



UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

FACULTAD DE CIENCIAS QUÍMICAS
POSGRADO EN CIENCIAS EN BIOPROCESOS



CARACTERIZACIÓN METAGENÓMICA DE *Chlorella vulgaris* Y *Scenedesmus acutus* PARA EL
ESTUDIO DE RUTAS METABÓLICAS INTERRELACIONADAS CON BACTERIAS

Tesis que para obtener el grado de:

Maestría en Ciencias en Bioprocesos

Presenta:

Martínez Constantino Abner

Director de tesis:

Dra. Ruth Elena Soria Guerra

Co-director de tesis:

Dr. Mishaél Sánchez Pérez

San Luis Potosí, S. L. P.

Agosto 2026



REPOSITORIO INSTITUCIONAL



UASLP-Sistema de Bibliotecas

Repositorio Institucional Tesis digitales Restricciones de uso

DERECHOS RESERVADOS

PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en este Trabajo Terminal está protegido por la Ley Federal de Derecho de Autor (LFDA) de los Estados Unidos Mexicanos.

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde se obtuvo, mencionando el autor o autores. Cualquier uso distinto o con fines de lucro, reproducción, edición o modificación será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



Caracterización metagenómica de *Chlorella vulgaris* y *Scenedesmus acutus* para el estudio de rutas metabólicas interrelacionadas con bacterias © 2026 por Martínez Constantino, Abner se distribuye bajo una licencia Creative Commons Atribución–No Comercial–Sin Derivadas 4.0 Internacional

Declaraciones Institucionales

Este proyecto se realizó en Laboratorio de Biotecnología de Células Vegetales y Bioinformática, adscrito a la Facultad de Ciencias Químicas de la Universidad Autónoma de San Luis Potosí, en el periodo comprendido entre agosto 2024 y agosto 2026, bajo la dirección de la Dra. Ruth Elena Soria Guerra fue apoyado por el Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI).

El programa de Maestría en Ciencias en Bioprocesos de la Universidad Autónoma de San Luis Potosí pertenece al Sistema Nacional de Posgrados (SNP) del SECIHTI, registro 000588.

Número de la beca otorgada por SECIHTI: **4030879**

Número de CVU: **2013347**

Los datos del trabajo titulado "CARACTERIZACIÓN METAGENÓMICA DE *Chlorella vulgaris* Y *Scenedesmus acutus* PARA EL ESTUDIO DE RUTAS METABÓLICAS INTERRELACIONADAS CON BACTERIAS" se encuentran bajo el resguardo de la Facultad de Ciencias Químicas y pertenecen a la Universidad Autónoma de San Luis Potosí.



UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

Facultad de Ciencias Químicas

Centro de Investigación y Estudios de Posgrado

Posgrado en Ciencias en Bioprocesos

Programa de Maestría

Formato M13

Solicitud de Registro de Tesis Maestría

San Luis Potosí SLP a septiembre/ 20 /2024

Comité Académico

En atención a: Coordinador/a del Posgrado

Por este conducto solicito a Usted ~~se lleve a cabo el registro de tema de tesis de~~ Maestría, el cual quedo definido de la siguiente manera: “CARACTERIZACIÓN METAGENÓMICA DE *Chlorella vulgaris* Y *Scenedesmus acutus* PARA EL ESTUDIO DE RUTAS METABÓLICAS INTERRELACIONADAS CON BACTERIAS” que desarrollará el/la estudiante: Abner Martínez Constantino bajo la dirección y/o Co-dirección de: Dra. Ruth Elena Soria Guerra y Dr. Mishael Sánchez Pérez.

Asimismo, le comunico que el proyecto en el cual trabajará el alumno involucrará el manejo de animales de experimentación, estudios con seres humanos o muestras derivadas de los mismos, el manejo y/o generación de organismos genéticamente modificados y requiere de aval de Comité de Ética e investigación de la FCQ.

(Complete la opción que aplique en su caso):

() Sí debido a que: _____

(x) No

() No Aplica

Sin otro particular, quedo de Usted.

A T E N T A M E N T E

Abner Martínez Constantino

Dra. Ruth Elena Soria Guerra



UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ



FACULTAD DE CIENCIAS QUÍMICAS
POSGRADO EN CIENCIAS EN BIOPROCESOS

Título del trabajo

CARACTERIZACIÓN METAGENÓMICA DE *Chlorella vulgaris* Y *Scenedesmus acutus* PARA EL ESTUDIO DE RUTAS METABÓLICAS INTERRELACIONADAS CON BACTERIAS

Tesis que para obtener el grado de:

Maestro en Ciencias en Bioprocesos

Presenta:

Martínez Constantino Abner

Director de tesis:

Dra. Ruth Elena Soria Guerra

SINODALES:

Presidente:

Dra. Catalina Arenas Huertero.

Secretario:

Dra. Ruth Elena Soria Guerra.

Vocal:

Dr. Mishael Sánchez Pérez.

Vocal:

Dra. Luzmila Martínez González.

San Luis Potosí, S. L. P.

Agosto 2026

INTEGRANTES DEL COMITÉ TUTORIAL ACADÉMICO

Dra. Ruth Elena Soria Guerra: Director de tesis. Adscrita al Posgrado en Ciencias en Bioprocesos de la Universidad Autónoma de San Luis Potosí, San Luis Potosí, S.L.P.

Dr. Mishael Sánchez Pérez. Co-director de tesis. Adscrito a la Facultad de Ciencias del Deporte de la Universidad Autónoma del estado de Morelos, Cuernavaca, Morelos.

Dra. Catalina Arenas Huertero: Sinodal de tesis. Adscrita al Posgrado en Ciencias en Bioprocesos de la Universidad Autónoma de San Luis Potosí, San Luis Potosí, S.L.P.



UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

Facultad de Ciencias Químicas

Centro de Investigación y Estudios de Posgrado

Posgrado en Ciencias en Bioprocesos

Programa de Maestría

Formato M5

Carta Cesión de Derechos

San Luis Potosí SLP a

agosto / 21 / 2026

En la ciudad de San Luis Potosí, el día 21 del mes de agosto del año 2026. El que suscribe Abner Martínez Constantino Alumno(a) del programa de posgrado en Ciencias en Bioprocesos adscrito a la Facultad de Ciencias Químicas, manifiesta que es autor(a) intelectual del presente trabajo terminal, realizado bajo la dirección de: la Dra. Ruth Elena Soria Guerra y Dr. Mishael Sánchez Pérez y cede los derechos del trabajo titulado “ CARACTERIZACIÓN METAGENÓMICA DE *Chlorella vulgaris* Y *Scenedesmus acutus* PARA EL ESTUDIO DE RUTAS METABÓLICAS INTERRELACIONADAS CON BACTERIAS” a la **Universidad Autónoma de San Luis Potosí**, para su difusión con fines académicos y de investigación.

Los usuarios de la información no deben reproducir de forma total o parcial texto, gráficas, imágenes o cualquier contenido del trabajo si el permiso expreso del o los autores. Éste, puede ser obtenido directamente con el autor o autores escribiendo a la siguiente dirección abnermtz45@gmail.com. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

Nombre y firma del alumno



UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

Facultad de Ciencias Químicas
Centro de Investigación y Estudios de Posgrado
Posgrado en Ciencias en Bioprocesos
Programa de Maestría

Formato M28

Carta de Análisis de Similitud

San Luis Potosí SLP a 20/08/2026

LB. María Zita Acosta Nava
Biblioteca de Posgrado
Facultad de Ciencias Químicas
Universidad Autónoma de San Luis Potosí

Asunto: Reporte de porcentaje de similitud de tesis de grado

Por este medio me permito informarle el porcentaje de similitud obtenido mediante Ithenticate para la tesis titulada “CARACTERIZACION METAGENÓMICA DE *Chlorella vulgaris* Y *Scenedesmus acutus* PARA EL ESTUDIO DE RUTAS METABOLICAS INTERRELACIONADAS CON BACTERIAS”. presentada por el autor Abner Martínez Constantino. La tesis es requisito para obtener el grado de Maestría en el Posgrado en Ciencias en Bioprocesos. El análisis reveló un porcentaje de similitud de 6% excluyendo referencias y metodología.

Agradezco sinceramente su valioso tiempo y dedicación para llevar a cabo una exhaustiva revisión de la tesis. Quedo a su disposición para cualquier consulta o inquietud que pueda surgir en el proceso.

Sin más por el momento, le envío un cordial saludo.

ATENTAMENTE

Dra. Alicia Grajales Lagunes
Coordinador Académico del Posgrado
en Ciencias en Bioprocesos

Dedicatorias

Agradezco a todos aquellos que me acompañaron, pero principalmente a Dios por otorgarme los talentos necesarios para hacer lo que me gusta: la ciencia.

A mi familia por siempre ha estado conmigo con especial énfasis a mi madre.

A la Dra. Ruth Elena Soria Guerra quien desde mis inicios de la licenciatura me guio de una forma excepcional.

Al Dr. Mishael Sánchez Pérez porque, además de convertirse en un asesor, se ha convertido en un gran amigo y ejemplo a lo que me quiero dedicar.

A la Dra. Caty Arenas, sus pláticas y consejos siempre alentándome a seguir este camino de la ciencia nunca las olvidaré.

Al Dr. José Juan Torres Martínez por siempre estar ahí, que me sacaba cuando nada funcionaba y les daba claridad a mis ideas, fuiste tu quien me alentó y formó en estos temas, siendo mentor de muchos temas de los que ahora entiendo.

A la IBQ (y próximamente Maestra en Ciencias Químicas) María del Carmen Antonio Reyes por siempre confrontar mis ideas, escucharlas y ayudarme a siempre a ser mejor (el café sigue siendo solo café).

Al Dr. Jesús (El Chuy) por siempre estar ahí, en silencio, siempre apoyándome.

A todos los niños de servicio, verano de las ciencias y proyecto terminal con los que tuve el placer de convivir y enseñarles algo de lo poco que sé, ustedes fueron el motivo de quererme dedicar a la ciencia.

Al final solo puedo decir... xd

Resumen

Las interacciones entre microalgas y bacterias constituyen un componente fundamental de los ecosistemas acuáticos, ya que favorecen el intercambio de nutrientes, vitaminas y metabolitos que influyen en el crecimiento y la estabilidad de ambas comunidades. No obstante, los mecanismos metabólicos que sustentan estas asociaciones aún son poco comprendidos. En este trabajo se desarrolló una estrategia bioinformática integrativa para caracterizar el potencial funcional de bacterias asociadas a *Chlorella vulgaris* y *Scenedesmus acutus* mediante el análisis conjunto de datos metagenómicos y metatranscriptómicos.

A partir del ensamblaje metagenómico se reconstruyeron genomas bacterianos ensamblados por metagenoma (MAGs) de alta calidad, los cuales fueron evaluados y anotados funcionalmente utilizando bases de datos especializadas. Posteriormente, se integró la información de expresión génica para construir matrices funcionales enfocadas en rutas metabólicas relacionadas con la comunicación e intercambio de metabolitos, incluyendo biosíntesis de biotina, riboflavina, sideróforos y mecanismos de percepción de quórum. Asimismo, se desarrollaron redes metabólicas que permitieron visualizar la distribución de funciones entre los MAGs e inferir posibles interacciones ecológicas.

Como complemento, se implementó un modelo de aprendizaje automático basado en Random Forest para identificar las funciones metabólicas con mayor contribución a la importancia funcional de los MAGs dentro de la red, permitiendo priorizar procesos metabólicos potencialmente relevantes en la interacción microalga-bacteria. En conjunto, los resultados proporcionan un marco reproducible para explorar asociaciones microbianas mediante datos ómicos y aportan evidencia sobre los mecanismos metabólicos que podrían favorecer relaciones mutualistas con potencial aplicación en biotecnología y optimización de cultivos microalgales.

Palabras clave: Microalgas, metagenómica, bioinformática, aprendizaje automático.

Abstract.

Microalgae-bacteria interactions are fundamental components of aquatic ecosystems, promoting the exchange of nutrients, vitamins, and metabolites that influence the growth and stability of both partners. However, the metabolic mechanisms underlying these associations remain poorly understood. This study presents an integrative bioinformatics framework to characterize the functional potential of bacterial communities associated with *Chlorella vulgaris* and *Scenedesmus acutus* through the combined analysis of metagenomic and metatranscriptomic data.

Metagenomic assembly enabled the reconstruction of thirteen high-quality metagenome-assembled genomes (MAGs), which were subsequently assessed and functionally annotated using specialized databases. Gene expression data were integrated to generate functional matrices focused on metabolic pathways involved in microbial communication and metabolite exchange, including biotin biosynthesis, riboflavin production, siderophore metabolism, and quorum sensing. In addition, metabolic networks were constructed to visualize the distribution of these functions among MAGs and to infer potential ecological interactions.

To complement these analyses, a Random Forest machine learning model was implemented to identify the metabolic functions that contributed most strongly to the functional importance of each MAG within the reconstructed network, enabling the prioritization of pathways potentially involved in microalgae-bacteria interactions. Overall, this work provides a reproducible computational framework for studying microbial associations using multi-omics data and offers new insights into the metabolic mechanisms that may promote mutualistic interactions, with potential applications in biotechnology, biomass production, and the optimization of microalgal cultivation systems.

Keywords: Microalgae, metagenomics, bioinformatics, machine learning.

Índice general.

