

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias Biológicas y Ambientales

**Análisis de la composición de consorcios microbianos de
Galápagos con capacidad de remediación de aguas residuales**

Stefany Mishell Aguilar Herrera

Ingeniería en Biotecnología

Trabajo de fin de carrera presentado como requisito
para la obtención del título de
Ingeniería en Biotecnología

Quito, 17 de diciembre de 2021

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias Biológicas y Ambientales

HOJA DE CALIFICACIÓN DE TRABAJO DE FIN DE CARRERA

**Análisis de la composición de consorcios microbianos de Galápagos con capacidad de
remediación de aguas residuales**

Stefany Mishell Aguilar Herrera

Nombre del profesor, Título académico

María de Lourdes Torres, PhD

Nombre del profesor, Título académico

Juan José Guadalupe, MSc

Quito, 17 de diciembre de 2021

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombres y apellidos: Stefany Mishell Aguilar Herrera

Código: 00200842

Cédula de identidad: 0922865936

Lugar y fecha: Quito, 17 de diciembre de 2021

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETHeses>.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETHeses>.

RESUMEN

Las islas Galápagos son mundialmente conocidas por su alta diversidad y endemismo. No obstante, existe un problema con el manejo de las aguas residuales. Estas son desechos líquidos que presentan características físicas, químicas y biológicas que amenazan al medio ambiente y a la salud humana. Para su tratamiento, es posible emplear microorganismos que reducen los niveles de contaminantes de forma sustentable. Dentro de este enfoque, se ha demostrado la eficiencia de utilizar consorcios alga-bacteria por sus relaciones mutualistas, pero la información sobre su composición es limitada. Por ende, este estudio busca caracterizar la composición de un consorcio con potencial de remediación de aguas residuales. Se contó con un tratamiento (fotoperiodo 12:12) y dark control (ausencia de luz) para evidenciar el rol de las algas en el consorcio. Mediante la secuenciación con MinION y el análisis metagenómico con el método de shotgun de este consorcio, fue posible determinar su composición, diversidad y funcionalidad en los días 0, 9 y 12. La asignación taxonómica demostró la dominancia del Filo *Proteobacteria* y la Clase *Alphaproteobacteria* a lo largo del tiempo, tanto en los tratamientos como dark controls. La abundancia de *Chlorophyta* disminuyó notablemente tras 12 días y fue ligeramente menor en los dark controls. La alfa diversidad en todos los puntos analizados fue baja. Por su parte, se observó que los tratamientos de los días 9 y 12 fueron las muestras más similares entre sí, seguidas de los tratamientos y dark controls del mismo día. El perfil funcional del consorcio reflejó la existencia de varias vías metabólicas útiles para la degradación de contaminantes frecuentes en aguas residuales, en mayor cantidad en el día 9. Por consiguiente, el consorcio de Galápagos es potencialmente aplicable en el tratamiento de aguas residuales, pero se requiere continuar con su estudio a mayor profundidad.

Palabras clave: Consorcio, Alga, Bacteria, Galápagos, Metagenómica, Aguas Residuales, Biorremediación, Composición Taxonómica, Diversidad Microbiana, MinION.

ABSTRACT

The Galapagos Islands are worldwide known for their high diversity and endemism. However, there is a problem with wastewater management. These are liquid wastes that have physical, chemical, and biological characteristics that threaten the environment and human health. For its treatment, it is possible to use microorganisms that reduce the levels of pollutants sustainably. Within this approach, the efficiency of using bacterial-algae consortia has been demonstrated because of their mutualistic relationships, but information on their composition is limited. Therefore, this study aims to characterize the composition of a potential wastewater remediation consortium. There was a treatment (12:12 photoperiod) and dark control (absence of light) to reveal the role of algae in the consortium. By MinION sequencing and shotgun metagenomic analysis of this consortium, it was feasible to determine its composition, diversity, and functionality on days 0, 9, and 12. Taxonomic assignment demonstrated the dominance of the *Proteobacteria* Phylum and *Alphaproteobacteria* Class over time, both in treatments and dark controls. *Chlorophyta*'s abundance decreased notably after 12 days, and it was slightly lower in the dark controls. The alpha diversity at all points analyzed was low. Treatments on days 9 and 12 were found to be the most similar samples to each other, followed by treatments and dark controls from the same day. The functional profile of the consortium reflected the existence of several metabolic pathways useful for wastewater frequent pollutants degradation, in greater quantity at day 9. Therefore, the Galapagos consortium is potentially applicable in wastewater treatment, but further study is required.

Key words: Consortium, algae, bacteria, Galapagos, Metagenomics, Wastewater, Bioremediation, Taxonomic Composition, Microbial Diversity, MinION.

TABLA DE CONTENIDO

Introducción	11
Métodos.....	15
Recolección de muestras.....	15
Extracción de ADN, cuantificación con NanoDrop y electroforesis.....	15
Preparación de librerías y secuenciación con MinION.....	16
Demultiplexing y análisis de la calidad de las secuencias.....	16
Alfa y beta diversidad.....	16
Asignación taxonómica.....	17
Ensamblaje y predicción de genes funcionales.....	17
Resultados.....	18
Demultiplexing y calidad de las lecturas.....	18
Asignación taxonómica.....	19
Alfa diversidad.....	19
Beta diversidad.....	20
Perfiles funcionales.....	20
Discusión.....	22
Conclusiones.....	26
Tablas.....	27
Figuras.....	30
Referencias bibliográficas.....	35

ÍNDICE DE TABLAS

Tabla 1.- Número de lecturas, bases producidas, calidad media, N50 y lectura de mayor tamaño para cada Time Point (TP).....	27
Tabla 2.- Cantidad de lecturas clasificadas taxonómicamente y su porcentaje para cada Time Point (TP).	27
Tabla 3.- Valores resultantes para la alfa diversidad de los diferentes Time Points (TP).	28
Tabla 4.- Vías metabólicas encontradas y el número de asignaciones a cada una.	29

ÍNDICE DE FIGURAS

Figura 1.- Diagrama de puntos de la calidad media del secuenciamiento con respecto a la longitud de las lecturas.....	30
Figura 2.- Comparación de la composición taxonómica a nivel de dominio entre las 5 muestras.....	31
Figura 3.- Comparación de la composición taxonómica a nivel de los filos más representativos de cada muestra.....	32
Figura 4.- Krona chart de la composición de Bacteria y Eukaryota en cada TP.	32
Figura 5.- Matriz de distancia basada en el índice de Bray-Curtis.....	33
Figura 6.- Comparación de las vías metabólicas con importancia en biorremediación encontradas y el número de asignaciones de cada vía.	34

ÍNDICE DE ANEXOS

Anexo 1: Abundancia relativa de las Clases predominantes dentro de las muestras.....	47
Anexo 2: Abundancia relativa de los Órdenes predominantes dentro de las muestras	48

INTRODUCCIÓN

Tanto las industrias como las instituciones públicas y privadas, residencias y establecimientos generan residuos sólidos y líquidos, siendo los últimos conocidos como aguas residuales, cuya descarga a cuerpos de agua constituye una de las mayores causas de contaminación ambiental. En el Ecuador, solamente entre el 10 y 25% de aguas residuales reciben un tratamiento previo a su descarga (Velasco, et al., 2018). Algunas características de las aguas residuales son que presentan una alta demanda química y biológica de oxígeno, alta cantidad de sólidos totales disueltos y suspendidos; además de contener organismos patógenos y varios contaminantes orgánicos e inorgánicos recalcitrantes, cuya alta concentración y poca biodegradabilidad representan el mayor desafío para la seguridad ambiental y la protección de la salud humana (Bharagava, et al., 2019). Por estos motivos, es fundamental buscar alternativas sustentables para combatir esta problemática; algunas consisten en el uso de métodos físicos, químicos, bioquímicos y biológicos (Peña, et al., 2018).

La biorremediación consiste en procesos colectivos que emplean sistemas biológicos para reducir los niveles de contaminantes hasta que sean indetectables, no tóxicos o aceptables bajo la legislación pertinente. Con frecuencia, se reporta el uso de comunidades microbianas en estos procesos por su capacidad de oxidar, inmovilizar, mineralizar o transformar contaminantes (Vishwakarma, et al., 2020). Su aplicación depende, principalmente, de la naturaleza del contaminante, pero también del grado de contaminación, tipo de ambiente, localización, costos y políticas ambientales. En base al sitio de aplicación, se clasifica a la biorremediación como *ex situ* o *in situ*. La primera hace referencia a la extracción y transporte del sitio contaminado hacia el sitio de tratamiento. Por su parte, la segunda implica la reducción de la concentración de contaminantes en el mismo sitio contaminado, sin perturbar su estructura (Azubuike, et al., 2016).

En cuanto a los tratamientos biológicos para aguas residuales, los consorcios microbianos de microalgas y bacterias han demostrado un inmenso potencial, puesto que existen relaciones de mutualismo, comensalismo, competencia y comunicación entre estos microorganismos (Subashchandrabose, et al., 2011). Es de interés la relación mutualista basada en el intercambio nutricional, donde las bacterias suministran nitrógeno fijo, vitamina B, fitohormonas, derivados de tiamina y sideróforos para acelerar el crecimiento de microalgas y su producción de biomasa (Zhang, et al., 2020). El oxígeno molecular de la fotosíntesis de las algas es utilizado como aceptor de electrones por bacterias para degradar materia orgánica y el dióxido de carbono producido por la mineralización de las bacterias completa el ciclo fotosintético (Subashchandrabose, et al., 2011). Por esto, se incrementa la eficiencia metabólica, se promueve el crecimiento acelerado de los grupos, se confiere robustez al ambiente, se minimiza la invasión de otras especies, se mejora la absorción total de nutrientes (Zhang, et al., 2020), se reduce el costo de aireación, se ahorra energía y disminuye el consumo del reactor (Mu, et al., 2021).

Además, el crecimiento de las algas agota el amonio, nitrato y fosfato, lo que beneficia la remoción de nutrientes en el tratamiento de aguas residuales, mejorando la calidad del proceso (Subashchandrabose, et al., 2011). Estudios han demostrado que géneros bacterianos se asocian específicamente con ciertas algas verdes dependiendo de sus productos secretores (Ramanan, et al., 2016). Por ejemplo, la bacteria *Azobacter* S8 junto a la microalga *Chlorella vulgaris* reducen la concentración de cromo al encontrarse en una proporción 1:1 (Tangahu, et al., 2020). A su vez, los géneros *Roseobacter* y *Flavobacterium* se relacionan consistentemente con diátomeas (Zhang, et al., 2020). En consecuencia, el estudio de las especies presentes en los consorcios y sus proporciones relativas es imperativo para determinar los grupos microbianos aptos para su empleo en el tratamiento de aguas residuales (Mu, et al., 2021).

Dado que la mayoría de los microorganismos ambientales no son cultivables, la secuenciación de su material genético permite obtener información sobre estos (Wiegand, et al., 2021).

