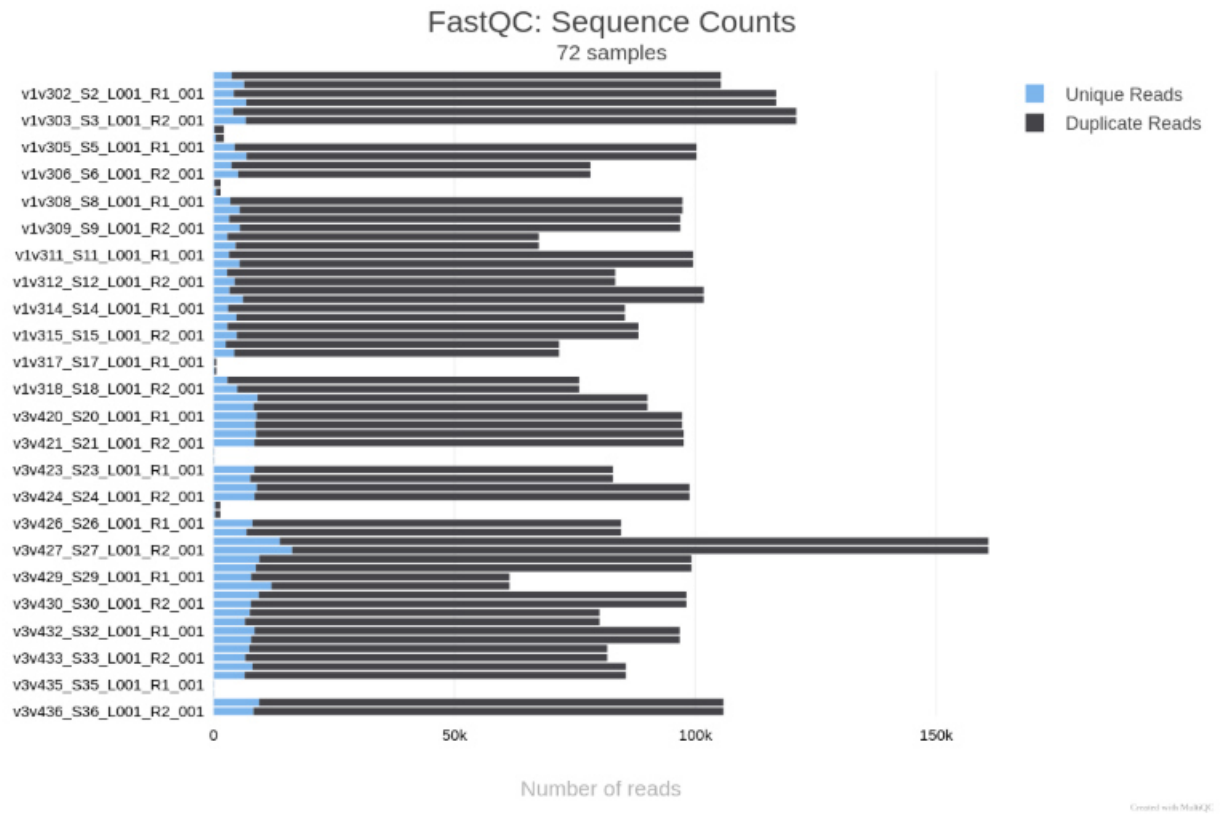


Figura S2. Gráfico de barras del número de lecturas de los datos de secuenciación