1. Introducción.....	1
2. Antecedentes.....	3
2.1. Consorcios de microorganismos.....	3
2.1.1. Consorcios microalga-bacteria frente a contaminantes de interés.....	4
2.2. Bioinformática e inteligencia artificial.....	6
2.3. <i>Chlorella vulgaris</i> y <i>Scenedesmus acutus</i> como candidatos en la remoción de contaminantes emergentes y metales pesados.....	9
3. Justificación.....	10
4. Hipótesis.....	10
5. Objetivo.....	10
6. Objetivos específicos.....	11
7. Materiales y métodos.....	11
7.1. Control de calidad y preprocesamiento.....	11
7.2. Ensamblaje del metagenoma.....	11
7.3. Análisis taxonómico.....	12
7.4. Anotación funcional.....	12
7.5. Inferencia de rutas metabólicas compartidas.....	12
7.6. Análisis exploratorio mediante aprendizaje automático.....	13
7.7. Visualización de resultados.....	13
8. Resultados.....	14
8.1. Calidad del ensamblado metagenómico.....	14
8.2. Análisis taxonómico.....	15
8.2.1. Recuperación de genomas bacterianos (binning metagenómico).....	18
8.2.4. Análisis ANI para la confirmación de genomas nucleares de <i>C. vulgaris</i> y <i>S. acutus</i>	29
8.3. Anotación funcional.....	32
8.3.1. Asignación funcional de proteínas hipotéticas y validación funcional mediante alineamientos con BLASTp.....	34
8.3.2. Detección de genes implicados en la biorremediación de fármacos y metales pesados.....	41
8.4. Elucidación de mapas metabólicos de interés.....	44
8.4.1. Ruta de biosíntesis de lípidos.....	44

8.4.2. Ruta de absorción de metales pesados.	49
8.4.3. Ruta de resistencia a antibióticos (β -lactámicos).....	51
8.5. Análisis de redes bipartita.	56
8.6. Análisis de expresión génica en los ensamblajes metagenómicos a partir de datos de RNA-seq.	58
8.6.1. Efecto de los tratamientos lumínicos en la actividad transcripcional de genes relacionados con la resistencia a antibióticos y biosíntesis de lípidos. .	63
8.7. Análisis exploratorio basado en algoritmos de IA (Machine Learning).....	70
8.7.1. Construcción de matrices funcionales.....	71
8.7.2. Cálculo de centralidad e influencia ecológica (Influence Score).....	73
8.7.3. Planteamiento del modelo de aprendizaje automático.	76
9. Conclusiones.	81
10. Bibliografía.	84

Índice de figuras.

Figura 1. Proceso de secuenciación masiva de última generación Illumina.....	3
<i>Figura 2. Análisis taxonómico (diagrama de Sankey) del metagenoma en C. vulgaris hecho con el software Kraken2 y Pavian, ordenado desde dominio hasta especie.</i>	16
Figura 3. Análisis taxonómico (diagrama de Sankey) del metagenoma en S. acutus hecho con el software Kraken2 y Pavian, ordenado desde dominio hasta especie.	17
Figura 4. Bins recuperados del metagenoma de C. vulgaris.	19
Figura 5. MAGs recuperados del metagenoma de S. acutus.	20
Figura 6. Resistomas encontrados en C. vulgaris.	24
Figura 7. Resistomas encontrados en S. acutus.	24
Figura 8. Genes de transferencia horizontal encontrados en el metagenoma de C. vulgaris.	26
Figura 9. Genes de transferencia horizontal encontrados en el metagenoma de S. acutus.	26
Figura 10. Análisis de completitud BUSCO.	27
Figura 11. Estadísticas de los genomas nucleares.	29
Figura 12. Porcentaje De identidad obtenido con el análisis ANI comparando diferentes cepas de ambas microalgas (C. vulgaris y S. acutus), de igual manera, la misma comparación con el genoma de Chlamydomonas reinhardtii como ancenstro común.	30
Figura 13. Porcentaje de cobertura obtenido con el análisis ANI con los mismos genomas antes mencionados.	31
Figura 14. Distribución de categorías funcionales COG en el metagenoma de C. vulgaris.	33
Figura 15. Distribución de categorías funcionales COG en el metagenoma de S. acutus.	34
Figura 16. Dispersión de Bitscores contra -Log10(E-value) en el alineamiento con BLASTp en C. vulgaris.	35

Figura 17. Dispersión de Bitscores contra -Log10(E-value) en el alineamiento con BLASTp en <i>S. acutus</i> .	36
Figura 18. Distribución del porcentaje de identidad contra su frecuencia en <i>C. vulgaris</i> .	37
Figura 19. Distribución del porcentaje de identidad contra su frecuencia en <i>S. acutus</i> .	38
Figura 20. Calidad de los alineamientos en <i>C. vulgaris</i> .	39
Figura 21. Calidad de los alineamientos en <i>S. acutus</i> .	40
Figura 22. Mapa metabólico de la biosíntesis de ácidos grasos. Verde: <i>Pseudomonas oryzae</i> . Rojo: <i>Microvirga</i> Sp VF16.	45
Figura 23. Mapa metabólico de la biosíntesis de ácidos grasos. Verde: <i>Chlorella vulgaris</i> .	45
Figura 24. Mapa metabólico de la biosíntesis de ácidos grasos. Verde: <i>Methylobacterium bullatum</i> . Rojo: <i>Cellulomonas</i> sp. Y8. Azul: <i>Streptomyces</i> sp. Verde: <i>Sphingomonas</i> sp.	46
Figura 25. Mapa metabólico de la biosíntesis de lípidos neutros. Verde: <i>Scenedesmus acutus</i> .	47
Figura 26. Mapa metabólico de la biosíntesis de lípidos neutros. Verde: <i>Stutzerimonas stutzeri</i> . Rojo: <i>Roseomonas fluvialis</i> . Azul: <i>Sphingomonas koreensis</i> .	47
Figura 27. Mapa metabólico de la biosíntesis de lípidos neutros. Verde: <i>Sphingopyxis</i> sp. Rojo: <i>Flavobacterium</i> sp. Azul: <i>Nordella</i> sp HKS 07.	48
Figura 28. Mapa metabólico de la biosíntesis de lípidos neutros. Verde: <i>Enhydrobacter</i> sp. Rojo: <i>Dyadobacter fermentans</i> DSM 18053.	48
Figura 29. Mapa metabólico de la captación y secuestro de metales pesados en <i>C. vulgaris</i> .	49
Figura 30. Mapa metabólico de la captación y secuestro de metales pesados en <i>S. acutus</i> .	50
Figura 31. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: <i>Pseudomonas oryzae</i> . Rojo: <i>Microvirga</i> Sp VF16.	51

Figura 32. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: <i>Chlorella vulgaris</i>	52
Figura 33. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: <i>Methylobacterium bullatum</i> . Rojo: <i>Cellulomonas</i> sp. Y8. Azul: <i>Streptomyces</i> sp. Verde: <i>Sphingomonas</i> sp.	52
Figura 34. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: <i>Scenedesmus acutus</i>	53
Figura 35. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: <i>Stutzerimonas stutzeri</i> . Rojo: <i>Roseomonas fluvialis</i> . Azul: <i>Sphingomonas koreensis</i>	54
Figura 36. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: <i>Sphingopyxis</i> sp. Rojo: <i>Flavobacterium</i> sp. Azul: <i>Nordella</i> sp HKS 07.	54
Figura 37. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: <i>Enhydrobacter</i> sp. Rojo: <i>Dyadobacter fermentans</i> DSM 18053.	55
Figura 38. Red bipartita del metagenoma de <i>C. vulgaris</i>	57
Figura 39. Red bipartita del metagenoma de <i>S. acutus</i>	58
Figura 40. Análisis de Componentes Principales (PCA) de los transcriptomas en <i>Chlorella vulgaris</i> en los tratamientos con luz azul y roja.	59
Figura 41. Análisis de Componentes Principales (PCA) de los transcriptomas en <i>Scenedesmus acutus</i> en los tratamientos con luz azul y roja.	60
Figura 42. Heatmap con los genes con mayor variabilidad en <i>C.vulgaris</i>	61
Figura 43. Heatmap con los genes con mayor variabilidad en <i>S. acutus</i>	62
Figura 44. Expresión de genes en la ruta de Resistencia a β -lactámicos en el metagenoma de <i>C. vulgaris</i> bajo luz blanca (control), azul y roja.	63
Figura 45. Coexpresión de genes asociados a la resistencia de β -lactámicos bajo luz azul y roja en el metagenoma de <i>C. vulgaris</i>	64
Figura 46. Expresión de genes en la ruta de Resistencia a β -lactámicos en el metagenoma de <i>S. acutus</i> bajo luz blanca (control), azul y roja.	65
Figura 47. Coexpresión de genes asociados a la resistencia de β -lactámicos bajo luz azul y roja en el metagenoma de <i>S. acutus</i>	66

Figura 48. Expresión de genes en la ruta de biosíntesis de ácidos grasos en el metagenoma de <i>C. vulgaris</i> bajo luz blanca (control), azul y roja.	67
Figura 49. Coexpresión de genes asociados a la biosíntesis de ácidos grasos bajo luz azul y roja en el metagenoma de <i>C. vulgaris</i>	68
Figura 50. Expresión de genes en la ruta de biosíntesis de ácidos grasos en el metagenoma de <i>S. acutus</i> bajo luz blanca (control), azul y roja.	69
Figura 51. Coexpresión de genes asociados a la biosíntesis de ácidos grasos bajo luz azul y roja en el metagenoma de <i>S. acutus</i>	69
Figura 52. Gráfico de importancia en el modelamiento con Random Forest de las diferentes rutas metabólicas en el metagenoma de <i>C. vulgaris</i>	77
Figura 53. Gráfico de valores observados contra valores predichos por el modelo en el análisis en el metagenoma de <i>C. vulgaris</i>	78
Figura 54. Gráfico de importancia en el modelamiento con Random Forest de las diferentes rutas metabólicas en el metagenoma de <i>S. acutus</i>	78
Figura 55. Gráfico de valores observados contra valores predichos por el modelo en el análisis en el metagenoma de <i>S. acutus</i>	79

Índice de tablas.

Tabla 1. Características del ensamblado del metagenoma de <i>C. vulgaris</i>	14
Tabla 2. Características del ensamblado del metagenoma de <i>S. acutus</i>	15
Tabla 3. Parámetros de calidad de cada ensamble donde N50 y N90 es la longitud de contigs con la cual se completa el 50% y 90% del ensamblado respectivamente. L50 y L90 es el número de contigs con la cual se completó el 50% y 90% del ensamblado. auN (área under Nx curve) representa todos los valores de N50 hasta N100 en un ensamble.	15
Tabla 4. Bacterias más representativas de cada metagenoma.....	18
Tabla 5. Bins clasificados taxonómicamente en el metagenoma de <i>C. vulgaris</i> . ..	21
Tabla 6. Bins clasificados taxonómicamente en el metagenoma de <i>S. acutus</i>	21
Tabla 7. Genes de resistencia a antibióticos detectados en el metagenoma de <i>C. vulgaris</i>	22
<i>Tabla 8. Genes de resistencia a antibióticos detectados en el metagenoma de S. acutus</i>	23
Tabla 9. Proteínas anotadas e hipotéticas de cada metagenoma.....	32
<i>Tabla 10. Principales hits de proteínas implicados en la biorremediación de metales pesados en el metagenoma de C. vulgaris</i>	41
Tabla 11. Principales hits de proteínas implicados en la biorremediación de fármacos en el metagenoma de <i>C. vulgaris</i>	42
Tabla 12. Principales hits de proteínas implicados en la biorremediación de metales pesados en el metagenoma de <i>S. acutus</i>	42
Tabla 13. Principales hits de proteínas implicados en la biorremediación de fármacos en el metagenoma de <i>S. acutus</i>	43
Tabla 14. Matriz de puntuación de funcionalidad en el metagenoma de <i>C. vulgaris</i>	72
Tabla 15. Matriz de puntuación de funcionalidad en el metagenoma de <i>S. acutus</i>	73
Tabla 16. Métricas de centralidad para cada MAG en el metagenoma de <i>C. vulgaris</i>	75

Tabla 17. Métricas de centralidad para cada MAG en el metagenoma de <i>S. acutus</i> .	
.....	76

1. Introducción.

El desarrollo de la genómica moderna comenzó con la implementación del método de secuenciación de Sanger durante la década de los 1970. Este avance impulsó los primeros proyectos de secuenciación a gran escala, incluyendo el Proyecto Genoma Humano. Asimismo, permitió el estudio de organismos modelo como *Escherichia coli* y *Saccharomyces cerevisiae*, esta último considerado el primer organismo eucariota cuyo genoma fue secuenciado por completo. Estos avances permitieron no sólo explorar secuencias nucleotídicas, sino también comprender su funcionalidad mediante la identificación de genes, regiones promotoras y el aprovechamiento de estas, sentando las bases de la biología molecular y la ingeniería genética modernas.

El avance continuo de las tecnologías de secuenciación ha permitido desarrollar plataformas con mayor capacidad, precisión y rendimiento. Entre ellas, la tecnología Illumina se ha consolidado como una de las más utilizadas debido a su elevada fidelidad y capacidad para generar millones de lecturas en una sola corrida de secuenciación. Estas características facilitan la secuenciación de genomas completos y el estudio de comunidades microbianas complejas, al reducir tanto los tiempos de procesamiento como los costos, por lo tanto, son deseables para la secuenciación de genomas pertenecientes a organismos con una carga genética considerable a un menor tiempo y costos razonables.

La tecnología de secuenciación de nueva generación (next- generation sequencing o NGS) comprende una serie de técnicas capaces de generar de manera paralela y de forma masiva millones de fragmentos de ácido desoxirribonucleico (ADN) en un único proceso de secuenciación, optimizando el rendimiento y los costos.

En términos generales, el procedimiento inicia con una fragmentación del ADN asistida por ultrasonificación, rompiendo las cadenas de esta en fragmentos pequeños de 100 a 300 pb. Una vez finalizada la fragmentación, se procede a la ligación de adaptadores a los fragmentos de ADN. Estos adaptadores son

esenciales para el proceso de secuenciación y permiten identificar los fragmentos. El conjunto de fragmentos de ADN ligados a adaptadores constituye una biblioteca de ADN. Una vez preparadas las librerías, comienza la hibridación de estas a la superficie sólida donde se llevará a cabo la amplificación; uno de los adaptadores contiene extremos compatibles con sondas hibridadas covalentemente al material donde se llevará a cabo la amplificación (superficie del flow cell), en este paso comienzan ciclos de amplificación por *puentes*. Su nombre es debido a que los adaptadores crearán un *punte* entre sondas hibridadas covalentemente en el material, formando un *arco* y exponiendo el fragmento de ADN a amplificar, una vez obtenidos los amplicones, estos se hibridarán con otros adaptadores y repitiendo el mismo proceso hasta obtener clusters de millones de copias de un mismo fragmento. Finalizada la amplificación, los amplicones que conforman los clusters son sometidas a una secuenciación por síntesis, en donde se incorporan nucleótidos marcados con fluoróforos que emiten señales detectables para el secuenciador, de tal manera que en función del tipo de señal detectado será el nucleótido correspondiente a la secuencia. Por último, estas secuencias se alinean y ensamblan entre sí para formar *contigs*, es decir, secuencias largas y continuas tal como se observa en la Figura 1.