La metagenómica es una tecnología empleada para el estudio del material genético obtenido directamente de las muestras ambientales, sin aislamiento ni cultivo de microorganismos (Datta, et al., 2020). El ADN obtenido de cada muestra ambiental se conoce como metagenoma y constituye la fuente de información genética (Datta, et al., 2020). Uno de los enfoques más utilizados es el “shotgun” que consiste en secuenciar todo el ADN presente en una muestra (Garrido, et al., 2017). Esta metodología permite realizar un análisis de genes funcionales, diversidad de genes relevantes, genes catabólicos y operones completos taxonómica y filogenéticamente (Bharagava, et al., 2019). De esta manera, la exploración de comunidades microbianas y sus funciones es más confiable e informativa (Sidhu, et al., 2017).

MinION de Oxford Nanopore Technologies (ONT) es un dispositivo de bolsillo que aplica la tecnología de secuenciación por nanoporos para el análisis de ácidos nucleicos de manera rápida y en tiempo real (Tyler, et al., 2018). MinION identifica las bases del ADN al medir los cambios en la conductividad eléctrica generados a medida que una hebra de ADN pasa por un poro biológico (Lu, et al., 2016). Previo a la secuenciación, se unen adaptadores a la secuencia de ADN para facilitar la captura y entrada del extremo 5' de una hebra en la enzima que asegura el desplazamiento unidireccional del ADN a través del nanoporo en milisegundos (Jain, et al., 2016).

La información generada gracias a esta tecnología permite determinar la composición y los perfiles funcionales de la comunidad durante el tratamiento de aguas residuales. Por ejemplo, Sidhu, et al. en el 2017 encontraron que en agua residual no tratada el metabolismo de ADN y proteínas, quimiotaxis, respiración y degradación de compuestos aromáticos fueron abundantes. Por otro lado, en el 2013, Silva y colaboradores reportaron la presencia vías del metabolismo de nitrógeno, azufre y fósforo, además de compuestos aromáticos.

Las Islas Galápagos son islas volcánicas ubicadas a 930 km de la costa ecuatoriana mundialmente reconocidas por su gran biodiversidad y endemismo (Jones, et al., 2021). A su vez, las islas se encuentran en una ubicación oceanográfica que crea diversos hábitats marinos con gradientes de temperatura y condiciones que cambian las estructuras de las comunidades microbianas (Gifford, et al., 2020). A pesar de que el archipiélago es uno de los lugares más estudiados del planeta, muchas de las especies que lo habitan y su ecología todavía son desconocidas (Walsh & Mena, 2013). Por su parte, el tratamiento de aguas residuales en el archipiélago es casi nulo, representando únicamente al 0.48% de PTAR en el país, siendo la región con menor tratamiento. Asimismo, la poca agua tratada es descargada en el mar y pantanos artificiales; mientras que el 100% del agua residual sin tratamiento se deposita en el mar (INEC, 2016). Adicionalmente, los análisis del suministro de agua en Galápagos demuestran la presencia elevada de coliformes fecales, detergentes, nitratos, materia orgánica y un nivel insuficiente de cloro desinfectante (Parque Nacional Galápagos, 2017; Alarcón, 2019). Considerando la falta de tratamiento de aguas contaminadas en las islas, es crucial caracterizar a los microorganismos autóctonos de la zona para conocer su potencial metabólico, metabolitos y procesos de degradación (Chandran, et al., 2020), debido a que estos están adaptados a dichas condiciones ambientales (Maitra, 2018).

Con estos antecedentes, y tomando en cuenta que la biorremediación es una herramienta emergente con un enfoque sustentable para el manejo de la contaminación ambiental, existe la necesidad de generar más investigaciones en esta área, especialmente, enfocada en los recursos biológicos presentes en una zona con tanta riqueza como las islas Galápagos. Por consiguiente, el objetivo del presente estudio es caracterizar la diversidad genética dentro de un consorcio alga-bacteria de Galápagos a lo largo del tiempo mediante su análisis metagenómico, para comprender la composición taxonómica, la dinámica microbiana y las posibles vías metabólicas implicadas en la remoción de compuestos contaminantes en aguas residuales.

MÉTODOS

Recolección de muestras

La muestra del consorcio se obtuvo de un arroyo de agua dulce en la Encañada Gutiérrez, San Cristóbal, cuya elevación es de 179 m.s.n.m y coordenadas geográficas 0°55'42.57" S 89°30'41.08" O (Suárez, 2019). Se criopreservó la muestra hasta su inoculación en agua residual sintética, la cual permite obtener mediciones reproducibles, pues se formula a partir de una concentración fija de carbono y nutrientes (O'Flaherty & Gray, 2013). Esto se hizo para evaluar la capacidad de remoción de la muestra a lo largo del tiempo en el Laboratorio de Ingeniería Ambiental USFQ. El consorcio correspondiente al tratamiento se mantuvo durante 15 días a 24°C, con agitación constante y sometido a un fotoperiodo 12:12. Mientras el dark control (DC) se mantuvo bajo las mismas condiciones, pero no se sometió a luz para evaluar la remoción de contaminantes al restringir el crecimiento de las algas en el consorcio.

Extracción de ADN, cuantificación con NanoDrop y electroforesis

Para la extracción de ADN, se transfirieron 15 ml del agua residual sintética con el inóculo a un tubo Falcon. Esto se realizó en los Time Points (TPs) 0, 9 y 12 tanto para los tratamientos como dark controls, ya que en estos puntos se evidenció la mayor remoción de contaminantes. Primero, se centrifugó la muestra por 8 minutos a 9000 rpm, para concentrar la biomasa celular; posteriormente, se descartó el sobrenadante. Se colocó la biomasa en un mortero, se añadió nitrógeno líquido para el paso de lisis celular, se maceró y se pesó la biomasa resultante. Luego, se extrajo el ADN con el kit DNeasy® PowerSoil® Pro (Qiagen, Hilden, Alemania, 2020), en base a las indicaciones del fabricante; con ciertas modificaciones: se realizó un vórtex por 40 minutos en la etapa de lisis celular y se utilizaron 30 µl de agua de PCR como buffer de elución. Se colocó 1 µl de la muestra en el espectrofotómetro NanoDrop para cuantificar la cantidad y calidad del ADN extraído. Finalmente, se cargaron 2 µl de ADN en un gel de agarosa al 1.5% para visualizar que el ADN no esté fragmentado.

Preparación de librerías y secuenciación con MinION

Para iniciar, se realizó un lavado del ADN con beads magnéticas AMPure XP en una proporción de 1.8X, con el fin de mantener fragmentos de un amplio rango de tamaños. Se cuantificó el ADN extraído con Qubit. Luego, se llevó a cabo la preparación de la librería genómica siguiendo el protocolo “Native barcoding genomic DNA”, generado por ONT (2021), para el cual se siguieron todas las directrices, excepto la repetición del lavado con etanol al 70% para limpiar las beads magnéticas. Los adaptadores y códigos de barras se añadieron de acuerdo con el protocolo y se asignaron aleatoriamente para cada muestra. La secuenciación se realizó cargando los buffers especificados en el protocolo y la librería de ADN en una celda de flujo por aproximadamente 24 horas. Cabe mencionar que la secuenciación de cada TP se realizó en un día distinto.

Demultiplexing y análisis de la calidad de las secuencias

La remoción de adaptadores y la clasificación de las lecturas en base a su código de barras se realizaron en el programa Porechop (Github, 2021), el cual alinea subconjuntos de lecturas con conjuntos de adaptadores y analiza si existe al menos el 90% de coincidencia para considerar que el adaptador se encuentra en la muestra (Wick, et al., 2017; Github, 2021). Para el análisis de la calidad de las secuencias se empleó el programa Nanoplot (De Coster, et al., 2019) que produce un resumen estadístico y gráficos QC informativos, pues contienen varios aspectos sobre la longitud de las lecturas, rendimiento acumulado, puntajes de calidad, calidad a lo largo del tiempo, entre otros (De Coster, et al., 2019).

Alfa y beta diversidad

La alfa diversidad indica la riqueza de especies en una comunidad individual (Arellano & Halffer, 2003). Para su cálculo se empleó el paquete Vegan (Oksanen, et al., 2020) en R, el mismo que provee herramientas para realizar estudios de ecología descriptiva de comunidades, mediante análisis de diversidad y disimilitudes (Oksanen, et al., 2020). El

paquete cuenta con los índices de Shannon y Simpson que fueron empleados; adicionalmente, se calculó el índice de Pielou.

La beta diversidad corresponde al grado de diferencia entre comunidades distintas (Arellano & Halffer, 2003). Su análisis se realizó en el programa Simka (Benoit, et al., 2016) que compara metagenomas *de novo*, al calcular distancias ecológicas estándar de una colección de datos y reemplazando los recuentos de especies por recuentos de *k*-mer lo suficientemente largos para que sean específicos de un genoma (Benoit, et al., 2016). Dentro de Simka, se generaron valores para la disimilitud de Bray-Curtis a partir de la diferencia relativa ponderada, donde la contribución de cada especie al índice se pondera por la suma de las abundancias de las especies en las dos unidades de muestra que se comparan (Roberts, 2017).

Asignación taxonómica

Para la asignación taxonómica se utilizó el programa Kaiju (Menzel, et al., 2016) que clasifica las lecturas taxonómicamente basado en proteínas, es decir, traduce las secuencias de aminoácidos y las compara con el catálogo disponible de genes microbianos (Menzel, et al., 2016).

Ensamblaje y predicción de genes funcionales

Se realizó el ensamblaje de los genomas en Flye (Kolmogorov, et al., 2019), mismo que ensambla genomas *de novo* a partir de lecturas largas. Este genera disjointigs que representan concatenaciones de segmentos genómicos disjuntos y construye un gráfico de ensamblaje preciso basado en el concatenado resultante. Luego, realiza gráficos de repeticiones para obtener contigs precisos (Kolmogorov, et al., 2019). Posteriormente, los archivos resultantes de Flye se utilizaron para correr el pipeline de análisis metagenómico automático SqueezeMeta. Este permite ensamblar conjuntamente los metagenomas, recuperarlos mediante binning y realizar anotaciones funcionales (Tamames & Puentes, 2017).

RESULTADOS

A continuación, se presentan los resultados obtenidos del secuenciamiento y análisis bioinformático de tres Time Points analizados del consorcio alga-bacteria de Galápagos.

Demultiplexing y calidad de las lecturas

A partir del demultiplexing y la remoción de adaptadores con Porechop, se obtuvo que el total de lecturas asignadas al código de barras correspondiente al TP0 fueron 1,013,704 y las bases fueron 1,581,333,319. Por otro lado, el tratamiento del TP9 contó con 662,584 lecturas y 1,004,618,802 bases; mientras que el dark control de este punto obtuvo un total de 902,632 lecturas con 1,477,805,585 bases. Por último, el número de lecturas en el tratamiento del TP12 fue de 283,792 con 285,561,800 bases y en el dark control de este punto se obtuvieron 333,553 lecturas y 276,530,740 bases (Tabla 1).