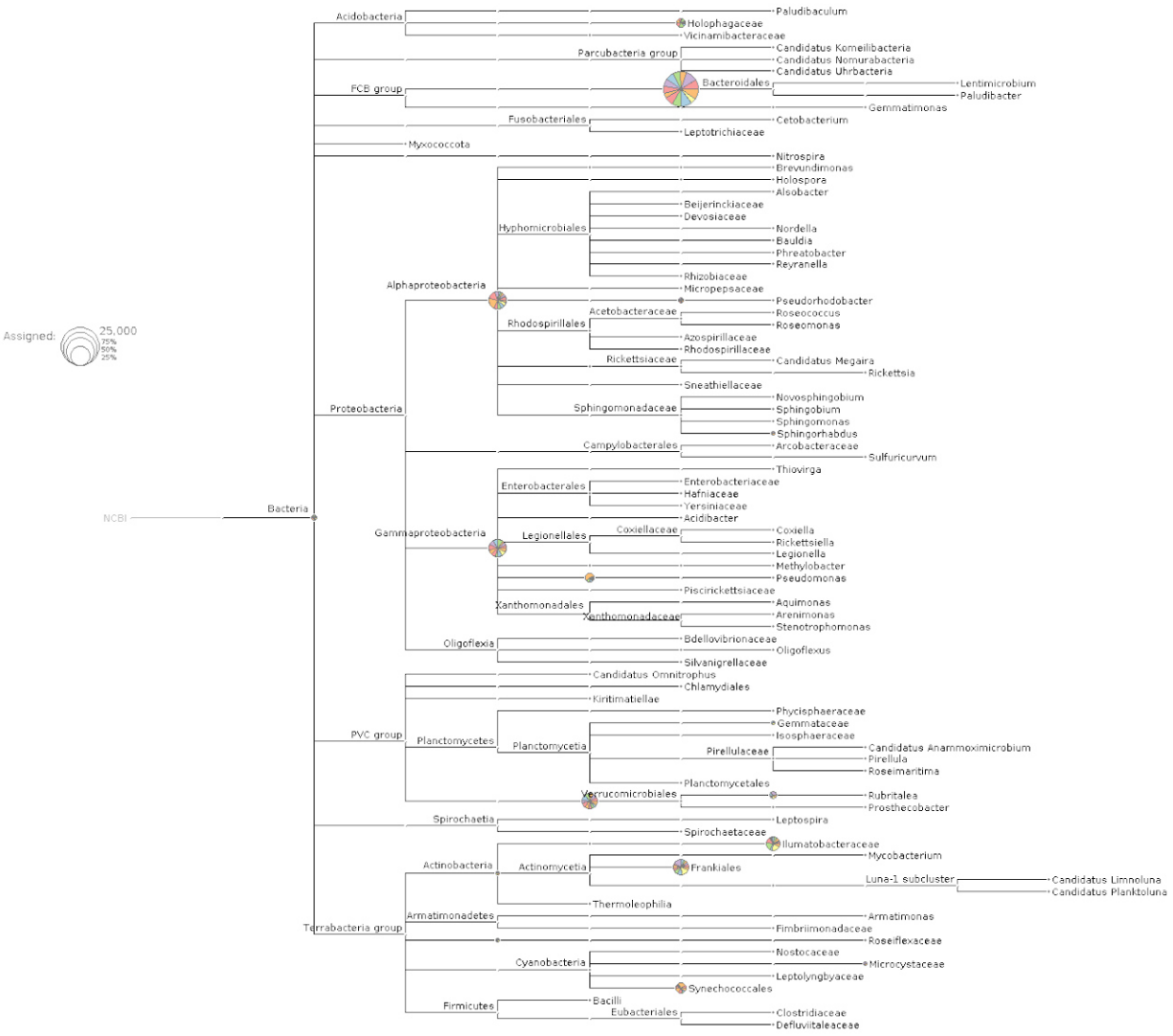


Figura S3. Filograma de la taxonomía de todas las muestras (MEGAN6)