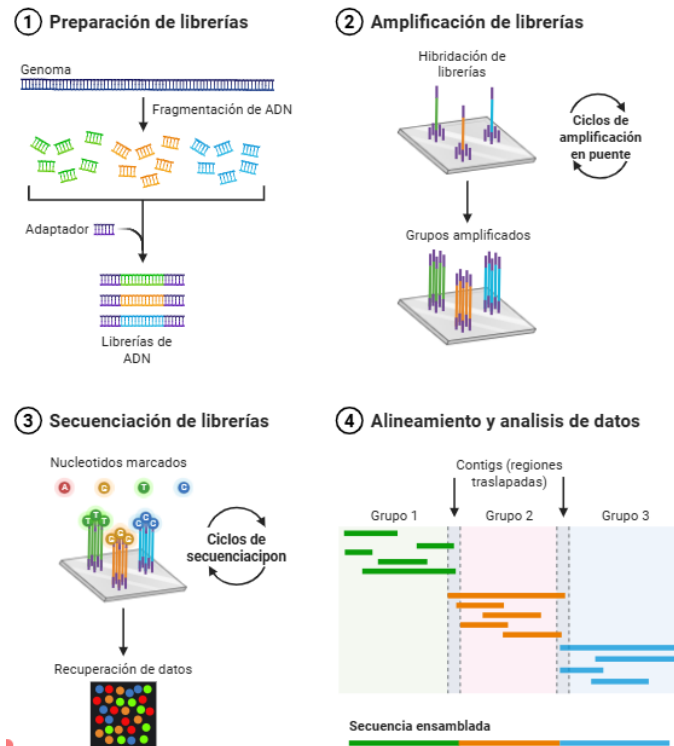


Figura 1. Proceso de secuenciación masiva de última generación Illumina.

2. Antecedentes.

2.1. Consorcios de microorganismos

El incremento en la resolución y sensibilidad de las tecnologías de secuenciación permitió la detección y caracterización simultánea múltiples genomas bacterianos inmersos en una misma matriz, incluso cuando algunos de estos se encuentren en bajas abundancias. En condiciones naturales, estos microorganismos no se encuentran aislados entre sí. Por el contrario, establecen interacciones con otros microorganismos que pueden contribuir a su supervivencia (Chen et al., 2020).

Este avance favoreció el estudio de comunidades microbianas completas y condujo al desarrollo de la metagenómica. Según el Instituto Nacional de Investigación Humana, el término “metagenoma” constituye a la estructura y función de todas las secuencias de nucleótidos aisladas y analizadas de una matriz en donde convergen y conviven varios microorganismos. A su vez, surge un nuevo campo de estudio relacionado con la genómica, denominado “metagenómica”, la cual, se dedica

analizar mediante recursos moleculares e informáticos las secuencias de ADN de un consorcio de algún microbioma a estudiar (*Metagenómica*, s. f.-b).

Un consorcio microbiano es una comunidad de microorganismos que coexisten en un mismo nicho ecológico y establecen interacciones que pueden ser de naturaleza mutualista, cooperativa, comensalista o antagonista, entre otras. Regularmente estas interacciones pueden estar mediadas por la producción e intercambio de metabolitos, la competencia por nutrientes o la producción de compuestos que modifican el ambiente con el fin de contribuir a la supervivencia, adaptación y estabilidad de un consorcio. Un ejemplo ampliamente estudiado de este tipo de interacción se encuentra entre los hongos micorrízicos del filo *Glomeromycota* y diversas especies árboles, donde el hongo incrementa la disponibilidad y adquisición de fósforo para la planta mediante la movilización de formas orgánicas e inorgánicas del nutriente, a cambio de compuestos carbonados derivados de la fotosíntesis de la planta (Chen et al., 2020).

Estos consorcios no solamente son propios de los suelos o de potenciación en la absorción de nutrientes para algunos productos agrícolas de interés, también son de interés ecológico debido a su potencial carácter remediador en cuerpos de agua. Por ejemplo, hay consorcios bacterianos para la eliminación de nitrógeno en altas concentraciones mediante procesos de nitrificación y desnitrificación usando bacterias como *Nitrosomas*, *Nitrobacter* y *Pseudomonas* (Cao et al., 2022). Los mismos corales en colaboración con algunas especies de cianobacterias (*Prochlorococcus*, *Synechococcus*) constituyen uno de los sistemas fototróficos más importantes en los mares a nivel de producción de oxígeno (Dong et al., 2022).

Estas características de fijación y aumento de biodisponibilidad de nutrientes hacen a estos consorcios sumamente atractivos para la biorremediación; tal es el ejemplo de algunas especies de bacterias como *Alcanivorax*, *Pseudomonas* y *Rhodococcus* capaces de degradar hidrocarburos (Juárez et al., 2023). Estas características resaltan el potencial biotecnológico de dichas comunidades para la recuperación de ambientes contaminados.

2.1.1. Consorcios microalga-bacteria frente a contaminantes de interés.

Entre los diferentes tipos de consorcios microbianos, las asociaciones establecidas entre microalgas y bacterias han despertado interés creciente debido a su capacidad para favorecer procesos de biorremediación y producción de metabolitos de interés biotecnológico. La presencia creciente de contaminantes constituye un problema ambiental relevante debido a sus efectos sobre los ecosistemas terrestres y acuáticos, que incluyen desde alteraciones de las comunidades biológicas hasta la acumulación de metales pesados en las cadenas tróficas (Ali & Khan, 2018). En este contexto, ciertos desechos que antes se pensaban como *inocuos* o que al ser descartados de manera normal por la baja que concentración que representarían en cuerpos acuáticos, pasaron a ser de importancia ecológica, a los cuales denominaron como *contaminantes emergentes*, estos tienen la característica de no estar regulados por alguna normativa federal, siendo que su diseminación se ha descrito en varios lugares del mundo, convirtiéndose en un factor de aumento a problemáticas graves como lo es la resistencia a antibióticos (Deblonde et al., 2011).

Se ha observado ampliamente como microalgas del género *Chlorophyta* son capaces de la degradación y/o biotransformación de antibióticos en presencia de ciertas bacterias. Estudios recientes apuntan que estas microalgas en simbiosis con bacterias del género *Silanimonas* son capaces de degradar antibióticos de amplio espectro como lo es la ceftriaxona en agua residual aún estando bajo concentraciones de nitrógeno y fósforo (Zaytsev et al., 2024).

Uno de los ejemplos más relevantes es la interacción entre microalgas y bacterias, donde se ha observado que ciertas bacterias pueden incrementar la biomasa algal, en algunos estudios se ha demostrado que la producción de biomasa de algunas microalgas se ve beneficiada por la presencia de algunas bacterias. También la remoción de ciertos contaminantes tales como nitrógeno (N), fósforo (P), carbono (C) y algunos metales pesados como el plomo (Pb), el níquel (Ni) y el cobre (Cu) son optimizados por este tipo de consorcios (Aditya et al., 2022).

Aunado a estos estudios, Peng y colaboradores en el 2021 reportaron que la bacteria *Azospirillum brasilense* reducía el estrés oxidativo de la microalga *Chlorella*

sorokiniana ante diferentes tipos de estrés como las altas concentraciones de salinidad y concentraciones altas de metales pesados. Recientemente en el 2024, Cai y colaboradores generaron un consorcio entre las bacterias *Rhodococcus erythropolis* y *Cupriavidus metallidurans* junto con la microalga *Coelastrella terrestris* y observaron un aumento en la producción de biomasa, así como la remoción de benceno, tolueno y fenol, llegando a porcentajes de remoción de 95.75% 92.7% y 100% respectivamente a los 7 días de crecimiento.

La interacción entre microalgas y bacterias en consorcios simbólicos es una de las estrategias más prometedoras para el desarrollo de soluciones biotecnológicas con enfoque a la biorremediación y compuestos con interés industrial. Las microalgas *Chlorella vulgaris* y *Scenedesmus acutus* han demostrado ser una buena opción en la remoción de contaminantes y síntesis de metabolitos valiosos con capacidad de ser potenciados significativamente en la presencia de ciertos géneros de bacterias. Estos consorcios microbianos se presentan de forma natural en cuerpos de agua y constituyen un sistema complejo donde las interacciones aún no se comprenden completamente. En este contexto, la integración de herramientas de secuenciación masiva, análisis bioinformáticos e inteligencia artificial representa una estrategia prometedora para estudiar sistemas biológicos, identificar rutas metabólicas compartidas, inferir relaciones funcionales y comprender los procesos moleculares que sustentan estas asociaciones, siendo importantes para encontrar herramientas innovadoras para la biorremediación de contaminantes tradicionales como los metales pesados e hidrocarburos y para contaminantes emergentes como fármacos y compuestos orgánicos de uso médico.

2.2. Bioinformática e inteligencia artificial.

La bioinformática es una disciplina que integra herramientas computacionales para el análisis e interpretación de datos biológicos, particularmente aquellos generados mediante tecnologías de secuenciación masiva. Para ello, emplea recursos y metodologías provenientes de áreas como la programación, el desarrollo de software, el análisis estadístico y las bases de datos, permitiendo procesar y analizar grandes volúmenes de información genómica, transcriptómica, proteómica

y metagenómica. Estos datos se encuentran disponibles en repositorios públicos como el Centro Nacional de Información Biotecnológica (NCBI), que proporciona acceso a bases de datos y recursos para el almacenamiento, consulta y análisis de información biológica.

Uno de los retos más complejos de esta rama es el volumen y la complejidad de datos que se obtienen después de una secuenciación y estos datos crudos requieren flujos de trabajo robustos para realizar tareas como la limpieza de secuencias, el ensamblaje del metagenoma, anotación de genes, clasificación taxonómica y su predicción funcional. Todos estos resultados del tratamiento de los datos crudos no solo permiten identificar la diversidad microbiana presente, sino también inferir su potencial metabólico y ecológico. De esta manera, la bioinformática no es solamente un recurso auxiliar, sino el motor analítico que hace posible la exploración y el análisis de comunidades microbianas complejas (Jovel et al., 2016).

La bioinformática ha sido esencial para el procesamiento de los grandes volúmenes de datos generados por la secuenciación masiva, estos conjuntos de datos pueden alcanzar un espacio de varios cientos de gigabytes incluso algunos cuantos terabytes de solamente datos crudos. Herramientas como ensambladores, anotadores funcionales, clasificadores taxonómicos y bases de datos de la Kyoto Encyclopedia of Genes and Genomes (KEGG) permiten transformar estos datos crudos en conocimiento biológico significativo (Stephens et al., 2015). Sin embargo, el procesamiento de estos datos puede requerir un tiempo considerable, pudiendo demorarse horas y días de procesamiento continuo inclusive en sistemas computacionales avanzados con altas prestaciones. En este contexto, la inteligencia artificial ofrece un enfoque complementario y altamente escalable, mediante el uso de algoritmos de machine learning es posible automatizar la detección de patrones, inferir relaciones funcionales y acelerar tareas de clasificación y predicción en menores periodos de tiempo (Eraslan et al., 2019).

La inteligencia artificial (IA) comprende un conjunto de métodos orientados al desarrollo de sistemas capaces de realizar tareas que requieren capacidades

asociadas con el razonamiento, el aprendizaje y la toma de decisiones. Dentro de la IA se encuentra el aprendizaje automático (machine learning), mediante el cual los algoritmos pueden identificar patrones y relaciones a partir de datos sin requerir que todas las reglas de decisión sean definidas explícitamente. Entre los algoritmos de aprendizaje automático se encuentra Random Forest, un método de aprendizaje supervisado que combina múltiples árboles de decisión para realizar tareas de clasificación o regresión y que puede ser utilizado para identificar las variables con mayor contribución a la predicción de una respuesta determinada (Schäferling, 2023).

En los últimos años, la integración de algoritmos propios de IA ha resultado ser de gran utilidad en sistemas complejos, desde planteamientos financieros y económicos, hasta sistemas de detecciones de enfermedades mediante el procesamiento de muestras provenientes de imagenología donde, el análisis de estos datos y el limitado recurso humano en estas áreas tan especializadas suelen ser un importante factor para el diagnóstico ya sea tardío (Behera et al., 2023). No obstante, también las ciencias ómicas se han beneficiado de estos algoritmos, en el ámbito de la bioinformática, dichos algoritmos se han empleado como herramientas complementarias para tareas como la clasificación taxonómica, predicción funcional de genes y proteínas e, inclusive, la identificación de patrones en grandes conjuntos biológicos.

La integración de técnicas bioinformáticas con algoritmos de IA ofrece una oportunidad sin precedentes para analizar grandes volúmenes de datos y la automatización de ciertos procesos los cuales implican un gran tiempo de procesamiento y recursos informáticos amplios. La bioinformática y la IA son dos herramientas que están relacionadas; al tener secuencias genómicas, las cuales codifican a una proteína, es sencillo otorgarles una identidad. No obstante, al tener una gran cantidad de variaciones *per se* en cada gen y según cada organismo, el clasificarlo da como resultado un trabajo extenuante, más aún si este involucra organismos eucariotas o consorcios complejos de microorganismos (Camacho et al., 2018). En este ámbito, los algoritmos de machine learning permiten analizar

múltiples variables simultáneamente e identificar relaciones o características relevantes que pueden resultar difíciles de detectar mediante aproximaciones convencionales. Estos algoritmos han potenciado el estudio de consorcios microbianos donde, la elevada dimensionalidad y complejidad de la información dificultan su interpretación. Bajo este punto, se presenta este trabajo con el objetivo de analizar metagenomas complejos aplicando herramientas bioinformáticas y de machine learning de manera exploratoria, destacando el flujo de trabajo utilizado para llegar a un análisis con estas nuevas herramientas.

2.3. *Chlorella vulgaris* y *Scenedesmus acutus* como candidatos en la remoción de contaminantes emergentes y metales pesados.