Los resultados de la calidad del secuenciamiento de los diferentes TPs se encuentran en la Figura 1, la cual contiene un diagrama de puntos que relaciona la longitud de las lecturas con la calidad de las mismas. En cuanto al análisis del TP0, se obtuvo que la media de la longitud de las lecturas fue de 1,560.0, la calidad media fue 11.1 y el N50 fue 5,040.0; adicionalmente, la lectura de mayor longitud tuvo un tamaño de 52652 y una calidad de 9.1. En el tratamiento del TP9 la media de la longitud fue de 1,516.2, la calidad media fue 11.1 y el N50 resultó ser 3,613.0; la lectura más larga fue de 44710 con una calidad de 11.1. Por su parte, en el dark control del TP9 se registró una media de la longitud de las lecturas de 1,637.2, una calidad media de 11.0 y un N50 de 3,944.0; mientras que la lectura más larga de esta muestra tuvo un tamaño de 41111 y una calidad de 11.2. En el día 12, la longitud media fue de 1,006.2, la calidad promedio de 10.6, el N50 de 2,081.0 para el tratamiento, donde la lectura más larga tuvo un tamaño de 44946 con una calidad de 11.7, Por su parte, el dark control presentó una longitud media de 829.0, una calidad de 10.6 y un N50 de 1,291.0; siendo su lectura más larga de 44391 con 11.7 de calidad.

Asignación taxonómica

Posterior a la asignación taxonómica realizada en el programa Kaiju, se evidenció que la muestra del tratamiento del TP9 alcanzó el mayor porcentaje de clasificación (47.962%) dentro del grupo de muestras. En la Tabla 2 se observan los porcentajes de lecturas clasificadas en cada TP, destacando el hecho de que, en todas las muestras, el porcentaje de lecturas no clasificadas fue mayor que el de lecturas clasificadas. Por su parte, los dominios más abundantes en los TPs fueron Bacteria y Eukarya, los mismos que empezaron con proporciones muy similares (cerca al 50%) en el TP0, como se visualiza en la Figura 4. No obstante, a lo largo del tiempo, las bacterias colonizaron el consorcio y se volvieron dominantes tras 12 días, lo que se observa tanto para los tratamientos como para los dark controls.

En cuanto a los filos más abundantes encontrados, destacan principalmente *Proteobacteria* y *Chlorophyta*, con porcentajes de 41% en el TP0, entre 78% y 75% del filo bacteriano y 8% de eucariotas en el TP9; y entre 69% y 78% de *Proteobacteria* y 9% de *Chlorophyta* en el TP12. Otros filos presentes, pero en menor concentración, fueron: *Terrabacteria*, *Opisthokonta*, Grupo FCB y Grupo PVC. A nivel de clase, las más abundantes fueron: *Trebouxiophyceae*, *Alphaproteobacteria*, *Betaproteobacteria*, *Chlorophyceae*, *Planctomycetes*, *Actinobacteria* y *Firmicutes*, en este orden para el TP0. En el TP9, la clase de *Alphaproteobacteria* dominó el consorcio, seguida por *Trebouxiophyceae*, *Betaproteobacteria* y *Dikarya*. En el día 12, el porcentaje de *Proteobacteria* disminuye y se observó que el Grupo FCB presentó un aumento considerable (Anexo A).

Alfa diversidad

Para el análisis de la alfa diversidad en las diferentes muestras (Tabla 3), se emplearon los índices de Shannon, Simpson y Pielou. Inicialmente, los valores obtenidos para el índice de Shannon para el TP0, tratamiento TP9, dark control TP9, tratamiento TP12 y dark control TP12 fueron 1.162, 0.728, 0.832, 0.775 y 1.033 respectivamente. Por otro lado, los valores

resultantes para el índice de Simpson fueron de 0.618, 0.328, 0.386, 0.344 y 0.478. Ambos índices coinciden en demostrar que en el TP0 existía una mayor diversidad. En cuanto al índice de Pielou, los datos obtenidos corresponden a 0.597, 0.374, 0.428, 0.398 y 0.531, lo que implica que existe dominancia de algunos grupos taxonómicos en la composición del consorcio en todos los puntos muestreados.

Beta diversidad

Con respecto a este parámetro, se empleó el índice de Bray-Curtis generado en el programa Simka, que provee una matriz de distancia. Se observaron valores entre 0.723 y 0.986 al comparar la composición y abundancia de las muestras en grupos de 2. Además, como se indica en la Figura 5, se identificó que las muestras correspondientes al tratamiento TP9 y tratamiento TP12 son las menos distantes entre sí, indicando que la composición de las dos es más similar. Por otro lado, las muestras del tratamiento y dark control del mismo TP (9 o 12) resultaron un poco menos similares y ambas muestras del TP9 fueron más cercanas al TP0 que el resto de muestras.

Perfiles funcionales

En lo que respecta a los perfiles funcionales y vías metabólicas encontradas, se asociaron 112 vías a *Alphaproteobacteria*, 114 al género *Sphingomonas* sp. y se encontraron 321 que no lograron clasificarse en el TP0. En el tratamiento TP9, se asignaron 116 vías metabólicas a *Proteobacteria* y 116 sin clasificación. Mientras que en el dark control TP9, se identificaron 94 vías metabólicas en el orden de *Hyphomicrobiales*, 100 en *Alphaproteobacteria*, 103 en *Sphingopyxis* sp. y 438 sin asignación. Por su parte, en el tratamiento TP12 se registraron 115 vías sin asignación, 78 asignadas a *Alphaproteobacteria* y 96 a *Sphingomonas* sp. Dentro de estos perfiles funcionales, son de interés aquellos involucrados en biorremediación, especialmente de aguas residuales. En la Tabla 4, se indican las vías metabólicas importantes para la biorremediación y la cantidad de veces que estas

fueron encontradas tanto en el TP0, como en el tratamiento y dark control del TP9 y TP12. Cabe destacar que, en la mayoría de las rutas, el número de asignaciones es mayor en el dark control TP9. Además, se encontró la presencia de la ruta del metabolismo de xenobióticos únicamente en esta muestra. Por otro lado, la vía de degradación de estireno se encontró en todas las muestras excepto el TP0. Finalmente, en el TP12 se presentó la ruta del metabolismo de butanoato y solamente en el dark control TP12 se detectó la vía de degradación de bifenilo, como se representa en la Figura 6.

DISCUSIÓN

A partir de los resultados del tamaño de las lecturas, se evidenció que estos se encuentran dentro del rango esperado para el secuenciamiento por nanoporos (500 pb - 2.3 Mb), siendo las más comunes entre 10 kb y 30 kb (Amarasinghe, et al., 2020). A su vez, las lecturas de mayor longitud presentaron tamaños cercanos a 50 kb en todos los TPs, demostrando que sí se generaron lecturas largas. El valor del N50 es la longitud de una secuencia para la cual todas las secuencias de esa longitud o mayores suman el 50% del tamaño total del conjunto (Shafin, et al., 2020). Los resultados son congruentes con la literatura donde el N50 de muestras microbianas producidas por ONT oscila entre 2102 pb y 4369 pb (Cuscó, et al., 2021). La calidad de las lecturas depende de alcanzar una velocidad de translocación óptima del ácido nucleico a través del nanoporo (Amarasinghe, et al., 2020). Para este análisis se emplea el Phred Quality Score (Q entre 4 y 60) que define la probabilidad de que el base call sea incorrecto (Lapidus, 2009; Delahaye & Nicolas, 2021). La calidad en todas las muestras analizadas es alta, pues incluso supera el mínimo establecido (7) y los valores reportados por Yadav y colaboradores (8.9) (2021).

La asignación taxonómica demostró un alto porcentaje de lecturas no clasificadas; para el cual se esperaba alrededor del 50% (Corvelo, et al., 2018). Esto se debe a que las lecturas son comparadas con una base de datos de referencia (Tovo, et al., 2020). Actualmente, aún no se cuenta con suficiente información sobre la mayoría de microorganismos ambientales, los que se conocen como “materia oscura microbiana” (Wiegand, et al., 2021). Por ende, la disponibilidad de genomas de referencia tiene un gran impacto en la clasificación de las especies encontradas en muestras metagenómicas (Menzel, et al., 2016). A pesar de esto, Kaiju demostró que, inicialmente, los consorcios contaban con altos porcentajes de *Proteobacterias* y *Chlorophyta*. A lo largo del tiempo, el filo bacteriano se vuelve dominante en el consorcio, mientras que las algas se mantienen en el segundo lugar, pero con porcentajes mucho menores.

La literatura indica que el filo predominante en aguas residuales es *Proteobacteria* y que su porcentaje incrementa a medida que avanza el tratamiento (Zhang, et al., 2019), con abundancias relativas entre 28% y 67% (Plaza, et al., 2021). Otros filos frecuentes son *Bacteroidetes*, *Actinobacteria*, *Firmicutes*, *Chloroflexi*, *Acidobacteria* (Zhang, et al., 2019; Plaza, et al., 2021; Yu, et al., 2021; Kaevska, et al., 2015). En el TP12, el Grupo FCB (*Fibrobacteres*, *Chlorobi* y *Bacteroidetes*) igualó al porcentaje de algas; puesto que los *Bacteroidetes* son anaerobios estrictos que descomponen compuestos orgánicos complejos como los encontrados en aguas residuales, por lo que son de los grupos más frecuentes (2.4-15.7%) en PTARs (Niestęski, et al., 2020). Adicionalmente, Kaevska y colaboradores reportaron que su abundancia relativa aumenta a lo largo el tratamiento (2021), al igual que en el presente estudio. Para las eucariotas, Barreiro y colaboradores (2021) también informan la presencia del filo *Chlorophyta* y la clase *Trebouxiophyceae* en comunidades alga-bacteria empleadas en el tratamiento de aguas residuales. En los dark controls, se observó que la abundancia de algas fue ligeramente menor a la de sus respectivos tratamientos. Esto se debe a que estas pueden presentar un crecimiento heterótrofo en ambientes oscuros (Morales, et al., 2015), por lo cual podrían sobrevivir bajo las condiciones del dark control; no obstante, su eficiencia en la remoción de contaminantes es menor a la observada cuando sí reciben luz.

A nivel de clase, se observó que las *Alphaproteobacterias* constituyeron más de la mitad de los grupos presentes, lo cual es consistente con lo encontrado por Ye y Zhang (2013), Ramanan y colaboradores (2015) y Liu y colaboradores (2018). Dado que estas utilizan compuestos orgánicos de baja masa molecular disueltos en aguas residuales como fuente de energía, además de contar con altas tasas de absorción de sustratos (Kragelund, et al., 2006), desempeña un rol fundamental en la fijación de nitrógeno (Tamminen, et al., 2021). Adicionalmente, son más afines a asociarse con algas verdes que otros grupos, ya que son funcionalmente más aptas (Ramanan, et al., 2015).