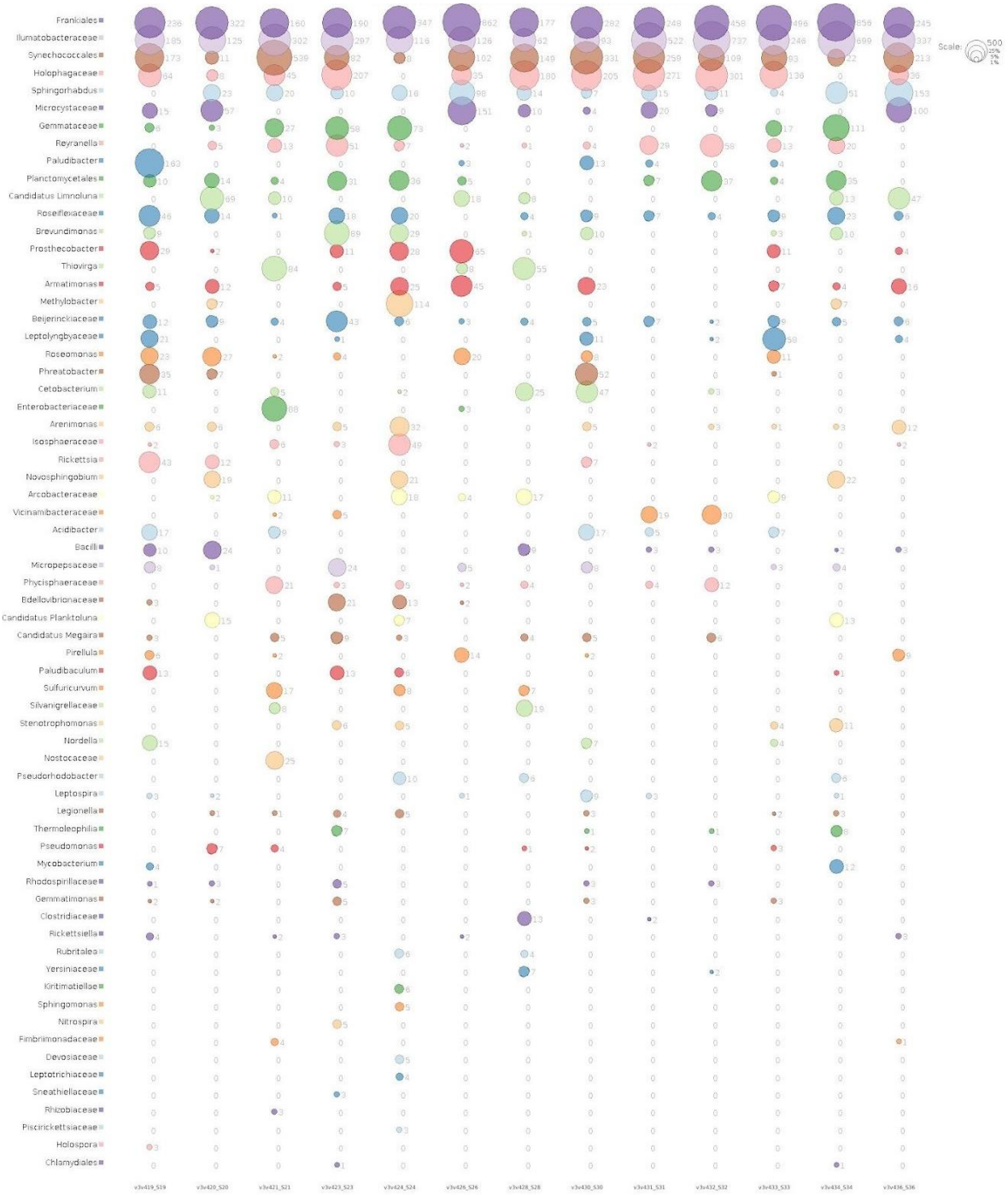


Figura S4. Gráfico de burbujas de los géneros bacterianos en todas las muestras (MEGAN6)

```
## Gráfico de barras

# Librerías
library(dplyr)
library(tidyr)
library(ggplot2)
library(readr)
library(RColorBrewer)

# 1. Importar datos
asv_table <- read_tsv("tabla_con_taxonomia.tsv")
metadata <- read_csv("metadata.csv")

# 2. Filtrar y ordenar
asv_table <- asv_table %>%
  separate(taxonomy,
           into = c("Domain", "Phylum", "Class", "Order", "Family", "Genus"),
           sep = ";",
           fill = "right")

asv_table <- asv_table %>%
  mutate(
    Domain = sub("d__", "", Domain),
    Phylum = sub("p__", "", Phylum),
    Class  = sub("c__", "", Class),
    Order  = sub("o__", "", Order),
    Family = sub("f__", "", Family),
    Genus  = sub("g__", "", Genus)
  )

asv_long <- asv_table %>%
  pivot_longer(cols = -c(Domain, Phylum, Class, Order, Family, Genus),
               names_to = "Sample",
               values_to = "Abundance") %>%
  left_join(metadata, by = c("Sample" = "sample_name")) %>%
  filter(Abundance > 0)


# 3. Función para los gráficos de barras
plot_barplot_taxon <- function(data,
                               nivel_taxonomico = "Phylum",
                               top_n = 6,
                               abundancia_relativa = TRUE,
                               ordenar_muestras = NULL,
                               paleta = NULL) {

  if(!(nivel_taxonomico %in% colnames(data))){
    stop("El nivel taxonómico no existe en el dataframe.")
  }

  df <- data %>%
```

```
  group_by(id2, .data[[nivel_taxonomico]]) %>%
  summarise(total = sum(Abundance), .groups = "drop")

colnames(df)[2] <- "Taxon"

if(abundancia_relativa){

  df <- df %>%
    group_by(id2) %>%
    mutate(total_sample = sum(total),
           Abundance = total / total_sample * 100) %>%
    ungroup()

  y_lab <- "Abundancia relativa (%)"

} else {

  df <- df %>%
    rename(Abundance = total)

  y_lab <- "Abundancia"
}

top_taxa <- df %>%
  group_by(Taxon) %>%
  summarise(total = sum(Abundance), .groups = "drop") %>%
  arrange(desc(total)) %>%
  slice(1:top_n) %>%