Chlorella vulgaris es una microalga verde del filo *Chlorophyta* de la familia *Chlorellaceae*, este organismo está presente en la mayoría de los cuerpos de agua, siendo un bioindicador de ecosistemas acuáticos sanos (Singh et al., 2025). Por otro lado, *Scenedesmus acutus* es una microalga verde del filo *Chlorophyta* pero perteneciente a la familia *Scenedesmaceae*. Sus usos biotecnológicos son bastos: desde la producción de lípidos para la síntesis verde de biodiesel (Torres, 2025), la remoción de contaminantes emergentes como fármacos y antibióticos (Ndlela et al., 2023) hasta la generación de péptidos antimicrobianos mediante la transformación génica de estas (Wang et al., 2025). Aunque ambas microalgas tengan una extensa lista de posibles usos biotecnológicos, estas no cuentan con la suficiente información genética suficiente en repositorios públicos tales como NCBI, donde apenas hay 7 genomas de *C. vulgaris* reportados y uno usado como referencia, por el contrario, en *S. acutus* no existe un genoma reportado en dicho repositorio.

En este contexto, el presente proyecto tiene como objetivo caracterizar el metagenoma asociado a *Chlorella vulgaris* y *Scenedesmus acutus*; las cuales fueron aisladas de una planta tratadora de aguas residuales. La secuenciación genómica de las microalgas *Chlorella vulgaris* y *Scenedesmus acutus* fueron obtenidos por colaboradores del grupo de trabajo interno (Gómez Torres, 2025). A partir del procesamiento de estos datos, se analizaron funciones relacionadas con las posibles interacciones microalga-bacteria que pudieran desarrollarse:

biosíntesis de biotina y riboflavina, adquisición de hierro mediante sideróforos y los mecanismos de comunicación bacteriana mediados por quorum sensing. En este enfoque, se integraron y compararon datos de ARN secuenciados por RNA-seq, obtenidos en el mismo grupo de trabajo (Torres, 2025) con el fin de complementar la caracterización del potencial funcional de las comunidades bacterianas y evaluar su posible contribución a procesos asociados con la biorremediación.

3. Justificación.

El potencial que tienen los consorcios entre bacterias y microorganismos complejos, tales como los que se encuentran inmersos en los metagenomas de *C. vulgaris* y *S. acutus* son de gran importancia biotecnológica debido a su alta capacidad de remoción de contaminantes. Sin embargo, la densidad de datos que implica esta clase de análisis es extenso, demandando una capacidad de cómputo superior. Debido a esta problemática, se plantea el uso de flujos de trabajo robustos, el uso de equipo de supercómputo y algoritmos de IA para develar interacciones microalga-bacteria que potencien la capacidad metabólica, como lo es la remoción de contaminantes.

4. Hipótesis.

Las herramientas bioinformáticas y de inteligencia artificial permitirán el análisis de caracterización de los metagenomas de *Chlorella vulgaris* y *Scenedesmus acutus*, elucidando interacciones bacterianas de interés que pueden participar de manera positiva o negativa en rutas metabólicas compartidas para la remoción de contaminantes como metales pesados, biotransformación de fármacos y resistencia a antibióticos.

5. Objetivo.

Analizar el metagenoma de *Chlorella vulgaris* y *Scenedesmus acutus* a partir de datos de secuenciación masiva para la elucidación de rutas metabólicas interrelacionadas entre microalgas y bacterias.

6. Objetivos específicos.

- Diseñar y aplicar un flujo de trabajo bioinformático para el procesamiento, ensamblaje y anotación de los metagenomas asociado a *Chlorella vulgaris* y *Scenedesmus acutus*.
- Ejecutar el diseño metagenómico para identificar y clasificar las bacterias asociadas a ambas microalgas.
- Identificar genes que pueden atribuir resistencia a antibióticos, biorremediación de metales pesados y biotransformación de fármacos.
- Elucidar rutas metabólicas compartidas entre microalgas y bacterias para el mapeo funcional y la predicción de vías metabólicas relacionadas a estos contaminantes.
- Modelar las redes de co-ocurrencia y patrones de interacción microalga-bacteria mediante algoritmos de aprendizaje automático (Machine Learning) o análisis de redes complejas.

7. Materiales y métodos.

7.1. Control de calidad y preprocesamiento.

Las lecturas obtenidas mediante secuenciación masiva, fueron sometidos a un control de calidad utilizando la herramienta FastQC con el propósito de evaluar la calidad, localización de adaptadores y posibles sesgos o errores del proceso de secuenciación. Posteriormente, se realizaron cortes de las secuencias que resultaron con una baja calidad utilizando el software Trimmomatic (Babraham Bioinformatics - FastQC A Quality Control Tool For High Throughput Sequence Data, s. f.).

7.2. Ensamblaje del metagenoma.

las lecturas filtradas se ensamblaron mediante MEGAHIT, con este programa se obtuvieron contigs de longitud variable y se filtraron aquellas con longitud menor a 500 pb (Dong et al., 2022). La calidad del ensamble fue evaluada mediante el software QUAST para la validación de métricas como N50, número de contigs y el tamaño total del ensamblado (Gurevich et al., 2013).

7.3. Análisis taxonómico.

Para la determinación de la composición taxonómica de ambos metagenomas se utilizó el ensamble obtenido en conjunto con el software Kraken2, utilizando la base de datos de RefSeq (Wood et al., 2019). Posteriormente, se realizó un análisis de abundancias relativas usando las herramientas KronaTools y siendo representadas por la paquetería grafica de ggplot en Rstudio (Ondov et al., 2011).

7.4. Anotación funcional.

Para la anotación funcional en ambos metagenomas, los datos fueron tratados con el software funannotate para la identificación de genes y secuencias codificantes (Seemann, 2014). Para obtener anotaciones funciones basados en ortología, se empleó el software eggNOG-mapper, asignando funciones COG (Clusters of Orthogous Groups) y rutas metabólicas (Huerta-Cepas et al., 2018). Para el aislamiento de genomas bacterianos (binning metagenómico) en formato de Genoma Ensamblado a partir de Metagenoma (Metagenome-Assembled Genome, MAG) se usó el software MaxBin2 y el software CheckM2 para evaluar parámetros de calidad, posteriormente, con el software Prokka se anotaron estos genomas y se les asignaran categorías KO (KEGG Orthology) para la inferencia de rutas metabólicas compartidas.

De manera complementaria, se integraron datos de ARN obtenidos mediante RNA-seq de cada microalga. Estos datos metatranscriptomicos fueron alineados contra los ensamblajes metagenómicos utilizando el software Rsubread y cuantificados por otro software llamado featureCounts (Liao et al., 2013). Se normalizaron los datos en forma de TPM (transcripts per million), permitiendo validar la expresión activa de genes de interés (Liao et al., 2013b).

7.5. Inferencia de rutas metabólicas compartidas.

Una vez obtenidas las anotaciones funcionales, los genes anotados se mapearon con la base de datos KEGG usando BlastKoala para inferir rutas metabólicas compartidas entre microalgas y bacterias presentes en ambos consorcios (Kanehisa, 2000). Se identificaron rutas relacionadas a resistencia de metales pesados y síntesis de metabolitos de interés biológico (Kanehisa et al., 2015).

7.6. Análisis exploratorio mediante aprendizaje automático.

7.7. Visualización de resultados.

Los resultados fueron integrados y visualizados con el lenguaje de programación R, con paqueterías como ggplot2, Phyloseq y Vegan para demostrar la diversidad taxonomica, distribución funcional y redes de co-ocurrencia microalga-bacteria (McMurdie & Holmes, 2013). También se generaron mapas metabólicos interactivos con KEGG Mapper y se diseñaron diagramas de Sankey para visualizar las rutas metabolicas compartidas. Todos los softwares mencionados son de uso libre.

8. Resultados.

8.1. Calidad del ensamblado metagenómico.

El proceso de ensamblado *de novo* se llevó a cabo con el ensamblador MEGAHIT, especificando sensibilidad para un ensamble metagenómico. Con el software Quast en su modalidad para metagenomas (metaQuast) obteniendo los siguientes resultados.

En la Tabla 1 y 2 podemos observar las características de cada uno de los ensambles. En estos, podemos observar el número de contigs generados a partir de las lecturas de la secuenciación masiva. Se observa que el metagenoma de *S. acutus* tiene un mejor ensamblado debido a poseer menos contigs pequeños y tener un mayor número de contigs más grandes. En la Tabla 3 podemos observar valores métricos de interés tales como el estadístico N50 y N90, estos se definen como una mediana tal que el 50% o 90% del ensamble está contenido en contigs que son iguales a ese valor. Las métricas L50 y L90 están relacionadas con las anteriores, estas representan el número de contigs que constituyen el 50% y el 90% del ensamble. Para este caso, podemos observar que las métricas de *S. acutus* son superiores con respecto a las de *C. vulgaris*, teniendo contigs más grandes en los valores de N50 y completando su ensamblaje al 90% con un menor número de contigs, además de que su contig más largo es el de mayor tamaño entre ambos ensambles.

Tabla 1. Características del ensamblado del metagenoma de *C. vulgaris*

Tabla 2. Características del ensamblado del metagenoma de *S. acutus*.

Tabla 3. Parámetros de calidad de cada ensamble donde N50 y N90 es la longitud de contigs con la cual se completa el 50% y 90% del ensamblado respectivamente. L50 y L90 es el número de contigs con la cual se completó el 50% y 90% del ensamblado. auN (área under Nx curve) representa todos los valores de N50 hasta N100 en un ensamble.

8.2. Análisis taxonómico.

En ambas Figuras (2 y 3) se observa una amplia gama de microorganismos; desde bacterias, arqueas y algunos virus, esto denota la capacidad de resolución de la tecnología de secuenciación. Sin embargo, podemos observar que, tanto para *C. vulgaris* y *S. acutus* no hay una clasificación hacia sus especies, en el caso de *C. vulgaris* se le clasifica como *Chlamydomonas reinhardtii* y para *S. acutus* como *Parachlorella kessleri*. Esto es debido a las bases de datos utilizadas para la clasificación taxonómica, si bien esta base de datos es una de las completas a disposición, estas no cuentan con la información necesaria para clasificar organismos eucariontes como lo es *S. acutus*, en el cual, no existe un genoma de referencia en NCBI para realizar una comparación más clara y dirigida.

Figura 2. Análisis taxonómico (diagrama de Sankey) del metagenoma en C. vulgaris hecho con el software Kraken2 y Pavian, ordenado desde dominio hasta especie.

Figura 3. Análisis taxonómico (diagrama de Sankey) del metagenoma en S. acutus hecho con el software Kraken2 y Pavian, ordenado desde dominio hasta especie.

No obstante, estos resultados son relevantes debido a la cantidad de bacterias que existen en ambos metagenomas. Para ambos metagenomas, se tomaron en consideración solo las bacterias que tuvieran un número de lecturas mayores a 100,000 pares de bases (bp).

En la Tabla 4 podemos observar las bacterias más representativas en cada metagenoma, siendo las más representativas en *C. vulgaris* la bacteria *Pseudomonas oryzae* el cual es un bacilo gram negativo de carácter saprófito, se ha observado que en co-cultivo con microalgas ayuda a la absorción de grandes cantidades de nitrógeno en ausencia de carbono (Tian et al., 2025), también esta bacteria común observarla en relaciones simbióticas entre microalgas y bacterias (Ashraf et al., 2022). En caso de *S. acutus* la bacteria con mayor

abundancia relativa fue *Stutzerimonas stutzeri* la cual es un bacilo gram negativo de carácter saprófito, el cual es un gran objetivo de interés biotecnológico ya que esta potencia la remoción de fármacos como el acetaminofén y diclofenaco (Vargas-Ordóñez et al., 2023).

Tabla 4. Bacterias más representativas de cada metagenoma.

8.2.1. Recuperación de genomas bacterianos (binning metagenómico).

MaxBin2 se utilizó para recuperar bins a partir de los ensamblajes metagenómicos. Este software es capaz de detectar y aislar potenciales genomas de bacterias de un metagenoma, la completitud y contaminación se evaluaron mediante el software CheckM2. Se recuperaron los genomas (bins) que obtuvieran valores mayores a 90% de completitud y menores a 20% de contaminación. Estos valores son

deseables debido a que estos pueden ser aportaciones de gran valor científico en bases de datos publicas como el NCBI.

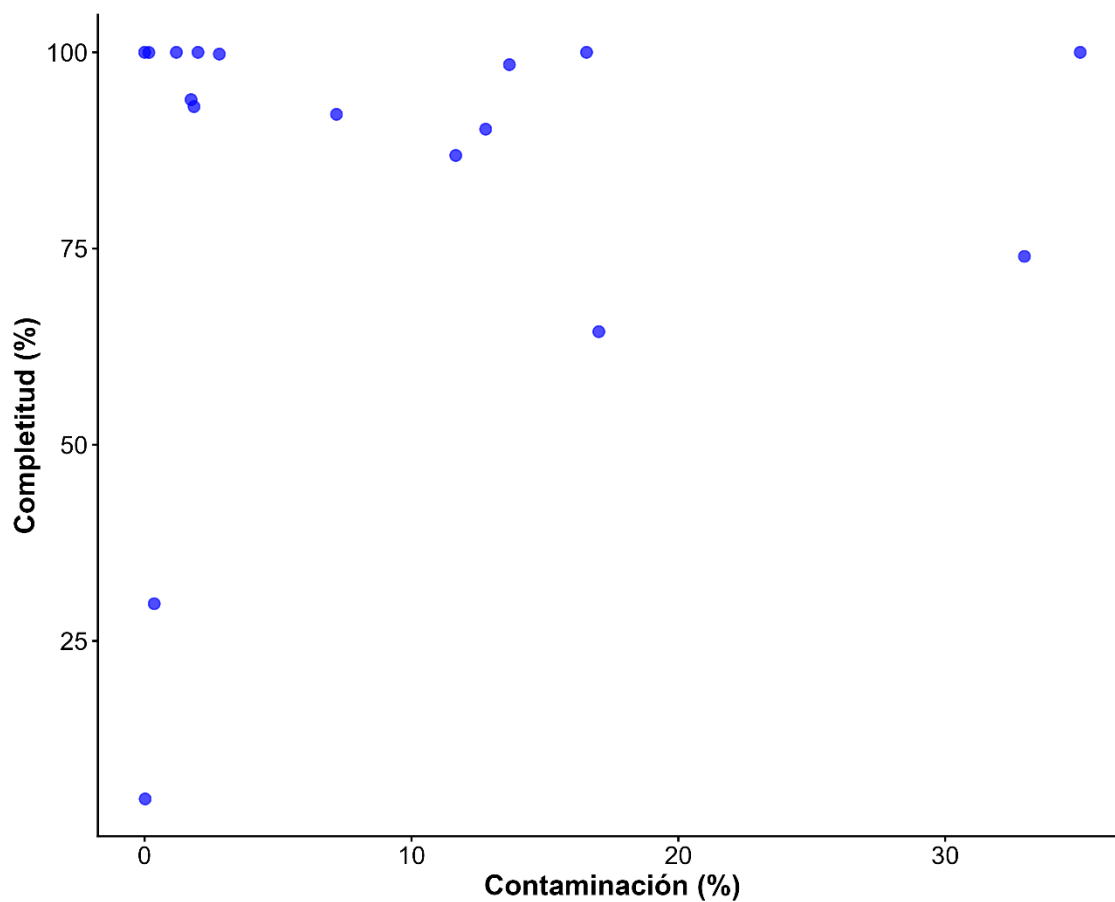


Figura 4. Bins recuperados del metagenoma de *C. vulgaris*.