La alfa diversidad corresponde al análisis de la estructura ecológica de una comunidad en base a su riqueza (número de grupos taxonómicos) y/o homogeneidad (distribución de la abundancia de los grupos) (Willis, 2019). Se evaluó el índice de Shannon, que refleja la entropía y el grado de incertidumbre (entre 0 y 5) de seleccionar al azar un individuo en la comunidad, donde valores mayores indican mayor diversidad (Pla, 2006). El índice de Simpson mide la probabilidad de que dos individuos seleccionados al azar correspondan a especies diferentes, donde 1 implica una diversidad infinita y 0 ausencia de diversidad (Roswell, et al., 2021). Ante esto, se observaron valores menores a los reportados para microbiomas de aguas residuales, por ejemplo, para Shannon entre 1.83 y 5 en muestras de Shanxi, China (Ma, et al., 2015), y 5 en muestras de Sao Paulo, Brasil (Nascimento, et al., 2018); y Simpson en muestras de Brasil (0.99) (Nascimento, et al., 2018) y Asia (0.74-0.99) (Gerhard & Gunsch, 2019). La disminución de la alfa diversidad implica que existe un menor número de especies y con abundancias distintas (Cyzdik & Zielinska, 2016). En PTAR, esto resulta beneficioso pues la dominancia de ciertos grupos incrementa la eficiencia del tratamiento (Plaza, et al., 2021).

Adicionalmente, el índice de Pielou es un indicador de homogeneidad o uniformidad, dependiente de la riqueza de especies y varía entre 0 (una especie dominante) y 1 (las especies son igual de abundantes) (Daly, et al., 2018). Inicialmente, en el TP0 se presenta una mayor uniformidad, pero con el curso del tiempo las *Proteobacterias* se vuelven dominantes dentro del consorcio. Sin embargo, el valor registrado también es menor al obtenido en estudios similares (0.71) (Zerva, et al., 2021). No obstante, todos los índices coinciden en que el TP0 demuestra mayor diversidad, debido a que este correspondía al inóculo inicial, antes de exponerlo a un medio de agua residual sintética; ya que, al inocular al consorcio en agua residual, la composición de este cambia por la complejidad y recalcitrancia de los contaminantes (Ma, et al., 2015). Otras posibles razones son el tratamiento empleado (Nascimento, et al., 2018) y la zona de muestreo (Tiwari, et al., 2021).

Por su parte, la beta diversidad es la variación en la composición de dos muestras y mide la ganancia o pérdida de especies entre ambas (Cyzdik & Zielinska, 2016). En base a la disimilitud de Bray-Curtis, los tratamientos TP9 y TP12 son más similares entre sí. Esto tiene sentido debido a que fueron sometidos a las mismas condiciones experimentales. Al comparar los tratamientos con los dark controls, se registran menos semejanzas. Por su parte, las muestras del TP9 son más similares al consorcio inicial (TP0), lo cual es consistente con la literatura donde se reporta que existen más diferencias entre la muestra inicial y las muestras después del ensayo (Plaza, et al., 2021).

Por último, los perfiles funcionales de interés para biorremediación fueron asignados a *Alphaproteobacterias*, en concordancia con lo reportado por Joshi y colaboradores (2017), donde las señales más intensas de degradación de contaminantes orgánicos se encontraron en este grupo, al igual que genes de degradación funcionales. Dependiendo de la fuente de descarga, el agua residual varía en su composición, pero algunos de los compuestos orgánicos más frecuentemente detectados son: bifenilos policlorados, pesticidas, herbicidas, fenoles, hidrocarburos aromáticos policíclicos, compuestos alifáticos y heterocíclicos (Zheng, et al., 2013; Roy, et al., 2018; Wu, et al., 2021; Studnik, et al., 2015; Shi, et al., 2014; Aliabadi, et al., 2012). En todos los puntos, se contó con asignaciones a vías metabólicas de degradación de estos compuestos; principalmente, de compuestos alifáticos y fenoles, los cuales se encuentran en combustibles y productos de uso diario (Loughrin, et al., 2009). Por ende, es muy probable su presencia en aguas de Galápagos. En el TP9, se registraron los valores más altos, sugiriendo que este día es crucial en el proceso de remoción de contaminantes; lo cual es consistente con los resultados de Sidhu y colaboradores (2017). No obstante, se detectó una alta asignación a vías del metabolismo de nitrógeno y azufre en todos los puntos, pues estas son necesarias para la supervivencia de los microorganismos en este ambiente (Joshi, et al., 2017).

CONCLUSIONES

Tras el análisis de los resultados en este estudio preliminar sobre la composición taxonómica, diversidad microbiana y perfiles funcionales del consorcio alga-bacteria de Galápagos, se concluye que el presente consorcio es prometedor para su aplicación en el tratamiento de agua residual. Esto se debe a que la composición del consorcio es similar a la reportada en la literatura sobre plantas de tratamiento de agua residual con resultados efectivos. Además, se evidenció la elevada presencia del filo *Proteobacteria* en la muestra inicial del consorcio y a lo largo del tiempo. Este grupo bacteriano es de gran interés en la remoción de contaminantes de fuentes de agua, ya que cuenta con la capacidad de adaptarse, sobrevivir, proliferar y reducir el nivel de contaminantes en agua residual. A pesar de que el consorcio no presentó una alta diversidad, esto no representa una limitación para su aplicación, pues se reporta que en plantas de tratamiento de agua residual es beneficioso que exista dominancia de ciertos grupos como las *Proteobacterias*. Adicionalmente, se observó que los microorganismos del consorcio cuentan con vías metabólicas útiles para la degradación de contaminantes frecuentes en aguas residuales y potencialmente presentes en el agua de Galápagos. Por consiguiente, se recomienda continuar con el análisis del consorcio a mayor profundidad para determinar su eficiencia en la remoción de contaminantes de aguas residuales.

TABLAS

Tabla 1.- Número de lecturas, bases producidas, calidad media, N50 y lectura de mayor tamaño para cada Time Point (TP).

	TP0	TP9		TP12	
	Tratamiento	Tratamiento	Dark Control	Tratamiento	Dark Control
Número de lecturas	1,013,704.00	662,584.00	902,632.00	283,792.00	333,553.00
Bases totales	1,581,333,319.00	1,004,618,802.00	1,477,805,585.00	285,561,800.00	276,530,740.00
Calidad media	11.0	11.2	11.1	10.6	10.6
N50	5,040.0	3,613.0	3,944.0	2,081.0	1,291.0
Lectura de mayor tamaño (pb)	52652	44710	41111	44946	44381

Cantidad de lecturas reconocidas por el barcode asignado a cada muestra y el total de bases obtenido en cada muestra. Además de presentar otros parámetros de interés para analizar la calidad de las lecturas como la calidad media, N50 y lectura de mayor tamaño.

Tabla 2.- Cantidad de lecturas clasificadas taxonómicamente y su porcentaje para cada Time Point (TP).

	TP0	TP9		TP12	
	Tratamiento	Tratamiento	Dark Control	Tratamiento	Dark Control
Lecturas clasificadas	220479	151279	331261	40936	84431
Lecturas totales	1013704	315415	902632	111155	333553
Porcentaje clasificado	21.750%	47.962%	36.699%	36.828%	25.313%
Porcentaje no clasificado	78%	52%	63%	63%	75%

Cantidad y porcentaje de lecturas clasificadas y no clasificadas por Kaiju para cada uno de los TPs.

Tabla 3.- Valores resultantes para la alfa diversidad de los diferentes Time Points (TP).

	TP0	TP9_TRAT	TP9_DC	TP12_TRAT	TP12_DC
Shannon	1.1620742	0.727596	0.8322974	0.7746532	1.032605
Simpson	0.6183251	0.3288302	0.3864949	0.3439643	0.477578
Pielou	0.597	0.374	0.428	0.398	0.531

Cálculo de alfa diversidad en base a los índices de Shannon, Simpson y Pielou para las muestras analizadas.

Tabla 4.- Vías metabólicas encontradas y el número de asignaciones a cada una.

Vía encontrada	TP0	TP9		TP12	
	Tratamiento	Tratamiento	Dark Control	Tratamiento	Dark Control
Degradación del 1,2-dicloroetano	11	8	16	NP	NP
Degradación del 1,4-diclorobenceno	17	NP	NP	NP	NP
Degradación de 1- y 2-metilnaftaleno	17	15	19	NP	NP
Degradación del 2,4-diclorobenzoato	20	17	17	16	NP
Degradación del ácido 3-cloroacrílico	8	6	16	6	6
Degradación de benzoato por ligación de CoA	95	78	108	72	79
Degradación de benzoato por hidroxilación	70	55	75	49	56
Degradación del trinitrotolueno	27	22	22	22	22
Degradación del estireno	NP	24	23	23	23
Degradación de gamma-hexaclorociclohexano	38	28	45	27	25
Degradación de la caprolactama	27	19	39	20	17
Degradación de carbazol	12	12	13	11	12
Degradación de bifenilo	NP	NP	NP	NP	17
Degradación de estructuras de glicanos	35	29	34	31	NP
Metabolismo de butanoato	NP	NP	NP	80	85
Metabolismo de los aminofosfonatos	15	NP	15	NP	NP
Metabolismo de xenobióticos por citocromo P450	NP	NP	23	NP	NP
Metabolismo del metano	75	57	93	54	54
Metabolismo del nitrógeno	181	127	201	139	121
Metabolismo del propanoato	125	78	127	70	75
Metabolismo del azufre	55	38	69	38	35

Vías metabólicas de interés para la biorremediación de aguas residuales y la de asignaciones a cada una para las 5 diferentes muestras. *NP: No

Presenta.