  pull(Taxon)

df <- df %>%
  mutate(Taxon = ifelse(Taxon %in% top_taxa, Taxon, "Otros")) %>%
  group_by(id2, Taxon) %>%
  summarise(Abundance = sum(Abundance), .groups = "drop")

if(!is.null(ordenar_muestras)){
  df$id2 <- factor(df$id2, levels = ordenar_muestras)
}

if(is.null(paleta)){
  paleta <- c(brewer.pal(9, "Paired"),
              brewer.pal(8, "Set2"))
}

ggplot(df, aes(x = id2, y = Abundance, fill = Taxon)) +
  geom_bar(stat = "identity") +
  theme_minimal() +
  labs(x = "", y = y_lab, fill = nivel_taxonomico) +
  scale_fill_manual(values = paleta) +
  theme(
    axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1),
```

```
      axis.text = element_text(size = 12),
      axis.title = element_text(size = 14),
      legend.position = "right",
      legend.text = element_text(size = 10),
      legend.title = element_text(size = 12, face = "bold"),
      panel.grid.major = element_line(color = "gray90"),
      panel.grid.minor = element_blank()
    )
}
```

Figura S5. Código en R para generar gráfico de barras

```r
## Diagrama de Venn

# Librerías
library(tidyverse)
library(conflicted)
library(ggplot2)
library(ggvenn)
library(readr)

# 1. Importar datos
tabla_abund <- read.delim("feature-table-with-taxonomy.tsv")
metadata <- read.delim("metadata.tsv")

colnames(tabla_abund)[1] <- "Taxon"

# 2. Convertir a formato largo y filtrar
tabla_long <- tabla_abund %>%
  pivot_longer(
    cols = -Taxon,
    names_to = "Sample",
    values_to = "Abundance"
  )

tabla_presencia <- tabla_long %>%
  filter(Abundance > 0)

# 3. Agregar metadata y agrupar taxones
tabla_presencia <- tabla_presencia %>%
  left_join(metadata, by = c("Sample" = "sample_name"))

lista_taxones <- tabla_presencia %>%
  group_by(Localidad) %>%
  summarise(Taxones = list(unique(Taxon))) %>%
  deframe()

# 4. Graficar Diagrama de Venn
ggvenn(
  lista_taxones,
  fill_color = c("#E69F00", "#56B4E9", "#009E73", "#F0E442", "#D55E00"),
  stroke_size = 0.5,
  set_name_size = 4
) +
  theme_bw() +
  labs(title = "Taxones compartidos entre localidades") +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold")
  )
```