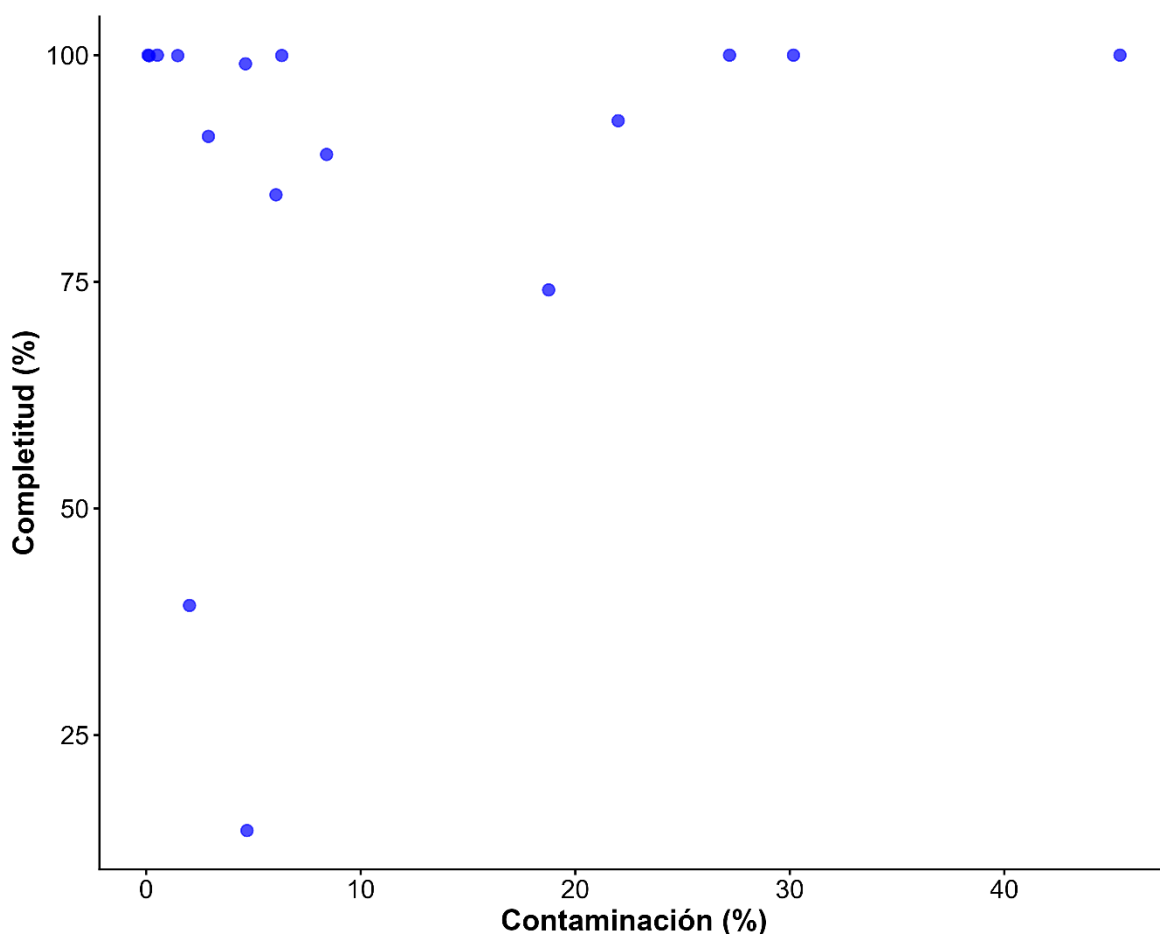


Figura 5. MAGs recuperados del metagenoma de *S. acutus*.

Se recuperaron los bins de origen bacteriano de ambos metagenomas y se seleccionaron aquellos que cumplían con los criterios de calidad antes mencionados. Se observan una amplia variedad de bins bacterianos que corresponden al metagenoma de *C. vulgaris* (Figura 4) al igual que en el metagenoma de *S. acutus* (Figura 6), en ambos casos, se observan genomas con porcentajes de completitud elevadas y porcentajes de contaminación que cumplen con los parámetros establecidos, no obstante, también existen genomas con porcentajes de contaminación elevadas o incompletos, descartando estos. Se realizó su clasificación taxonómica con el software Kraken2 usando la base de datos Refseq (2024) y se anotó su genoma con el software Prokka. Para finalizar, con los genes anotados se inició una búsqueda masiva de genes implicados en la resistencia de antibióticos usando el software ABRicate con las bases de datos Card y Resfinder, obteniendo las Tablas 5, 6, 7 y 8.

Tabla 5. Bins clasificados taxonómicamente en el metagenoma de C. vulgaris.

Tabla 6. Bins clasificados taxonómicamente en el metagenoma de S. acutus.

Tabla 7. Genes de resistencia a antibióticos detectados en el metagenoma de C. vulgaris.

Tabla 8. Genes de resistencia a antibióticos detectados en el metagenoma de S. acutus.

En las Tablas 7 y 8 podemos visualizar una amplia gama de genes implicados en la resistencia a antibióticos (concretamente a los β -lactámicos) y a algunos péptidos antimicrobianos catiónicos lo cual converge en consorcio microbiano sumamente capacitado para degradar antibióticos. Este tipo de genes suelen aparecer debido a una presión selectiva ambiental, ya sea por una presencia de estos antimicrobianos o por la alta concentración de metales pesados presentes en su medio natural.

Una vez identificados posibles los genes de resistencia a antibióticos es necesario determinar los resistomas presentes en ambos metagenomas. Un resistoma corresponde al conjunto de genes relacionados con resistencia antimicrobiana presentes en un genoma o comunidad microbiana (Biocodex Microbiota Institute, 2025). Estos resistomas son identificados con un software llamado RGI (Resistance Gen Identifier) el cual usa un compendio de bases de datos con los cuales identificar

y develar secuencias de proteínas de las enzimas implicadas en este proceso (Kwatra, 2021).

Figura 6. Resistomas encontrados en C. vulgaris.

Figura 7. Resistomas encontrados en S. acutus.

Las Figuras 6 y 7 fueron generadas a partir del software RGI (Resistance Gene Identifier), dichos mapas de calor (heatmaps) muestran la variedad de genes que, en conjunto, conforman a los resistomas de cada metagenoma, el color verde indica un valor conservador en los conteos de cada gen, el color amarillo indica una presencia mayor a comparación de los demás genes. En el caso de *S. acutus*, se identificaron genes que le otorgan resistencia a los aminoglicosidos, diaminoprimidinas, fluoroquinolonas, glucopéptidos, fenicoles y tetracilinas. Por otra parte, en el metagenoma de *C. vulgaris*, aparte de que se identificaron los mismos genes de resistencia de *S. acutus*, destacan genes capaces de degradar antibióticos de diferentes grupos como las cefalosporinas, monobetalactámicos penicilinas y fenicoles; aunque este último es un quelante para metales, también posee actividad antimicrobiana (Goda et al., 2025). En conjunto, el metagenoma de *C. vulgaris* podría presentar una mayor capacidad para tolerar espectro más amplio de antibióticos en comparación con el metagenoma de *S. acutus*. No obstante, estas observaciones deben validarse bajo experimentos controlados, evaluando efectos de estos antibióticos en comparación con cepas puras.

8.2.2.1. Genes de transferencia horizontal.

Con el fin de determinar si estos genes encontrados pueden ser transferidos de bacterias a otras, se determinó realizar ensayos BLAST con una base de datos conocida como “MOB-Suite”, la cual, contiene varias secuencias de nucleótidos relacionados con proteínas de transferencia horizontal (Robertson, 2023). Mediante la herramienta Diamond BLASTp con la base de datos ya antes mencionada, se obtuvieron hits positivos, se depuraron los datos de hits con un bitscore mayores a 100% y con un porcentaje de identidad mínimo de 70%, obteniendo los siguientes resultados.

Figura 8. Genes de transferencia horizontal encontrados en el metagenoma de C. vulgaris.

Figura 9. Genes de transferencia horizontal encontrados en el metagenoma de S. acutus.

En las Figuras 8 y 9, podemos observar cómo cada uno de los genomas aislados (BINs) comparten proteínas de transferencia horizontal, concretamente “MOBP”. Este, no es más que una familia de proteínas llamadas “relaxasas” que están muy relacionadas con la transferencia de genes de bacteria a bacteria. Estas proteínas son enzimas que inician el proceso de transferencia conjugativa de ADN y son piezas centrales del moviloma bacteriano (Guzmán-Herrador & Llosa, 2019). Estas son fundamentales para la transferencia horizontal, que existan en varios genomas bacterianos de ambos metagenomas denota una transferencia activa de genes y posibles plásmidos entre bacterias.

8.2.3. Análisis de potenciales genomas nucleares.

Las Tablas 5 y 6 muestran que los metagenomas de *C. vulgaris* y *S. acutus* contienen diversos genomas (bins) correspondientes a genomas bacterianos. No obstante, entre ellos destacan el bin005 y el bin003, los cuales se les considera como candidatos a representar los genomas nucleares de *C. vulgaris* y *S. acutus*, respectivamente. Esta interpretación se sustenta debido a que los pesos de estos bins son significativamente mayores en comparación a los potenciales bins bacterianos, además, dichos valores de son concordantes con tamaños de genomas reportados en microalgas Para evaluar esta posibilidad, se utilizó BUSCO para estimar la completitud de los bins candidatos y determinar el número de ortólogos conservados. Los resultados se muestran en la Figura 10.

Figura 10. Análisis de completitud BUSCO.

Figura 11. Estadísticas de los genomas nucleares.

8.2.4. Análisis ANI para la confirmación de genomas nucleares de *C. vulgaris* y *S. acutus*.

Los genomas encontrados y aislados con Maxbin se tienen que evaluar con genomas reportados en otras bases de datos para corroborar su identidad. El análisis de identidad nucleotídica (ANI) es un método bioinformático el cual compara por ortología de genes la similitud entre 2 o más genomas. En este caso, se utilizó el software Pyani en su modalidad de “ANIm” en conjunto con varios genomas reportados como genomas de referencia.

Figura 12. Porcentaje De identidad obtenido con el análisis ANI comparando diferentes cepas de ambas microalgas (C. vulgaris y S. acutus), de igual manera, la misma comparación con el genoma de Chlamydomonas reinhardtii como ancestro común.

No obstante, también es necesario corroborar el porcentaje de cobertura. El porcentaje de cobertura en un análisis ANI representa la proporción del genoma que logró alinearse con el genoma evaluado con el cálculo de identidad nucleotídica. Este valor indica el alcance real de la comparación de identidad nucleotídica, es decir, que fracción del genoma participó realmente en el análisis.

Figura 13. Porcentaje de cobertura obtenido con el análisis ANI con los mismos genomas antes mencionados.

8.3. Anotación funcional

Para la anotación funcional se utilizó el software Prokka, posteriormente, se utilizó el software EggNOG para otorgarle una funcionalidad a las secuencias de proteínas anotadas basadas en ortología. En la tabla 9 se muestra la cantidad de proteínas anotadas con dicho software.

Tabla 9. Proteínas anotadas e hipotéticas de cada metagenoma.

Microalga	Proteínas anotadas	Proteínas hipotéticas
<i>Chlorella vulgaris</i>	29204	175884
<i>Scenedesmus acutus</i>	36862	129879

Podemos apreciar en las Figuras 14 y 15 las distribuciones de las funciones de proteínas anotadas presentes en cada metagenoma, proporcionando un visión general desde las funciones metabólicas de cada una. En ambos casos existen un porcentaje importante de proteínas con una función desconocida, esta tendencia refleja lo poco caracterizados (o exclusivos) que suelen ser este tipo de consorcios de microorganismos, sugiriendo una diversidad funcional sin explorar.

Al comparar ambos metagenomas, también observamos un comportamiento diferencial en la distribución de categorías funcionales, por ejemplo, en la transducción de señales, en el metagenoma de *C. vulgaris* representa el 22.33% de toda la anotación, mientras que en el metagenoma de *S. acutus* solo representa el 11.17%. Sin embargo, en la producción y conversión de energía, el metagenoma de *S. acutus* tiene una mayor cantidad de proteínas anotadas (13.17%) a comparación del metagenoma de *C. vulgaris*, esto demuestra el nivel de especialización entre ambos metagenomas y como estos se adaptan según las necesidades de cada

comunidad. No obstante, las funciones dominantes en ambos metagenomas son las metabólicas, desde el metabolismo y transporte de aminoácidos, lípidos, aniones metálicos y carbohidratos, hasta proteínas relacionadas con la transcripción y modificaciones post traducciones.

Figura 14. Distribución de categorías funcionales COG en el metagenoma de C. vulgaris.

Figura 15. Distribución de categorías funcionales COG en el metagenoma de S. acutus.

8.3.1. Asignación funcional de proteínas hipotéticas y validación funcional mediante alineamientos con BLASTp.

Con la finalidad de otorgar una funcionalidad a las proteínas hipotéticas de ambos metagenomas se utilizó el software Diamond BLASTp con la base de datos NR, obteniendo los siguientes resultados.