FIGURAS

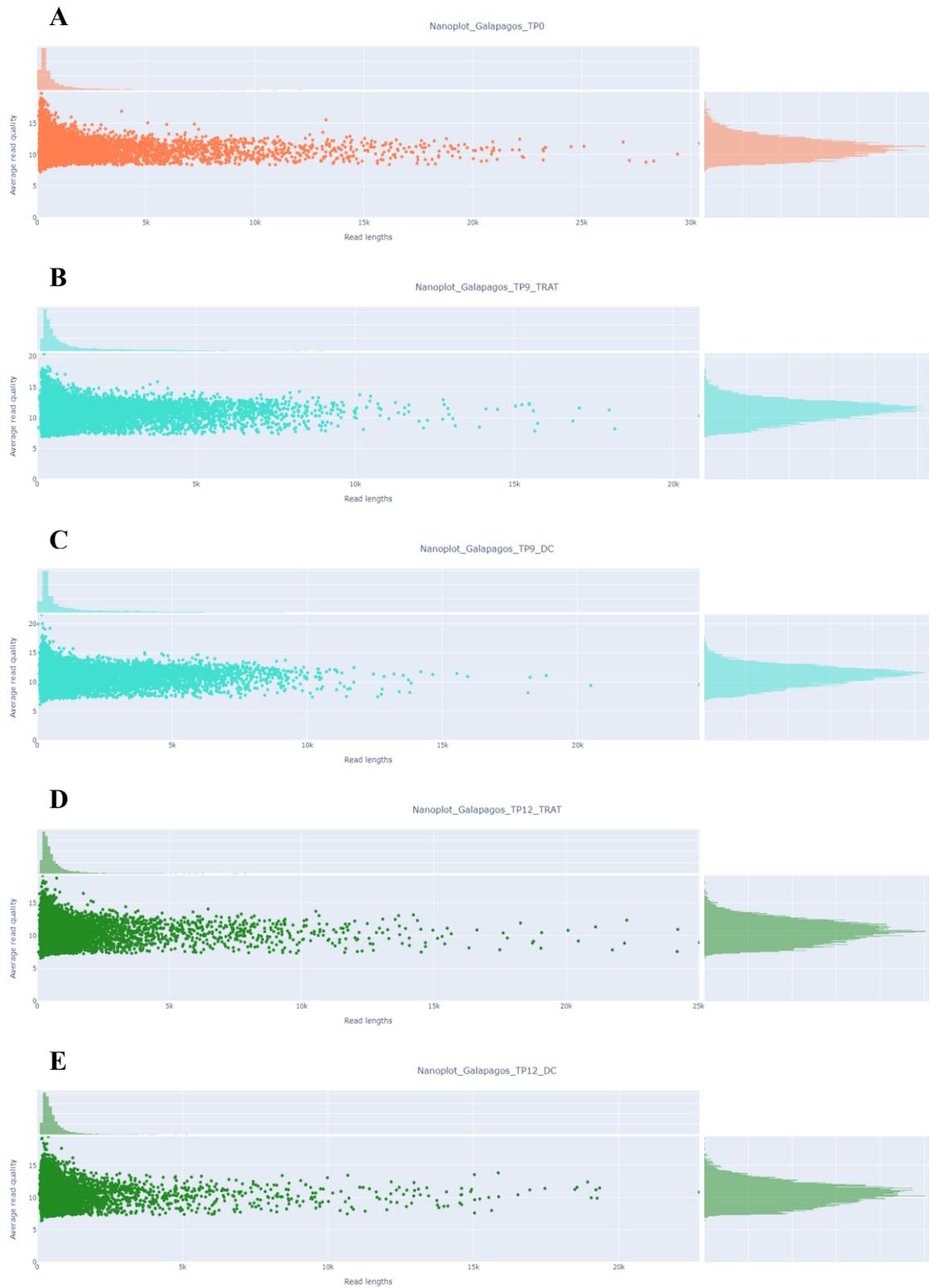


Figura 1.- Diagrama de puntos de la calidad media del secuenciamiento con respecto a la longitud de las lecturas.

A) Diagrama correspondiente al secuenciamiento del TP0. B) Diagrama correspondiente al secuenciamiento tratamiento del TP9. C) Diagrama correspondiente al secuenciamiento dark control del TP9. D) Diagrama correspondiente al secuenciamiento tratamiento del TP12. E) Diagrama correspondiente al secuenciamiento dark control del TP12.

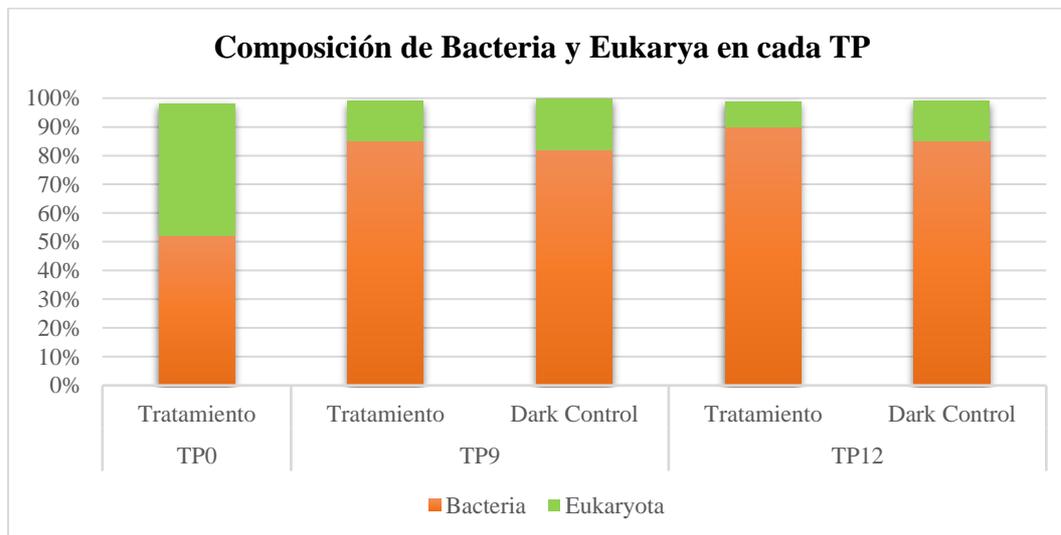


Figura 2.- Comparación de la composición taxonómica a nivel de dominio entre las 5 muestras.

Porcentaje de Bacteria y Eukarya encontradas en los diferentes tiempos de muestreo.

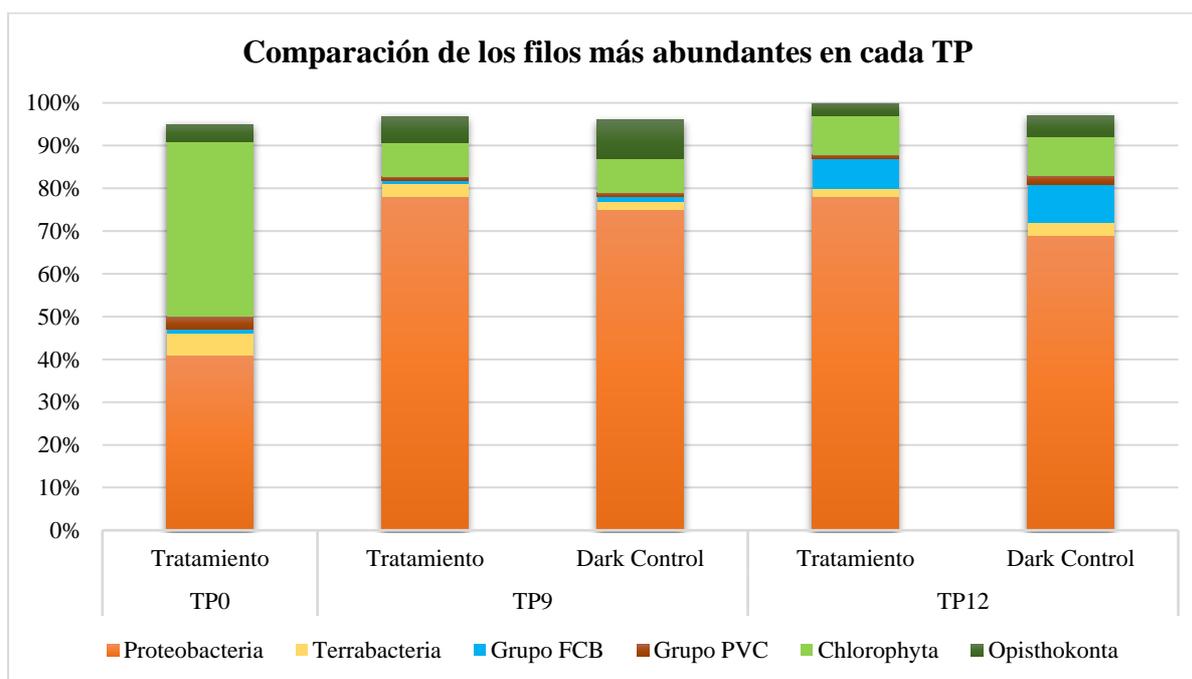


Figura 3.- Comparación de la composición taxonómica a nivel de los filos más representativos de cada muestra.

Porcentaje de lecturas asignadas a los filos más abundantes de las 5 muestras analizadas.

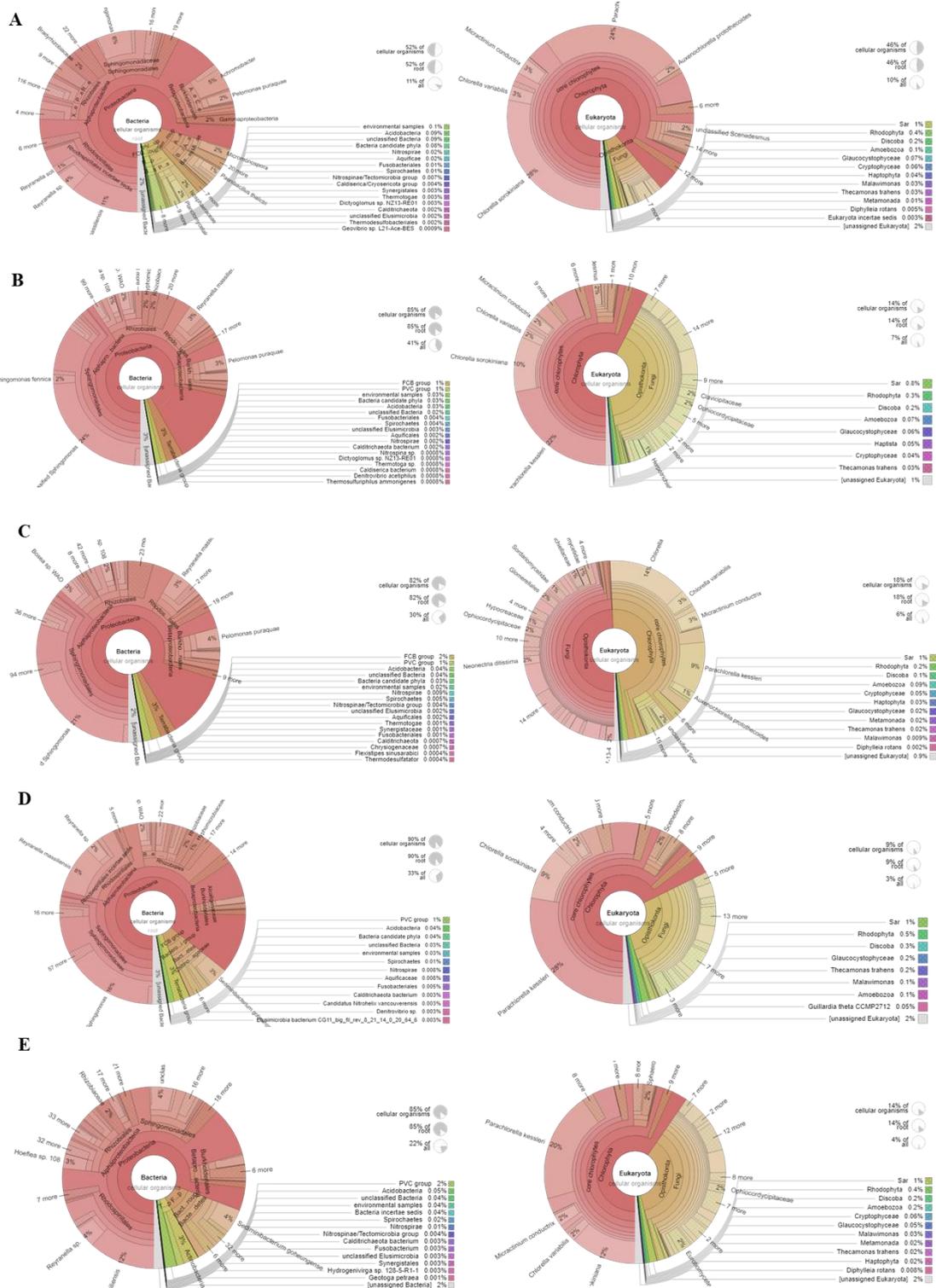


Figura 4.- Krona chart de la composición de Bacteria y Eukaryota en cada TP.