Figura S6. Código en R para generar Diagrama de Venn

```r
## Gráfico de cajas y bigotes

# Librerías
library(ggplot2)
library(dplyr)
library(readr)

# 1. Importar datos
shannon_table <- read_tsv("ruta/alpha-diversity.tsv")
metadata <- read_csv("ruta/metadata.csv")

# 2. Unir datos
colnames(shannon_table)[1] <- "sample_name"

shannon_data <- shannon_table %>%
  inner_join(metadata, by = "sample_name")

colores_localidad <- c(
  "Localidad_1" = "#87CEFA",
  "Localidad_2" = "#4682B4",
  "Localidad_3" = "#FF6666",
  "Localidad_4" = "#66C266",
  "Localidad_5" = "#D299FF"
)

# 3. Gráfico de cajas y bigotes
ggplot(shannon_data,
       aes(x = locality,
           y = shannon_entropy,
           fill = locality)) +
  geom_boxplot() +
  scale_fill_manual(values = colores_localidad) +
  labs(title = "Distribución del Índice de Shannon",
       x = "Localidad",
       y = "Índice de Shannon") +
  theme_minimal() +
  theme(
    axis.text = element_text(size = 12),
    axis.title = element_text(size = 14),
    legend.position = "none"
  )
```

Figura S7. Código en R para generar gráfico de cajas y bigotes

```r
## Análisis de Coordenadas Principales (PCoA)

# Librerías
library(tidyverse)
library(conflicted)
library(ggplot2)
library(dplyr)
library(readr)

# 1. Importar datos
pcoa_data <- read_tsv("ordination.txt", skip = 6)
metadata <- read.csv("metadata.csv")

colnames(pcoa_data)[1:3] <- c("SampleID", "PC1", "PC2")
pcoa_data <- pcoa_data %>%
  filter(!grepl("Biplot|Site constraints", SampleID))

# 3. Unir datos
pcoa_data <- pcoa_data %>%
  left_join(metadata, by = c("SampleID" = "sample_name"))

# 4. Graficar PCoA

ggplot(pcoa_data, aes(x = PC1, y = PC2, color = locality, shape = season)) +
  geom_point(size = 4) +
  labs(
  title = "",
  x = "PC1 (%)",
  y = "PC2 (%)",
  color = "Localidad",
  shape = "Época"
  ) +
  scale_color_manual(values = c(
    "Localidad_1" = "#1f77b4",
    "Localidad_2" = "#E3893A",
    "Localidad_3" = "#259993",
    "Localidad_4" = "#BB3F35",
    "Localidad_5" = "#825DA3"
  )) +
  scale_shape_manual(values = c(16, 17)) +
  theme_minimal() +
  theme(
    legend.position = "right",
    legend.text = element_text(size = 10),
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    panel.grid.major = element_line(color = "gray95"),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(size = 0.8, fill = "transparent")
```

Figura S8. Código en R para generar Análisis de Coordenadas Principales

```r
## Mapa de calor (heatmap)

# Librerías
library(tidyverse)
library(phetmap)

# 1. Importar datos
faprotax_data <- read.csv("ruta/tabla_faprotax.csv",
                          header = TRUE,
                          row.names = 1)

metadata <- read.csv("ruta/metadata.csv",
                     header = TRUE)

# 2. Filtrar
if("OTU.ID" %in% colnames(faprotax_data)){
  faprotax_data <- faprotax_data %>% select(-OTU.ID)
}

faprotax_long <- faprotax_data %>%
  rownames_to_column(var = "Function") %>%
  pivot_longer(-Function,
               names_to = "Sample",
               values_to = "Abundance") %>%
  left_join(metadata, by = c("Sample" = "sample_name")) %>%
  filter(Abundance > 0)

# Generar matriz
faprotax_matrix <- faprotax_long %>%
  group_by(Function, id2) %>%
  summarise(Mean_Abundance = mean(Abundance), .groups = "drop") %>%
  pivot_wider(names_from = id2,
              values_from = Mean_Abundance,
              values_fill = 0) %>%
  column_to_rownames(var = "Function") %>%
  as.matrix()

# Graficar mapa de calor (heatmap)
pheatmap(faprotax_matrix,
         scale = "row",
         clustering_distance_rows = "euclidean",
         clustering_distance_cols = "euclidean",
         clustering_method = "ward.D2",
         color = colorRampPalette(c("blue", "white", "red"))(100),
         border_color = NA,
         cutree_cols = 3,
         cutree_rows = 3)
```

Figura S9. Código en R para generar mapa de calor