En las Figuras 16 y 17 podemos evaluar la calidad global de los alineamientos obtenidos. El bitscore refleja la fuerza de los alineamientos, mientras que el valor de $-\log_{10}$ (E-value) representa la significancia estadística, entre más altos sean estos, son más confiables. En ambas figuras se observa una tendencia lineal, en donde

los alineamientos de importancia se encuentran con un mayor bitscore y un mayor $-\log_{10}(\text{E-value})$. Los alineamientos restantes presentaron menor significancia.

Figura 16. Dispersión de Bitscores contra $-\text{Log}_{10}(\text{E-value})$ en el alineamiento con BLASTp en C. vulgaris.

*Figura 17. Dispersión de Bitscores contra $-\text{Log}_{10}(\text{E-value})$ en el alineamiento con BLASTp en *S. acutus*.*

En las Figuras 18 y 19 se aprecia la distribución de las frecuencias de identidades y proteínas homologas más cercanas contenidas en la base de datos. Podemos observar que algunos alineamientos muestran identidades moderadas (de 60% a 80%), no obstante, La mayoría de los hits correspondientes a proteínas hipotéticas presentó porcentajes de identidad elevados (>90%).

Figura 18. Distribución del porcentaje de identidad contra su frecuencia en C. vulgaris.

Figura 19. Distribución del porcentaje de identidad contra su frecuencia en S. acutus.

Para una mejor visualización de estos resultados, se elaboró un gráfico de burbujas, combinando el porcentaje de identidad, el bitscore y la significancia estadística, mostrando los siguientes resultados.

Figura 20. Calidad de los alineamientos en C. vulgaris.

Figura 21. Calidad de los alineamientos en S. acutus.

En las Figuras 20 y 21 se puede apreciar cada uno de los alineamientos obtenidos, cada circulo representa un hit positivo con la base de datos, los puntos más grandes poseen una mayor significancia estadística, a su vez, cuanto más estén a la izquierda, estos hits presentarán un porcentaje de identidad más elevado y estos que tengan una mayor coloración azul poseen un bitscore mayor. Comparando ambas figuras podemos observar que el análisis en las proteínas hipotéticas del metagenoma de *C. vulgaris* tiende a ser mayor en cuestión de numero de hits, sin embargo, las proteínas hipotéticas del metagenoma en *S. acutus* hay una mejor homogenización entre los hits en función el bitscore y el porcentaje de identidad. En cualquiera de los casos, en ambos análisis se obtuvieron hits de alta calidad. Cada uno de los hits positivos fueron guardados en un archivo txt en formato tsv.

8.3.2. Detección de genes implicados en la biorremediación de fármacos y metales pesados.

Para la detección de secuencias genéticas implicadas en la biorremediación de metales pesados y fármacos se comparó la anotación obtenida y se corrieron el software de alineamiento BLASTp con una base de datos interna del laboratorio, los cuales contienen genomas de *Arabidopsis thaliana*, *Chlamydomonas reinhardtii* y *Chlorella vulgaris*. Se tomaron los principales hits con los mejores valores de bitscore y E-value, obteniendo los siguientes resultados.

Tabla 10. Principales hits de proteínas implicados en la biorremediación de metales pesados en el metagenoma de C. vulgaris.

Tabla 11. Principales hits de proteínas implicados en la biorremediación de fármacos en el metagenoma de C. vulgaris.

Tabla 12. Principales hits de proteínas implicados en la biorremediación de metales pesados en el metagenoma de *S. acutus*.

Tabla 13. Principales hits de proteínas implicados en la biorremediación de fármacos en el metagenoma de S. acutus.

Por otro lado, la presencia de mono oxigenasas dependientes de flavinas (FMOs) indica que ambas microalgas participan en la biotransformación de compuestos orgánicos y fármacos, afirmando que ambas microalgas son capaces de degradar/transformar fármacos u otros xenobióticos (Huang et al., 2021). También, la detección de peptidasas (como aminopeptidasas, oligopeptidasas A, prolina endopeptidasas) apunta a la capacidad de degradar péptidos antimicrobianos, estos mecanismos son conocidos por inactivar péptidos bioactivos y modular resistencia microbiana (De la Nieta et al.,

2020). Del mismo modo, dichas predicciones se encuentran en un archivo txt en formato tsv.

8.4. Elucidación de mapas metabólicos de interés.

Para la elucidación de los mapas interactivos de interés se utilizaron las anotaciones de las funciones KO por parte de EggNOG y se limpiaron usando el lenguaje de programación R con paqueterías como Dplyr, Stringr y Tidy. Por último, los mapas fueron reconstruidos por repositorio en línea llamado KEGG Mapper. Debido a las limitaciones que presenta este repositorio, los mapas metabólicos se dividieron en tres mapas metabólicos de interés: biosíntesis de lípidos, resistencia a antimicrobianos y secuestro de metales pesados (antineoplásicos), a su vez, cada grupo de mapas metabólicos se subdividió en los diferentes genomas bacterianos aislados, esto para lograr una correcta visualización y análisis de cada uno de dichos mapas metabólicos.

8.4.1. Ruta de biosíntesis de lípidos.

Esta ruta es una de las más importantes desde un punto de vista biotecnológico debido al alto valor que tienen los lípidos, desde usos como biocombustibles, bioplásticos, aditivos, etc. La diferenciación de colores entre verde, rojo y azul dan lugar a qué microorganismos poseen el gen que está siendo coloreado, si este lo comparten entre los mismos microorganismos o estos mismos completan su ruta metabólica.

Figura 22. Mapa metabólico de la biosíntesis de ácidos grasos. Verde: Pseudomonas oryzae. Rojo: Microvirga Sp VF16.

Figura 23. Mapa metabólico de la biosíntesis de ácidos grasos. Verde: Chlorella vulgaris.

Figura 24. Mapa metabólico de la biosíntesis de ácidos grasos. Verde: Methylobacterium bullatum. Rojo: Cellulomonas sp. Y8. Azul: Streptomyces sp. Verde: Sphingomonas sp.

En las Figuras 22, 23 y 24 podemos observar los mapas metabólicos referentes a la biosíntesis de lípidos en el metagenoma de *C. vulgaris*, cada bacteria y microalga cuentan con rutas completas para la formación de ácidos grasos con diferente longitud de cadena. Todas las rutas están completas excepción de *Streptomyces sp* y *Sphingomonas sp*.

Figura 25. Mapa metabólico de la biosíntesis de lípidos neutros. Verde: Scenedesmus acutus.

*Figura 26. Mapa metabólico de la biosíntesis de lípidos neutros. Verde: Stutzerimonas stutzeri.
Rojo: Roseomonas fluvialis. Azul: Sphingomonas koreensis.*

Figura 27. Mapa metabólico de la biosíntesis de lípidos neutros. Verde: Sphingopyxis sp. Rojo: Flavobacterium sp. Azul: Nordella sp HKS 07.

Figura 28. Mapa metabólico de la biosíntesis de lípidos neutros. Verde: Enhydrobacter sp. Rojo: Dyadobacter fermentans DSM 18053.

En las Figuras 25, 26, 27 y 28 se muestran el mapa metabólico de la biosíntesis de ácidos grasos dividido en las bacterias presentes en el metagenoma de *S. acutus*. Podemos observar que, al igual que en metagenoma de *C. vulgaris*, las rutas enzimáticas para la elongación de las cadenas alifáticas de ácidos grasos se mantienen similares entre ambos metagenomas.

8.4.2. Ruta de absorción de metales pesados.

La ruta de captación y secuestro de metales pesados es una de las redes metabólicas más importantes cuando se habla de biorremediación de contaminantes. Es necesario describir con exactitud estos sistemas debido a la importancia que tiene la recuperación de estos metales mediante el procesamiento de la biomasa residual posterior a algún tipo de tratamiento.

Figura 29. Mapa metabólico de la captación y secuestro de metales pesados en C. vulgaris.

Figura 30. Mapa metabólico de la captación y secuestro de metales pesados en S. acutus.

En las Figuras 29 y 30 se muestra el mecanismo de secuestro de platino. Es importante enfatizar que esta ruta metabólica es propiamente para la resistencia de fármacos que tiene como principal componente al platino, sin embargo, esta ruta comparte bombas de flujo y eflujo que son necesarias para mantener una concentración de metales que permita mantener una homeostasis. Por ejemplo, en las figuras ya antes mencionadas, se observa la bomba *hCTR1* el cuales es un transportador transmembranal de cobre, sin embargo, este transportador funciona con diferentes metales pesados, tales como el plomo y el hierro. Esta ruta metabólica hace homología a la captación y secuestro de fármacos asociados con el platino divalente (Zhou & Gitschier, 1997), subsecuentemente, participa una gran familia de genes los cuales están definidos como *ATP7A/B*; los cuales son *HSA*, *PTR*, *PPS*, *GGO*, *PON*, *PPYG*, *NLE*, *HMH*, *SSYN* y *MCC*, estos genes actúan como secuestradores de metales de diferentes valencias para la internalización (y posterior contención) de metales, algunos metales pudieran ser excretados en forma de vacuolas, sin embargo, la gran mayoría permanecerán dentro de la célula en forma de biocontenedores (Vulpe et al., 1993). Este tipo de transporte no es propio de organismos eucariontes, en algunas bacterias como *E. coli* se han descrito estos poros selectivos de iones metálicos, no obstante, estos poros solo funcionan como manutención homeostática y detoxificación de cobre y otros metales (Fan & Rosen, 2002).

8.4.3. Ruta de resistencia a antibióticos (β -lactámicos).

Aunque se encontraron diversos resistomas en ambos metagenomas, la ruta de resistencia a los β -lactámicos es una de las más importantes, debido a que esta familia de antibióticos es una de las usadas en la medicina general por sus constantes actualizaciones y nuevos antibióticos de cuarta generación como los son las cefalosporinas (Tilanus & Drusano, 2023). Sin embargo, el uso indiscriminado de estos antibióticos y la presión ambiental derivada del estrés constate por la presencia de contaminantes o de antibióticos a baja concentración generan mecanismos de resistencia a estos mismos (Fu et al., 2022).

Figura 31. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: Pseudomonas oryzihabitans. Rojo: Microvirga Sp VF16.

Figura 32. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: Chlorella vulgaris.

Figura 33. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: Methylobacterium bullatum. Rojo: Cellulomonas sp. Y8. Azul: Streptomyces sp. Verde: Sphingomonas sp.

En las Figuras 31, 32 y 33 se observan los diferentes mapas metabólicos relacionados con la resistencia a los antibióticos β -lactámicos, en ellos, podemos observar diferentes genes que presentes relacionados con el procesamiento de esta familia de antibióticos, sin embargo, en la Figura 32 perteneciente a *C. vulgaris*, podemos observar una ausencia de estos genes, tanto de poros especializados para internalizar estas moléculas (*oprD*) hasta enzimas especializadas para la degradación de estas (*blaTEM*).

Figura 34. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: Scenedesmus acutus.

Figura 35. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: Stutzerimonas stutzeri. Rojo: Roseomonas fluvialis. Azul: Sphingomonas koreensis.

Figura 36. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: Sphingopyxis sp. Rojo: Flavobacterium sp. Azul: Nordella sp HKS 07.

Figura 37. Mapa metabólico relacionado a la resistencia a los β -lactámicos. Verde: Enhydrobacter sp. Rojo: Dyadobacter fermentans DSM 18053.

No obstante, debido a las limitaciones inherentes al ensamblaje metagenómico y la dificultad para confirmar la posible integración de estos genes en organismos eucariontes y solo partiendo de evidencia bioinformática, estos resultados deben interpretarse con cautela. Por lo tanto, estos hallazgos constituyen una evidencia potencial a la transferencia horizontal de genes, más que una demostración concluyente de la integración estable de estos genes en el genoma microalgal.

8.5. Análisis de redes bipartita.

Una red bipartita es un tipo de gráfico cuyos nodos se dividen en conjuntos, donde las aristas solo se conectan con nodos de conjuntos diferentes, pero nunca entre sí. En metagenómica, se utiliza para visualizar el análisis funcional de conjuntos de genes en de varios genomas bacterianos recuperados y como estos se interconectan entre sí. Con fines de mostrar dichas redes en este trabajo, se procesaron cada genoma bacteriano anotado con el software Prokka, dichas anotaciones se procesaron y filtraron solamente los KOs relacionados a genes la biorremediación de metales pesados y genes relacionados a resistencia a antibióticos en R para generar un archivo graphml, donde los nodos son los

genomas recuperados y cada vértice son estos genes implicados, este análisis se llevó a cabo en el software Gephi (versión: 0.10.1).

Figura 38. Red bipartita del metagenoma de C. vulgaris.

Figura 39. Red bipartita del metagenoma de S. acutus.

Las Figuras 38 y 39 muestran las redes bipartitas correspondientes a los metagenomas de *C. vulgaris* y *S. acutus*, respectivamente. En ambos casos, diversos MAGs comparten genes asociados con la respuesta a metales. En contraste, algunos genes relacionados con resistencia a antibióticos se concentran principalmente en determinados MAGs. En *C. vulgaris*, este patrón es particularmente evidente en *Pseudomonas oryzae*, mientras que en *S. acutus* se observa en *Stutzerimonas stutzeri*. Este patrón podría estar relacionado con las condiciones ambientales de origen de los consorcios; sin embargo, esta posibilidad deberá evaluarse experimentalmente (Joseph et al., 2023).

8.6. Análisis de expresión génica en los ensamblajes metagenómicos a partir de datos de RNA-seq.

Para el análisis de expresión génica se partieron de los ensamblajes metagenómicos originados por el software MEGAHIT, a su vez, se prepararon las lecturas pareadas de RNA-seq, los cuales contienen el transcriptoma en condiciones basales de cada

microalga, así también como el sometimiento de 2 tratamientos: luz azul y roja. Para conocer la ubicación exacta de cada gen anotado, se usaron las anotaciones originado por Prokka. Los alineamientos se llevaron a cabo en Rstudio con las paqueterías “Rsubread” y “Biostrings” (Liao et al., 2019. Aboyoun, 2017).