Clasificación taxonómica dentro de los 2 dominios principales encontrados. A) TP0. B) Tratamiento del TP9. C) Dark control del TP9. D) Tratamiento del TP12. E) Dark control del TP12.

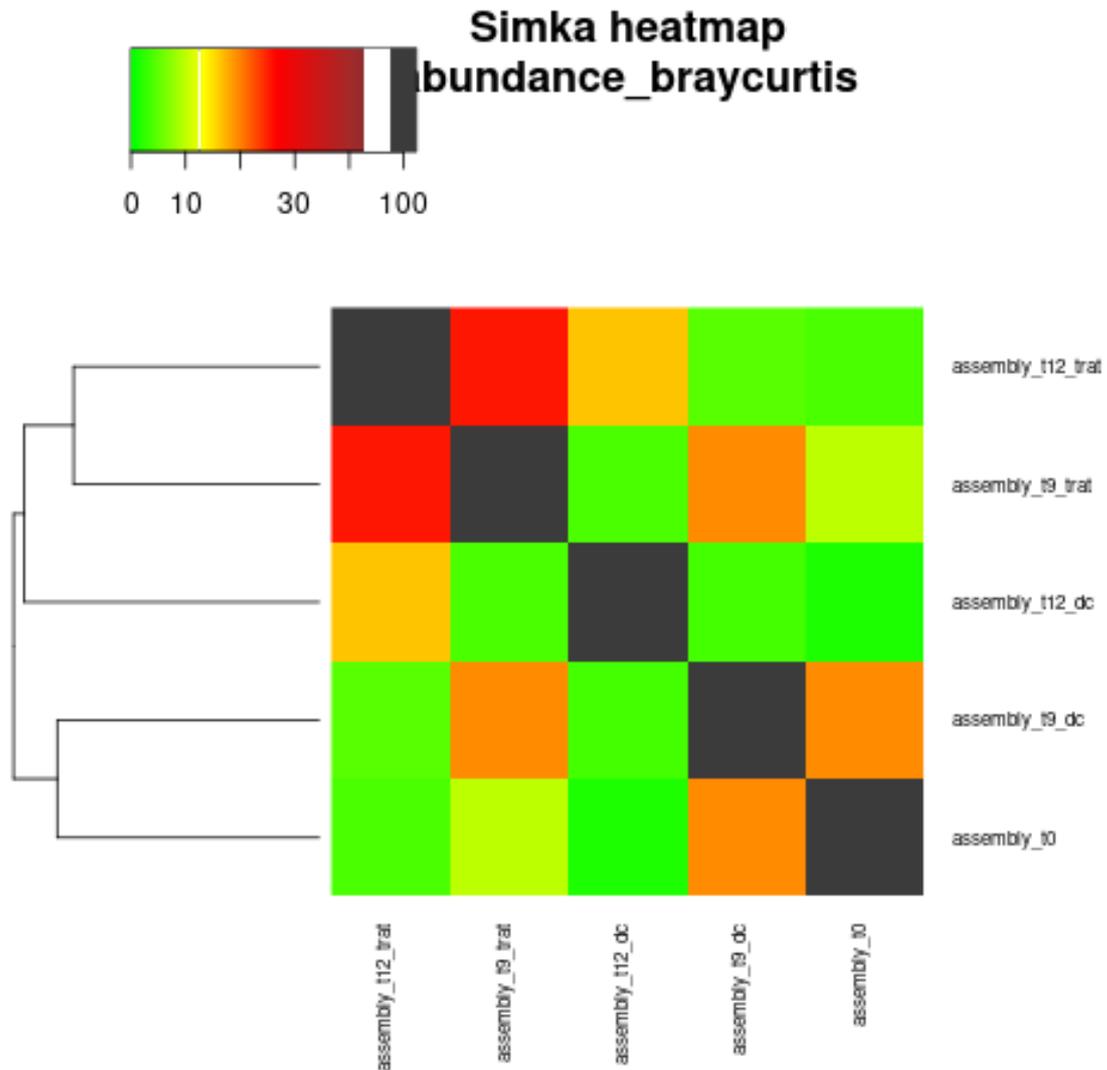


Figura 5.- Matriz de distancia basada en el índice de Bray-Curtis.

Comparación de la Beta diversidad entre las muestras donde valores de diversidad idénticos se representan con negro, valores cercanos corresponden al color rojo, valores distantes son naranjas y valores muy distantes se identifican con verde.

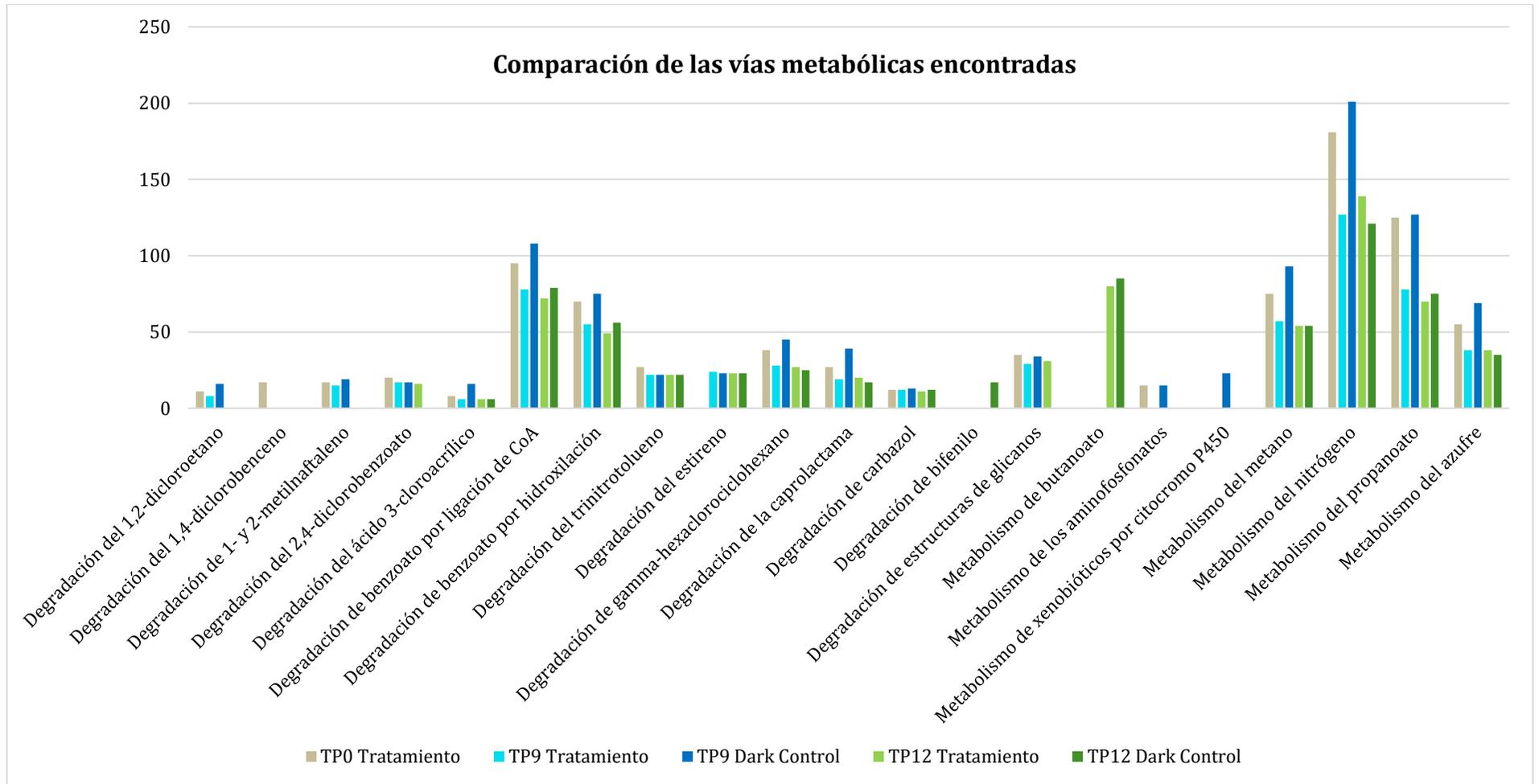


Figura 6.- Comparación de las vías metabólicas con importancia en biorremediación encontradas y el número de asignaciones de cada vía.

Vías metabólicas implicadas en el tratamiento de aguas residuales encontradas en las muestras estudiadas.