Una vez creados los alineamientos, se obtuvieron transcritos por millón (TPMs) mediante FeatureCounts de genes participantes en la ruta KEGG 01501 que describe la ruta hidrolítica de los antibióticos β -lactámicos, los cuales se trataron por el método estadístico de Spearman para observar expresiones diferenciales en cada tratamiento. Posteriormente con los KOs presentes en la ruta ya antes mencionada, se mapearon con los MAGs bacterianos aislados para crear redes de co-expresión, obteniendo los siguientes resultados.

Figura 40. Análisis de Componentes Principales (PCA) de los transcriptomas en Chlorella vulgaris en los tratamientos con luz azul y roja.

Figura 41. Análisis de Componentes Principales (PCA) de los transcriptomas en Scenedesmus acutus en los tratamientos con luz azul y roja.

Figura 42. Heatmap con los genes con mayor variabilidad en C.vulgaris.

Figura 43. Heatmap con los genes con mayor variabilidad en S. acutus.

8.6.1. Efecto de los tratamientos lumínicos en la actividad transcripcional de genes relacionados con la resistencia a antibióticos y biosíntesis de lípidos.

Para este análisis, se utilizaron las anotaciones KO originadas por EggNOG, seleccionando dos rutas metabólicas de interés: Resistencia a antibióticos β -lactámicos (ruta en KEGG: map01501) y biosíntesis de ácidos grasos (ruta en KEGG: map00061), se trataron con el método de Pearson para correlacionar los niveles de expresión con los tratamientos.

Figura 44. Expresión de genes en la ruta de Resistencia a β -lactámicos en el metagenoma de C. vulgaris bajo luz blanca (control), azul y roja.

Figura 45. Coexpresión de genes asociados a la resistencia de β -lactámicos bajo luz azul y roja en el metagenoma de C. vulgaris.

En la Figura 44, podemos observar los niveles de expresión relativa ($\text{Log}_2(\text{TPM} + 1)$) en *C. vulgaris*. Podemos apreciar que en ninguna replica de las muestras de los transcriptomas se pudieron formar clústeres de expresión en función del tratamiento, esto puede ver que la heterogeneidad en el perfil transcripcional de estos genes no se ve influenciada por el tipo de luz irradiada. En la Figura 45, se puede confirmar estas observaciones, podemos ver en el heatmap de coexpresión que muy pocos genes responden entre sí, infiriendo que la respuesta transcriptómica derivada del estrés por los tratamientos lumínicos es mínima.

Figura 46. Expresión de genes en la ruta de Resistencia a β -lactámicos en el metagenoma de S. acutus bajo luz blanca (control), azul y roja.

Figura 47. Coexpresión de genes asociados a la resistencia de β -lactámicos bajo luz azul y roja en el metagenoma de S. acutus.

Figura 48. Expresión de genes en la ruta de biosíntesis de ácidos grasos en el metagenoma de C. vulgaris bajo luz blanca (control), azul y roja.

Figura 49. Coexpresión de genes asociados a la biosíntesis de ácidos grasos bajo luz azul y roja en el metagenoma de C. vulgaris.

En la Figura 48 podemos observar la expresión génica de cada uno de los genes implicados en la biosíntesis de ácidos grasos, podemos observar que cada uno de estos genes predomina una expresión basal, infiriendo que, sin importar el tratamiento lumínico, habrá una actividad transcripcional basal. En la Figura 49 podemos observar el comportamiento transcriptómico de estos genes en comparación consigo mismos en la misma ruta metabólica, podemos observar que la coexpresión de estos genes se comporta de manera irregular, existiendo genes que se sobre expresan de manera diferencial en comparación con otros genes. No obstante, esto no significa que la ruta metabólica esté inactiva, si no que varios genes se expresan más en nodos enzimáticos importantes, según las necesidades del microorganismo.

Figura 50. Expresión de genes en la ruta de biosíntesis de ácidos grasos en el metagenoma de S. acutus bajo luz blanca (control), azul y roja.

Figura 51. Coexpresión de genes asociados a la biosíntesis de ácidos grasos bajo luz azul y roja en el metagenoma de S. acutus.

8.7. Análisis exploratorio basado en algoritmos de IA (Machine Learning).

8.7.1. Construcción de matrices funcionales.

$$FS = \frac{KO_{MAG}}{KO_{Dir}} \quad (1)$$

Donde:

FS: Puntuación de funcionalidad (Functional Score).

KO_{MAG}: Numero de KOs presentes del diccionario en el MAG.

KO_{Dir}: Numero de KOs totales del diccionario.

Podemos ver en la Ecuación (1) como es calculado la puntuación de funcionalidad (FS), generando valores de 0 a 1, el cual infiere lo siguiente: valores más cercanos a 1 tienen una participación y presencia mayores en el metagenoma, por consiguiente, la ruta metabólica se vuelve más importante (Romero et al., 2025), obteniendo los siguientes resultados:

Tabla 14. Matriz de puntuación de funcionalidad en el metagenoma de C. vulgaris.

Tabla 15. Matriz de puntuación de funcionalidad en el metagenoma de *S. acutus*.

8.7.2. Cálculo de centralidad e influencia ecológica (Influence Score).

Para comenzar el modelado exploratorio es necesario determinar variables de respuesta viables que describan la influencia que tienen estas rutas metabólicas en el desarrollo del consorcio bacteriano, para este caso, los valores que conforman la influencia ecológica son los siguientes: Grado (degree), intermediación (betweenness), cercanía (closeness) y Eigenvector. Cada una de estas métricas de centralidad permiten identificar “centros” (hubs) clave dentro de la comunidad microbiana, en este caso, estos centros son cual de estas 4 rutas metabólicas son las más importantes (Peschel et al., 2020).

El grado determina cuantas conexiones tiene un nodo y matemáticamente se calcula con la siguiente ecuación:

$$Degree(i) = \sum_j a_{ij} \quad (2)$$

Donde:

Degree(i): Grado

a_{ij} : valor de cada conexión.

La intermediación describe si los nodos conectan con regiones distintas a su red, es decir, que tanto un nodo funciona como intermediario entre las características de la red. Matemáticamente se representa de la siguiente manera:

$$BC(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (3)$$

Donde:

$BC(v)$: intermediación.

σ : Cambios mínimos.

En biología, un MAG con alta intermediación significa que puede conectar con procesos metabólicos distintos.

La cercanía es una métrica que permite calcular que tan cerca está un nodo de otros y matemáticamente se representa de la siguiente manera:

$$CC(i) = \frac{1}{\sum d(i,j)} \quad (4)$$

Donde:

$CC(i)$: Cercanía.

d : distancia de los MAGs.

El Eigenvector es una de las métricas que más considera características de la misma red. Este valor representa vectores de importancia en los nodos e identifica

que tantas ramificaciones tiene un nodo. Matemáticamente se ejemplifica de la siguiente manera:

$$Ax = \lambda x$$

(5)

Tabla 16. Métricas de centralidad para cada MAG en el metagenoma de C. vulgaris.

Tabla 17. Métricas de centralidad para cada MAG en el metagenoma de S. acutus.

8.7.3. Planteamiento del modelo de aprendizaje automático.

Figura 52. Gráfico de importancia en el modelamiento con Random Forest de las diferentes rutas metabólicas en el metagenoma de C. vulgaris.

Figura 53. Gráfico de valores observados contra valores predichos por el modelo en el análisis en el metagenoma de C. vulgaris.

Figura 54. Gráfico de importancia en el modelamiento con Random Forest de las diferentes rutas metabólicas en el metagenoma de S. acutus.

Figura 55. Gráfico de valores observados contra valores predichos por el modelo en el análisis en el metagenoma de S. acutus.

9. Conclusiones.

10. Bibliografia.

- ❖ Aditya, L., Mahlia, T. I., Nguyen, L. N., Vu, H. P., & Nghiem, L. D. (2022). Microalgae-bacteria consortium for wastewater treatment and biomass production. *The Science Of The Total Environment*, 838, 155871. <https://doi.org/10.1016/j.scitotenv.2022.155871>
- ❖ Ali, H., & Khan, E. (2018). Trophic transfer, bioaccumulation, and biomagnification of non-essential hazardous heavy metals and metalloids in food chains/webs—*Concepts and implications for wildlife and human health. Human And Ecological Risk Assessment An International Journal*, 25(6), 1353-1376. <https://doi.org/10.1080/10807039.2018.1469398>
- ❖ Ashraf, N., Ahmad, F., & Lu, Y. (2022). Synergy between microalgae and microbiome in polluted waters. *Trends In Microbiology*, 31(1), 9-21. <https://doi.org/10.1016/j.tim.2022.06.004>
- ❖ Babraham Bioinformatics - FastQC a quality Control tool for high throughput sequence data. (s. f.). <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- ❖ Balzano, S., Sardo, A., Blasio, M., Chahine, T. B., Dell'Anno, F., Sansone, C., & Brunet, C. (2020). Microalgal Metallothioneins and Phytochelatins and Their Potential Use in Bioremediation. *Frontiers In Microbiology*, 11, 517. <https://doi.org/10.3389/fmicb.2020.00517>

- ❖ Bauernfeind, A., Stemplinger, I., Jungwirth, R., Wilhelm, R., & Chong, Y. (1996). Comparative characterization of the cephamycinase blaCMY-1 gene and its relationship with other beta-lactamase genes. *Antimicrobial Agents And Chemotherapy*, 40(8), 1926-1930. <https://doi.org/10.1128/aac.40.8.1926>
- ❖ Behera, S., Nayak, S. C., & Kumar, A. V. S. P. (2023). A Comprehensive Survey on Higher Order Neural Networks and Evolutionary Optimization Learning Algorithms in Financial Time Series Forecasting. *Archives Of Computational Methods In Engineering*, 30(7), 4401-4448. <https://doi.org/10.1007/s11831-023-09942-9>
- ❖ Biocodex Microbiota Institute. (2025, 1 diciembre). *Microbioma intestinal y resistencia a los antimicrobianos*. <https://www.biocodexmicrobiotainstitute.com/es/pro/microbioma-intestinal-y-resistencia-los-antimicrobianos>
- ❖ Bradford, P. A. (2001). Extended-Spectrum β -Lactamases in the 21st Century: Characterization, Epidemiology, and Detection of This Important Resistance Threat. *Clinical Microbiology Reviews*, 14(4), 933-951. <https://doi.org/10.1128/cmr.14.4.933-951.2001> .
- ❖ Breiman, L., Cutler, A., Liaw, A., & Wiener, M. (2002). randomForest: Breiman and Cutlers Random Forests for Classification and Regression [Conjunto de datos]. <https://doi.org/10.32614/cran.package.randomforest>
- ❖ Cao, Q., Li, X., Chen, Y., Li, X., Xie, Z., & Li, D. (2022). Nitrification resistance and functional redundancy maintain the system stability of partial nitrification in high-strength ammonium wastewater system. *Bioresource Technology*, 365, 128157. <https://doi.org/10.1016/j.biortech.2022.128157>
- ❖ Cai, Z., Karunakaran, E., & Pandhal, J. (2024). Bottom-up construction and screening of algae-bacteria consortia for pollutant biodegradation. *Frontiers In Microbiology*, 15. <https://doi.org/10.3389/fmicb.2024.1349016>
- ❖ Camacho, D. M., Collins, K. M., Powers, R. K., Costello, J. C., & Collins, J. J. (2018). *Next-Generation Machine Learning for Biological Networks*. *Cell*, 173(7), 1581-1592. <https://doi.org/10.1016/j.cell.2018.05.015>

- ❖ Chen, Q., Ding, J., Zhu, Y., He, J., & Hu, H. (2020). Soil bacterial taxonomic diversity is critical to maintaining the plant productivity. *Environment International*, 140, 105766. <https://doi.org/10.1016/j.envint.2020.105766>
- ❖ Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., T, B., DO, Way, G. P., Ferrero, E., Agapow, P., Zietz, M., Hoffman, M. M., Xie, W., Rosen, G. L., Lengerich, B. J., Israeli, J., Lanchantin, J., Woloszynek, S., Carpenter, A. E., Shrikumar, A., Xu, J., . . . Greene, C. S. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal Of The Royal Society Interface*, 15(141), 20170387. <https://doi.org/10.1098/rsif.2017.0387>
- ❖ Deepa, N., Chauhan, S., Trivedi, P. K., & Singh, A. (2025). Endophytic *Pseudomonas oryzae* CB24 Boosts Photosynthesis Through Monogalactosyldiacylglycerol Driven Lipid Reprogramming. *Plant Cell & Environment*. <https://doi.org/10.1111/pce.70149>
- ❖ De la Nieta, R. S., Antoraz, S., Alzate, J. F., Santamaría, R. I., & Díaz, M. (2020). Antibiotic Production and Antibiotic Resistance: The Two Sides of AbrB1/B2, a Two-Component System of *Streptomyces coelicolor*. *Frontiers In Microbiology*, 11, 587750. <https://doi.org/10.3389/fmicb.2020.587750>
- ❖ Dong, X., Lan, H., Huang, L., Zhang, H., Lin, X., Weng, S., Peng, Y., Lin, J., Wang, J., Peng, J., & Yang, Y. (2022). Metagenomic Views of Microbial Communities in Sand Sediments Associated with Coral Reefs. *Microbial Ecology*, 85(2), 465-477. <https://doi.org/10.1007/s00248-021-01957-8>
- ❖ Ducousso-Détrez, A., Raveau, R., Fontaine, J., Hijri, M., & Sahraoui, A. L. (2022). Glomerales Dominate Arbuscular Mycorrhizal Fungal Communities Associated with Spontaneous Plants in Phosphate-Rich Soils of Former Rock Phosphate Mining Sites. *Microorganisms*, 10(12), 2406. <https://doi.org/10.3390/microorganisms10122406>
- ❖ Eraslan, G., Avsec, Ž., Gagneur, J., & Theis, F. J. (2019). Deep learning: new computational modelling techniques for genomics. *Nature Reviews Genetics*, 20(7), 389-403. <https://doi.org/10.1038/s41576-019-0122-6>