REFERENCIAS BIBLIOGRÁFICAS

- Alarcón, M. (2019). *El problema del agua para uso y consumo humano en Santa Cruz, Galápagos* [Tesis de maestría no publicada]. Universidad Andina Simón Bolívar.
- Aliabadi, M., Aroujalian, A., & Raisi, A. (2012). Removal of styrene from petrochemical wastewater using pervaporation process. *Desalination*, 284, 116-121. doi: 10.1016/j.desal.2011.08.044
- Amarasinghe, S., Su, S., Dong, X., Zappia, L., Ritchie, M., & Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biology*, 21(1). doi: 10.1186/s13059-020-1935-5
- Arellano, L., & Halffter, G. (2003). Gamma diversity: derived from and a determinant of Alpha diversity and Beta diversity. An analysis of three tropical landscapes. *Acta zoológica mexicana* (90), 27-76. Recuperado en 19 de noviembre de 2021, de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0065-17372003000300003&lng=es&tlng=en.
- Azubuike, C. C., Chikere, C. B., & Okpokwasili, G. C. (2016). Bioremediation techniques-classification based on site of application: principles, advantages, limitations and prospects. *World journal of microbiology & biotechnology*, 32(11), 180. <https://doi.org/10.1007/s11274-016-2137-x>
- Barreiro, S., González, C., & de Godos, I. (2021). Characterization of communities in a microalgae-bacteria system treating domestic wastewater reveals dominance of phototrophic and pigmented bacteria. *Algal Research*, 59, 102447. doi: 10.1016/j.algal.2021.102447
- Benoit, G., Peterlongo, P., Mariadassou, M., Drezen, E., Schbath, S., Lavenier, D., & Lemaitre, C. (2016). Multiple comparative metagenomics using multisetk-mer counting. *Peerj Computer Science*, 2, e94. doi: 10.7717/peerj-cs.94

- Bharagave, R., Purchase, D., Saxena, G., & Mulla, S. (2019). Applications of Metagenomics in Microbial Bioremediation of Pollutants: From Genomics to Environmental Cleanup. In S. Das & H. Ranjan (Ed.s), *Microbial Diversity in the Genomic Era* (pp. 459-477). Academic Press. <https://doi.org/10.1016/C2017-0-01759-7>
- Corvelo, A., Clarke, W., Robine, N., & Zody, M. (2018). taxMaps: comprehensive and highly accurate taxonomic classification of short-read data in reasonable time. *Genome Research*, 28(5), 751-758. doi: 10.1101/gr.225276.117
- Cuscó, A., Pérez, D., Viñes, J., Fàbregas, N., & Francino, O. (2021). Long-read metagenomics retrieves complete single-contig bacterial genomes from canine feces. *BMC Genomics*, 22(1). doi: 10.1186/s12864-021-07607-0
- Cydzik-Kwiatkowska, A., & Zielińska, M. (2016). Bacterial communities in full-scale wastewater treatment systems. *World Journal Of Microbiology And Biotechnology*, 32(4). doi: 10.1007/s11274-016-2012-9
- Chandran, H., Meena, M., & Sharma, K. (2020). Microbial Biodiversity and Bioremediation Assessment Through Omics Approaches. *Frontiers In Environmental Chemistry*, 1, 1-22. doi: 10.3389/fenvc.2020.570326
- Daly, A., Baetens, J., & De Baets, B. (2018). Ecological Diversity: Measuring the Unmeasurable. *Mathematics*, 6(7), 119. doi: 10.3390/math6070119
- Datta, S., Rajnish, K., Samuel, M., Pugazhendhi, A., & Selvarajan, E. (2020). Metagenomic applications in microbial diversity, bioremediation, pollution monitoring, enzyme and drug discovery. A review. *Environmental Chemistry Letters*, 18(4), 1229-1241. doi: 10.1007/s10311-020-01010-z
- De Coster, W., D'Hert, S., Schultz, D., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics*, 34(15), 2666-2669. doi: 10.1093/bioinformatics/bty149

- Delahaye, C., & Nicolas, J. (2021). Sequencing DNA with nanopores: Troubles and biases. *PLOS ONE*, *16*(10), e0257521. doi: 10.1371/journal.pone.0257521
- Fang, H., Zhang, H., Han, L., Mei, J., Ge, Q., Long, Z., & Yu, Y. (2018). Exploring bacterial communities and biodegradation genes in activated sludge from pesticide wastewater treatment plants via metagenomic analysis. *Environmental Pollution*. doi:10.1016/j.envpol.2018.09.080
- Garrido, J., Polo, M., & Oller, I. (2017). Advanced microbial analysis for wastewater quality monitoring: metagenomics trend. *Applied Microbiology And Biotechnology*, *101*(20), 7445-7458. doi: 10.1007/s00253-017-8490-3
- Gerhard, W., & Gunsch, C. (2019). Metabarcoding and machine learning analysis of environmental DNA in ballast water arriving to hub ports. *Environment International*, *124*, 312-319. doi: 10.1016/j.envint.2018.12.038
- Gifford, S., Zhao, L., Stemple, B., DeLong, K., Medeiros, P., Seim, H., & Marchetti, A. (2020). Microbial Niche Diversification in the Galápagos Archipelago and Its Response to El Niño. *Frontiers In Microbiology*, *11*. doi: 10.3389/fmicb.2020.575194
- GitHub. (2021). GitHub - rrwick/Porechop: adapter trimmer for Oxford Nanopore reads. Recuperado el 10 de Octubre de 2021, de <https://github.com/rrwick/Porechop>
- Instituto Nacional de Estadísticas y Censos. (2016). Estadística De Información Ambiental Económica en Gobiernos Autónomos Descentralizados Municipales 2015 (Agua y Alcantarillado). Recuperado el 4 de noviembre de 2021, de https://www.ecuadorencifras.gob.ec/documentos/web-inec/Encuestas_Ambientales/Municipios_2015/Documento_Tecnico-Gestion_de_Agua_y_Alcantarillado_2015.pdf

- Jain, M., Olsen, H., Paten, B., & Akeson, M. (2016). The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology*, *17*(1). doi: 10.1186/s13059-016-1103-0
- Jones, J. S., Porter, A., Muñoz-pérez, J. P., Alarcón-ruales, D., Galloway, T. S., Godley, B. J., Santillo, D., Vagg, J. And Lewis, C. (2021). Plastic contamination of a Galapagos Island (Ecuador) and the relative risks to native marine species. *Science Of The Total Environment*, *789*, 147704. doi: 10.1016/j.scitotenv.2021.147704
- Joshi, D. R., Zhang, Y., Gao, Y., Liu, Y., & Yang, M. (2017). Biotransformation of nitrogen- and sulfur-containing pollutants during coking wastewater treatment: Correspondence of performance to microbial community functional structure. *Water Research*, *121*, 338–348. doi:10.1016/j.watres.2017.05.045
- Kaevska, M., Videnska, P., & Vasickova, P. (2015). Changes in Microbial Composition of Wastewater During Treatment in a Full-Scale Plant. *Current Microbiology*, *72*(2), 128-132. doi: 10.1007/s00284-015-0924-5
- Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, *37*(5), 540-546. doi: 10.1038/s41587-019-0072-8
- Kragelund, C., Kong, Y., van der Waarde, J., Thelen, K., Eikelboom, D., & Tandoi, V. et al. (2006). Ecophysiology of different filamentous Alphaproteobacteria in industrial wastewater treatment plants. *Microbiology*, *152*(10), 3003-3012. doi: 10.1099/mic.0.29249-0
- Lapidus, A. (2009). Genome Sequence Databases: Sequencing and Assembly. En M. Schaechter (Ed.), *Encyclopedia of Microbiology* (pp. 196-210). San Diego: Academic Press.

- Liu, L., Zeng, Z., Bee, M., Gibson, V., Wei, L., Huang, X., & Liu, C. (2018). Characteristics and performance of aerobic algae-bacteria granular consortia in a photo-sequencing batch reactor. *Journal Of Hazardous Materials*, *349*, 135-142. doi: 10.1016/j.jhazmat.2018.01.059
- Loughrin, J., Vanotti, M., Szogi, A., & Lovanh, N. (2009). Evaluation of Second-Generation Multistage Wastewater Treatment System for the Removal of Malodors from Liquid Swine Waste. *Journal Of Environmental Quality*, *38*(4), 1739-1748. doi: 10.2134/jeq2008.0430
- Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics*, *14*(5), 265-279. doi: 10.1016/j.gpb.2016.05.004
- Ma, Q., Qu, Y., Zhang, X., Shen, W., Liu, Z., Wang, J., Zhang, Z. And Zhou, J. (2015). Identification of the microbial community composition and structure of coal-mine wastewater treatment plants. *Microbiological Research*, *175*, 1-5. doi: 10.1016/j.micres.2014.12.013
- Maitra, S. (2018). In Situ Bioremediation - An Overview. *Research Journal Of Life Sciences, Bioinformatics, Pharmaceutical And Chemical Sciences*, *04*(06), 576-598. doi: 10.26479/2018.0406.45
- Menzel, P., Ng, K., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications*, *7*(1). doi: 10.1038/ncomms11257
- Morales, D., Martinez, O., Kyndt, J., & Martinez, A. (2014). Heterotrophic growth of microalgae: metabolic aspects. *World Journal Of Microbiology And Biotechnology*, *31*(1), 1-9. doi: 10.1007/s11274-014-1773-2
- Mu, R., Jia, Y., Ma, G., Liu, L., Hao, K., Qi, F., & Shao, Y. (2021). Advances in the use of microalgal–bacterial consortia for wastewater treatment: Community structures,

- interactions, economic resource reclamation, and study techniques. *Water Environment Research*, 93(8), 1217-1230. doi: 10.1002/wer.1496
- Nascimento, A., Souza, A., Andrade, P., Andreote, F., Coscione, A., Oliveira, F., & Regitano, J. (2018). Sewage Sludge Microbial Structures and Relations to Their Sources, Treatments, and Chemical Attributes. *Frontiers In Microbiology*, 9. doi: 10.3389/fmicb.2018.01462
- Niestępski, S., Harnisz, M., Ciesielski, S., Korzeniewska, E., & Osińska, A. (2020). Environmental fate of Bacteroidetes, with particular emphasis on Bacteroides fragilis group bacteria and their specific antibiotic resistance genes, in activated sludge wastewater treatment plants. *Journal Of Hazardous Materials*, 394, 122544. doi: 10.1016/j.jhazmat.2020.122544
- O'Flaherty, E., & Gray, N. (2013). A comparative analysis of the characteristics of a range of real and synthetic wastewaters. *Environmental Science And Pollution Research*, 20(12), 8813-8830. doi: 10.1007/s11356-013-1863-y
- Oksanen, J., Blanchet, F., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P., O'hara, R., Simpson, G., Solymos, P., Stevens, M., Szoecs, E. And Wagner, H. (2020). Community Ecology Package. Recuperado en 19 de noviembre de 2021, de <https://cran.r-project.org/web/packages/vegan/vegan.pdf>
- Ordoñez, M. (2020). Análisis y comparación de metagenomas de consorcios microbianos de Galápagos con capacidad de remediación de aguas residuales [Tesis de pregrado no publicada]. Universidad San Francisco de Quito.
- Oxford Nanopore Technologies. (2021). Native barcoding genomic DNA (with EXP-NBD104, EXP-NBD114, and SQK-LSK109). Oxford, UK

- Parque Nacional Galápagos. (2017). Calidad de Agua – Dirección del Parque Nacional Galápagos. Recuperado el 4 de noviembre de 2021, de <http://www.galapagos.gob.ec/calidad-de-agua/>
- Peña, S., Mayorga, J., & Montoya, R. (2018). Propuesta de tratamiento de las aguas residuales de la ciudad de Yaguachi (Ecuador). *Ciencia e Ingeniería*, 39(2), 161-167.
- Pla, Laura. (2006). Biodiversidad: Inferencia basada en el índice de Shannon y la riqueza. *Interciencia*, 31(8), 583-590. Recuperado el 19 de noviembre de 2021, de http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0378-18442006000800008&lng=es&tlng=es.
- Płaza, G., Jałowiecki, Ł., Głowacka, D., Hubeny, J., Harnisz, M., & Korzeniewska, E. (2021). Insights into the microbial diversity and structure in a full-scale municipal wastewater treatment plant with particular regard to Archaea. *PLOS ONE*, 16(4), e0250514. doi: 10.1371/journal.pone.0250514
- Qiagen. (2020). DNeasy PowerSoil Pro Kit Handbook - QIAGEN. Recuperado el 10 de Octubre de 2021, de <https://www.qiagen.com/nl/resources/resourcedetail?id=9bb59b74-e493-4aeb-b6c1-f660852e8d97&lang=en>
- Ramanan, R., Kang, Z., Kim, B.-H., Cho, D.-H., Jin, L., Oh, H.-M., & Kim, H.-S. (2015). Phycosphere bacterial diversity in green algae reveals an apparent similarity across habitats. *Algal Research*, 8, 140–144. doi:10.1016/j.algal.2015.02.003
- Ramanan, R., Kim, B.-H., Cho, D.-H., Oh, H.-M., & Kim, H.-S. (2016). Algae–bacteria interactions: Evolution, ecology and emerging applications. *Biotechnology Advances*, 34(1), 14–29. doi:10.1016/j.biotechadv.2015.12.003
- Roberts, D. (2017). Distance, dissimilarity, and mean–variance ratios in ordination. *Methods In Ecology And Evolution*, 8(11), 1398-1407. doi: 10.1111/2041-210x.12739

- Roswell, M., Dushoff, J., & Winfree, R. (2021). A conceptual guide to measuring species diversity. *Oikos*, *130*(3), 321-338. doi: 10.1111/oik.07202
- Roy, A., Dutta, A., Pal, S., Gupta, A., Sarkar, J., Chatterjee, A., Saha, A., Sarkar, P., Sar, P. And Kazy, S. K. (2018). Biostimulation and bioaugmentation of native microbial community accelerated bioremediation of oil refinery sludge. *Bioresource Technology*, *253*, 22-32. doi: 10.1016/j.biortech.2018.01.004
- Shafin, K., Pesout, T., Lorig-Roach, R., Haukness, M., Olsen, H. E., Bosworth, C., Armstrong, J., Tigyi, K., Maurer, N., Koren, S., Sedlazeck, F. J., Marschall, T., Mayes, S., Costa, V., Zook, J. M., Liu, K. J., Kilburn, D., Sorensen, M., Munson, K. M., Vollger, M. R., Monlong, J., Garrison, E., Eichler, E. E., Salama, S., Haussler, D., Green, R. E., Akeson, M., Phillippy, A., Miga, K. H., Carnevali, P., Jain, M. And Paten, B. (2020). Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. *Nature Biotechnology*, *38*(9), 1044-1053. doi: 10.1038/s41587-020-0503-6
- Shi, S., Qu, Y., Ma, F., & Zhou, J. (2014). Bioremediation of coking wastewater containing carbazole, dibenzofuran and dibenzothiophene by immobilized naphthalene-cultivated *Arthrobacter* sp. W1 in magnetic gellan gum. *Bioresource Technology*, *166*, 79-86. doi: 10.1016/j.biortech.2014.05.036
- Sidhu, C., Vikram, S., & Pinnaka, A. (2017). Unraveling the Microbial Interactions and Metabolic Potentials in Pre- and Post-treated Sludge from a Wastewater Treatment Plant Using Metagenomic Studies. *Frontiers In Microbiology*, *8*. doi: 10.3389/fmicb.2017.01382
- Silva, C., Hayden, H., Sawbridge, T., Mele, P., De Paula, S., Silva, L., Vidigal, P., Vicentini, R., Sousa, M., Torres, A., Santiago, V., & Oliveira, V. (2013). Identification of Genes and Pathways Related to Phenol Degradation in Metagenomic Libraries from

- Petroleum Refinery Wastewater. *Plos ONE*, 8(4), e61811. doi: 10.1371/journal.pone.0061811
- Studnik, H., Liebsch, S., Forlani, G., Wieczorek, D., Kafarski, P., & Lipok, J. (2015). Amino polyphosphonates – chemical features and practical uses, environmental durability and biodegradation. *New Biotechnology*, 32(1), 1–6. doi:10.1016/j.nbt.2014.06.007
- Suárez, A. (2019). *A preliminary view of the biogeography, phylogeography and ecology of freshwater shrimps (Atyidae) from Isla San Cristóbal, Galápagos Archipelago* [Tesis de licenciatura no publicada]. Universidad San Francisco de Quito.
- Subashchandrabose, S. R., Ramakrishnan, B., Megharaj, M., Venkateswarlu, K., & Naidu, R. (2011). Consortia of cyanobacteria/microalgae and bacteria: Biotechnological potential. *Biotechnology Advances*, 29(6), 896–907. doi:10.1016/j.biotechadv.2011.07.009
- Tamames, J., & Puente, F. (2019). SqueezeMeta, A Highly Portable, Fully Automatic Metagenomic Analysis Pipeline. *Frontiers In Microbiology*, 9. doi: 10.3389/fmicb.2018.03349
- Tamminen, M., Spaak, J., Tlili, A., Eggen, R., Stamm, C., & Räsänen, K. (2021). Wastewater constituents impact biofilm microbial community in receiving streams. *Science Of The Total Environment*, 151080. doi: 10.1016/j.scitotenv.2021.151080
- Tangahu, B., Berlianto, M., & Kartika, A. (2020). Deconcentration of Chromium Contained in Wastewater Using a Bacteria and Microalgae Consortia with a High Rate Algal Reactor System. *Journal Of Ecological Engineering*, 21(8), 272-284. doi: 10.12911/22998993/126878
- Tiwari, A., Hokajärvi, A., Domingo, J. S., Elk, M., Jayaprakash, B., Ryu, H., Siponen, S., Vepsäläinen, A., Kauppinen, A., Puurunen, O., Artimo, A., Perkola, N., Huttula, T., Miettinen, I. T. And Pitkänen, T. (2021). Bacterial diversity and predicted enzymatic function in a multipurpose surface water system – from wastewater effluent discharges

- to drinking water production. *Environmental Microbiome*, 16(1). doi: 10.1186/s40793-021-00379-w
- Tovo, A., Menzel, P., Krogh, A., Cosentino Lagomarsino, M., & Suweis, S. (2020). Taxonomic classification method for metagenomics based on core protein families with Core-Kaiju. *Nucleic Acids Research*, 48(16), e93-e93. doi: 10.1093/nar/gkaa568
- Tyler, A., Mataseje, L., Urfano, C., Schmidt, L., Antonation, K., Mulvey, M., & Corbett, C. (2018). Evaluation of Oxford Nanopore's MinION Sequencing Device for Microbial Whole Genome Sequencing Applications. *Scientific Reports*, 8(1). doi: 10.1038/s41598-018-29334-5
- Velasco, G., Moncayo, J., & Chuquer, D. (2018). Diagnóstico Del Sistema De Tratamiento De Aguas Residuales De Manta. *Infoanalítica*, 7(1), 27-39. doi: 10.26807/ia.v7i1.90
- Vishwakarma, G., Bhattacharjee, G., Gohil, N., Singh, V. (2020). Current status, challenges and future of bioremediation. En V. Pandey & V. Singh (Ed.s), *Bioremediation of Pollutants* (pp. 445-462). Elsevier. <https://doi.org/10.1016/B978-0-12-819025-8.00020-X>
- Walsh, S., & Mena, C. (2013). *Science and Conservation in the Galapagos Islands* (pp. 1-2). New York: Springer.
- Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Completing bacterial genome assemblies with multiplex MinION sequencing. *Microbial genomics*, 3(10), e000132. <https://doi.org/10.1099/mgen.0.000132>
- Wiegand, S., Dam, H., Riba, J., Vollmers, J., & Kaster, A. (2021). Printing Microbial Dark Matter: Using Single Cell Dispensing and Genomics to Investigate the Patescibacteria/Candidate Phyla Radiation. *Frontiers In Microbiology*, 12. doi: 10.3389/fmicb.2021.635506

- Willis, A. (2019). Rarefaction, Alpha Diversity, and Statistics. *Frontiers In Microbiology*, 10. doi: 10.3389/fmicb.2019.02407
- Wu, B., Liu, F., Fang, W., Yang, T., Chen, G., He, Z., & Wang, S. (2021). Microbial sulfur metabolism and environmental implications. *Science Of The Total Environment*, 778, 146085. doi: 10.1016/j.scitotenv.2021.146085
- Xie, N., Zhong, L., Ouyang, L., Xu, W., Zeng, Q., & Wang, K. et al. (2021). Community Composition and Function of Bacteria in Activated Sludge of Municipal Wastewater Treatment Plants. *Water*, 13(6), 852. doi: 10.3390/w13060852
- Yadav, R., Rajput, V., & Dharne, M. (2021). Functional metagenomic landscape of polluted river reveals potential genes involved in degradation of xenobiotic pollutants. *Environmental Research*, 192, 110332. doi: 10.1016/j.envres.2020.110332
- Ye, L., & Zhang, T. (2012). Bacterial communities in different sections of a municipal wastewater treatment plant revealed by 16S rDNA 454 pyrosequencing. *Applied Microbiology And Biotechnology*, 97(6), 2681-2690. doi: 10.1007/s00253-012-4082-4
- Yu, J., Tang, S. N., & Lee, P. (2021). Microbial Communities in Full-Scale Wastewater Treatment Systems Exhibit Deterministic Assembly Processes and Functional Dependency over Time. *Environmental Science & Technology*, 55(8), 5312–5323. doi:10.1021/acs.est.0c06732
- Zerva, I., Remmas, N., Kagalou, I., Melidis, P., Ariantsi, M., Sylaios, G., & Ntougias, S. (2021). Effect of Chlorination on Microbiological Quality of Effluent of a Full-Scale Wastewater Treatment Plant. *Life (Basel, Switzerland)*, 11(1), 68. <https://doi.org/10.3390/life11010068>
- Zhang, B., Li, W., Guo, Y., Zhang, Z., Shi, W., & Cui, F. et al. (2020). Microalgal-bacterial consortia: From interspecies interactions to biotechnological applications. *Renewable And Sustainable Energy Reviews*, 118, 109563. doi: 10.1016/j.rser.2019.109563

- Zhang, B., Yang, T., Sun, C., & Wen, X. (2021). Drivers of microbial beta-diversity in wastewater treatment plants in China. *Journal Of Environmental Sciences*, 115, 341-349. doi: 10.1016/j.jes.2021.07.028
- Zhang, L., Shen, Z., Fang, W., & Gao, G. (2019). Composition of bacterial communities in municipal wastewater treatment plant. *Science Of The Total Environment*, 689, 1181-1191. doi: 10.1016/j.scitotenv.2019.06.432
- Zheng, C., Zhao, L., Zhou, X., Fu, Z., & Li, A. (2013). Treatment Technologies for Organic Wastewater. En Walid Elshorbagy (Ed.), *Water Treatment*. Emiratos Árabes Unidos: Universidad de Emiratos Árabes Unidos. Recuperado de: <https://www.intechopen.com/chapters/41953>

ANEXOS

ANEXO 1: ABUNDANCIA RELATIVA DE LAS CLASES PREDOMINANTES DENTRO DE LAS MUESTRAS