- ❖ Fan, B., & Rosen, B. P. (2002). Biochemical Characterization of CopA, the *Escherichia coli* Cu(I)-translocating P-type ATPase. *Journal Of Biological Chemistry*, 277(49), 46987-46992. <https://doi.org/10.1074/jbc.m208490200>
- ❖ Favero, V. O., Carvalho, R. H., Motta, V. M., Leite, A. B. C., Coelho, M. R. R., Xavier, G. R., Rumjanek, N. G., & Urquiaga, S. (2021). *Bradyrhizobium* as the only rhizobial inhabitant of mung bean (*Vigna radiata*) Nodules in Tropical Soils: A strategy based on microbiome for improving biological nitrogen fixation using bio-products. *Frontiers In Plant Science*, 11. <https://doi.org/10.3389/fpls.2020.602645>
- ❖ Fu, Y., Zhu, Y., Dong, H., Li, J., Zhang, W., Shao, Y., & Shao, Y. (2022). Effects of heavy metals and antibiotics on antibiotic resistance genes and microbial communities in soil. *Process Safety And Environmental Protection*, 169, 418-427. <https://doi.org/10.1016/j.psep.2022.11.020>
- ❖ Goda, E. S., Ameen, H., Abady, M. M., Kiey, S. A. A., Makhoulf, G., Gamal, H., Abdelkhalik, A., & Wang, D. H. (2025). N-methylene phosphonic acid gelatin encapsulated copper sulfide nanoparticles as smart antibacterial materials. *Journal Of Alloys And Compounds*, 1031, 180942. <https://doi.org/10.1016/j.jallcom.2025.180942>
- ❖ González-González, L. M., & De-Bashan, L. E. (2021). Toward the Enhancement of Microalgal Metabolite Production through Microalgae–Bacteria Consortia. *Biology*, 10(4), 282. <https://doi.org/10.3390/biology10040282>
- ❖ Gupta, A., Naseem, M., Gupta, E., Srivastava, P. K., & Barik, S. K. (2025). Metagenomic profiling of cyanide-degrading microbial communities in steel industry wastewater with an implication for bioremediation. *Frontiers Of Environmental Science & Engineering*, 19(10). <https://doi.org/10.1007/s11783-025-2057-9>
- ❖ Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072-1075. <https://doi.org/10.1093/bioinformatics/btt086>

- ❖ Guzmán-Herrador, D. L., & Llosa, M. (2019). The secret life of conjugative relaxases. *Plasmid*, 104, 102415. <https://doi.org/10.1016/j.plasmid.2019.102415>
- ❖ Hall, B. G., Salipante, S. J., & Barlow, M. (2003). The Metallo- β -Lactamases Fall into Two Distinct Phylogenetic Groups. *Journal Of Molecular Evolution*, 57(3), 249-254. <https://doi.org/10.1007/s00239-003-2471-0>
- ❖ Huang, S., Howington, M. B., Dobry, C. J., Evans, C. R., & Leiser, S. F. (2021). Flavin-Containing Monooxygenases Are Conserved Regulators of Stress Resistance and Metabolism. *Frontiers In Cell And Developmental Biology*, 9, 630188. <https://doi.org/10.3389/fcell.2021.630188>
- ❖ Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., Mende, D. R., Letunic, I., Rattei, T., Jensen, L. J., Von Mering, C., & Bork, P. (2018). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*, 47(D1), D309-D314. <https://doi.org/10.1093/nar/gky1085>
- ❖ Joseph, O. J., Ogunleye, G. E., Oyinlola, K. A., Balogun, A. I., & Olumeko, D. T. (2023). Co-occurrence of heavy metals and antibiotics resistance in bacteria isolated from metal-polluted soil. *Environmental Analysis Health And Toxicology*, 38(4), e2023024. <https://doi.org/10.5620/eaht.2023024>
- ❖ Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., Perry, T., Kao, D., Mason, A. L., Madsen, K. L., & Wong, G. K. (2016). Characterization of the gut microbiome using 16S or shotgun metagenomics. *Frontiers In Microbiology*, 7. <https://doi.org/10.3389/fmicb.2016.00459>
- ❖ Juárez, K., Reza, L., Bretón-Deval, L., Morales-Guzmán, D., Trejo-Hernández, M. R., García-Guevara, F., & Lara, P. (2023). Microaerobic degradation of crude oil and long chain alkanes by a new *Rhodococcus* strain from Gulf of Mexico. *World Journal Of Microbiology And Biotechnology*, 39(10). <https://doi.org/10.1007/s11274-023-03703-3>
- ❖ Kanehisa, M. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(1), 27-30. <https://doi.org/10.1093/nar/28.1.27>

- ❖ Kanehisa, M., Sato, Y., & Morishima, K. (2015). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal Of Molecular Biology*, 428(4), 726-731. <https://doi.org/10.1016/j.jmb.2015.11.006>
- ❖ Kwatra, B. B. (2021). In Silico Prediction and comparison of resistomes in model pseudomonas strains by Resistance Gene Identifier (RGI). *BioRxiv*. <https://doi.org/10.1101/2021.11.15.468576>
- ❖ LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- ❖ Li, D., Liu, C., Luo, R., Sadakane, K., & Lam, T. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31(10), 1674-1676. <https://doi.org/10.1093/bioinformatics/btv033>
- ❖ Liao, Y., Smyth, G. K., & Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Research*, 41(10), e108. <https://doi.org/10.1093/nar/gkt214>
- ❖ Liao, Y., Smyth, G. K., & Shi, W. (2013b). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923-930. <https://doi.org/10.1093/bioinformatics/btt656>
- ❖ Madhubalaji, C. K., Ravi, S., & Mudliar, S. N. (2021). Unraveling of Chlorella-associated bacterial load, diversity, and their imputed functions at high- and low-yield conditions through metagenome sequencing. *Journal Of Phycology*, 58(1), 133-145. <https://doi.org/10.1111/jpy.13225>
- ❖ McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible Interactive Analysis and graphics of microbiome census data. *PLoS ONE*, 8(4), e61217. <https://doi.org/10.1371/journal.pone.0061217>
- ❖ Min, S., Lee, B., & Yoon, S. (2016). Deep learning in bioinformatics. Briefings in bioinformatics, bbw068. <https://doi.org/10.1093/bib/bbw068>
- Ondov, B. D., Bergman, N. H., & Phillippy, A. M. (2011). Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics*, 12(1). <https://doi.org/10.1186/1471-2105-12-385>

- ❖ Naaz, S., Ahmad, N., Jameel, M. R., Al-Huqail, A. A., Khan, F., & Qureshi, M. I. (2023). Impact of Some Toxic Metals on Important ABC Transporters in Soybean (*Glycine max* L.). *ACS Omega*, 8(30), 27597-27611. <https://doi.org/10.1021/acsomega.3c03325>
- ❖ Ndlela, L. L., Schroeder, P., Genthe, B., & Cruzeiro, C. (2023). Removal of Antibiotics Using an Algae-Algae Consortium (*Chlorella protothecoides* and *Chlorella vulgaris*). *Toxics*, 11(7), 588. <https://doi.org/10.3390/toxics11070588>
- ❖ Pagès H, Aboyoun P, Gentleman R, DebRoy S (2026). Biostrings: Efficient manipulation of biological strings. doi:10.18129/B9.bioc.Biostrings. R package version 2.80.1, <https://bioconductor.org/packages/Biostrings>.
- ❖ Peng, H., De-Bashan, L. E., & Higgins, B. T. (2020). *Azospirillum brasilense* reduces oxidative stress in the green microalgae *Chlorella sorokiniana* under different stressors. *Journal Of Biotechnology*, 325, 179-185. <https://doi.org/10.1016/j.jbiotec.2020.10.029>
- ❖ Peschel, S., Müller, C. L., Von Mutius, E., Boulesteix, A., & Depner, M. (2020). NetCoMi: network construction and comparison for microbiome data in R. *Briefings In Bioinformatics*, 22(4). <https://doi.org/10.1093/bib/bbaa290>
- ❖ Pödör, Z., & Hekfusz, M. (2024). Comparing Feature Selection Methods on Metagenomic Data using Random Forest Classifier. *Transactions On Machine Learning And Artificial Intelligence*, 12(1). <https://doi.org/10.14738/tecs.121.16525>
- ❖ Poirel, L., Naas, T., & Nordmann, P. (2009). Diversity, Epidemiology, and Genetics of Class D β -Lactamases. *Antimicrobial Agents And Chemotherapy*, 54(1), 24-38. <https://doi.org/10.1128/aac.01512-08>
- ❖ Qin, S., Nie, B., Wei, F., Lin, Y., Jiang, L., & Zhang, Q. (2025). Quorum Sensing Molecules Promote the Nutrients Removal in a Hn-Ad Bacteria - *Chlorella* Symbiotic System. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5287602>

- ❖ Robertson. (2023). MOB-suite database v. 3.1.8 [Conjunto de datos]. En Zenodo (CERN European Organization for Nuclear Research). <https://doi.org/10.5281/zenodo.10304948>
- ❖ Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), 2068-2069. <https://doi.org/10.1093/bioinformatics/btu153>
- ❖ Romero, F., Labouyrie, M., Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Tedersoo, L., Bahram, M., Eisenhauer, N., Sünneemann, M., Guerra, C. A., Tao, D., Rog, I., Jiao, S., Mocali, S., Rillig, M. C., Lehmann, A., Delgado-Baquerizo, M., & Van Der Heijden, M. G. A. (2025). The soil microbiome as an indicator of ecosystem multifunctionality in European soils. *Nature Communications*, 17(1), 705. <https://doi.org/10.1038/s41467-025-67353-9>
- ❖ Schäferling, S. (2023). Artificial intelligence and automated decision-making: terminological clarifications. *En Law, governance and technology series* (pp. 13-21). https://doi.org/10.1007/978-3-031-48125-3_2
- ❖ Schmidt, H., Korfmann, G., Barth, H., & Martin, H. H. (1995). The signal transducer encoded by ampG is essential for induction of chromosomal AmpC -lactamase in Escherichia coli by -lactam antibiotics and unspecific inducers. *Microbiology*, 141(5), 1085-1092. <https://doi.org/10.1099/13500872-141-5-1085>
- ❖ Singh, S. K., Sharma, A., & Sundaram, S. (2025). Evaluation of Chlorella vulgaris as sensitive, cost-effective, and environmentally sustainable biosensor tools for heavy metal monitoring in aquatic ecosystems. *Environmental Monitoring And Assessment*, 197(5), 504. <https://doi.org/10.1007/s10661-025-13960-4>
- ❖ Stephens, Z. D., Lee, S. Y., Faghri, F., Campbell, R. H., Zhai, C., Efron, M. J., Iyer, R., Schatz, M. C., Sinha, S., & Robinson, G. E. (2015). Big data: astronomical or genomics? *PLoS Biology*, 13(7), e1002195. <https://doi.org/10.1371/journal.pbio.1002195>
- ❖ Tenenbaum, D. (2017). KEGGREST. NCI BioConductor. <https://doi.org/10.18129/b9.bioc.keggrest>

- ❖ Tian, J., Hu, J., Xiong, Y., Deng, X., Fang, Y., Wang, G., Chi, R., & Xiao, C. (2025). Metagenomic and metabolomic insights into microalgal-bacterial symbiosis under low carbon-to-nitrogen ratios. *Bioresource Technology*, 434, 132849. <https://doi.org/10.1016/j.biortech.2025.132849>
- ❖ Tilanus, A., & Drusano, G. (2023). Optimizing the Use of Beta-Lactam Antibiotics in Clinical Practice: A Test of Time. *Open Forum Infectious Diseases*, 10(7), ofad305. <https://doi.org/10.1093/ofid/ofad305>
- ❖ Torres, J. J. T. M. (2025). Análisis transcriptómico asociado a la producción de lípidos en *Chlorella vulgaris* y *Scenedesmus acutus* por efecto de la luz azul y roja [Tesis doctoral]. Universidad Autónoma de San Luis Potosí.
- ❖ Vargas-Ordóñez, A., Aguilar-Romero, I., Villaverde, J., Madrid, F., & Morillo, E. (2023). Isolation of Novel Bacterial Strains *Pseudomonas extremaustralis* CSW01 and *Stutzerimonas stutzeri* CSW02 from Sewage Sludge for Paracetamol Biodegradation. *Microorganisms*, 11(1), 196. <https://doi.org/10.3390/microorganisms11010196>
- ❖ Vulpe, C., Levinson, B., Whitney, S., Packman, S., & Gitschier, J. (1993). Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. *Nature Genetics*, 3(1), 7-13. <https://doi.org/10.1038/ng0193-7>
- ❖ Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, 20(1). <https://doi.org/10.1186/s13059-019-1891-0>
- ❖ Zaytsev, P. A., Shurygin, B. M., Rodin, V. A., Panova, T. V., Zvereva, M. I., Skripnikova, E. V., & Solovchenko, A. E. (2024). Comparative Metagenomics for Monitoring the Hidden Dynamics of the Algal-Bacterial Wastewater Community under the Influence of Drugs. *Nanobiotechnology Reports*, 19(3), 393-407. <https://doi.org/10.1134/s2635167624600470>
- ❖ Zhou, B., & Gitschier, J. (1997). hCTR1: A human gene for copper uptake identified by complementation in yeast. *Proceedings Of The National Academy Of Sciences*, 94(14), 7481-7486. <https://doi.org/10.1073/pnas.94.14.7481>

Reporte de similitud.



Página 1 de 107 - Portada

Identificador de la entrega: tmsoid::3117615712605

Abner Martínez

Tesis Maestría_2026_final_reporte_similitud_AMC.pdf



Universidad Autónoma de San Luis Potosí

Detalles del documento

Identificador de la entrega

tmsoid::3117615712605

Fecha de entrega

20 ago 2026, 9:27 a.m. GMT-6

Fecha de descarga

20 ago 2026, 9:43 a.m. GMT-6

Nombre del archivo

Tesis Maestría_2026_final_reporte_similitud_AMC.pdf

Tamaño del archivo

6.3 MB

103 páginas

18.060 palabras

109.896 caracteres



Página 1 de 107 - Portada

Identificador de la entrega: tmsoid::3117615712605

6% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- ▶ Bibliografía
- ▶ Texto citado
- ▶ Texto mencionado
- ▶ Coincidencias menores (menos de 10 palabras)
- ▶ Abstract
- ▶ Methods and Materials

Fuentes principales

- 6%  Fuentes de Internet
- 1%  Publicaciones
- 0%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.