

UNIVERSIDAD DEL VALLE



Viromas y plasmidomas asociados a tejido foliar y radicular de *Rhizophora mangle* L. (*Rhizophoraceae*) y *Avicennia germinans* L. (*Acanthaceae*) en la Bahía de Buenaventura, Colombia

Carlo Santiago Becerra Casierra

Director: Nelson Toro Perea, PhD

Maestría en Ciencias - Biología

Facultad de Ciencias Naturales y Exactas

Septiembre 01 de 2025

Contenido

Resumen general.	5
General abstract.	7
Capítulo 1: Introducción y métodos generales.	9
Introducción general.	9
Objetivos.	11
Objetivo general.	11
Objetivos específicos.	11
Hipótesis.	11
Métodos generales.	11
Toma de muestras.	11
Extracción de ácidos nucleicos.	12
Identificación molecular de las plantas hospederas.	12
Construcción de librerías y secuenciación.	13
Análisis bioinformático.	13
- Limpieza de secuencias de baja calidad.	13
- Eliminación del genoma de la planta hospedera.	13
Capítulo 2: Comunidad viral y viromas endófitos de ADN del tejido foliar y radicular de <i>R. mangle</i> y <i>A. germinans</i>.	14
Introducción.	14
Objetivo específico.	15
Métodos.	15
Ensamblaje de secuencias limpias.	16
Anotación taxonómica del viroma asociado a cada tejido y a cada especie de mangle.	16
Análisis estadísticos y de diversidad alfa y beta.	16
Detección, filtrado y anotación de genomas virales reconstruidos por especie de planta.	16
Anotación de genes auxiliares metabólicos en los vMAGs reconstruidos y ensamblados por planta.	17
Anotación estructural y funcional de los vMAGs reconstruidos.	18
Resultados.	18
Anotación taxonómica de la comunidad viral.	18
Análisis de diversidad alfa.	19
Análisis estadísticos y de diversidad beta.	21
Anotación taxonómica y análisis de calidad de los vMAGs reconstruidos y filtrados con VIRify	22
Anotación de genes auxiliares metabólicos en los vMAGs reconstruidos obtenidos por planta.	23
- Anotación funcional con la base de datos de PLaBAs de AMG de los genomas virales reconstruidos.	23
- Enriquecimiento funcional con PLaBAs de los AMGs de los genomas virales reconstruidos de <i>A. germinans</i> .	24

- Enriquecimiento funcional con PLaBAs de los AMGs de los vMAGs reconstruidos de <i>R. mangle</i> .	25
- Anotación funcional de CAZimas en los vMAGs reconstruidos de muestras de <i>A. germinans</i> y <i>R. mangle</i> .	27
Discusión.	28
Análisis taxonómico de comunidades virales asociadas a tejidos foliares y radiculares de <i>A. germinans</i> y <i>R. mangle</i> .	28
Análisis de diversidad de comunidades virales asociadas a tejidos foliares y radiculares de <i>A. germinans</i> y <i>R. mangle</i> .	30
Análisis de calidad de los genomas virales reconstruidos y filtrados con VIRify.	31
Análisis estructurales y funcionales de los genomas virales reconstruidos y anotados por especie de <i>mangle</i> .	31
- Análisis estructural de genomas virales de <i>A. germinans</i> y <i>R. mangle</i> .	31
- Análisis funcional de AMGs de vMAGs de <i>A. germinans</i> y <i>R. mangle</i> .	32
- Anotación funcional de CAZimas en los genomas virales reconstruidos de muestras de <i>A. germinans</i> y <i>R. mangle</i> .	34
Capítulo 3: Plasmidoma endófito del tejido foliar y radicular de <i>R. mangle</i> y <i>A. germinans</i>.	35
Introducción.	35
Objetivo específico.	37
Métodos.	37
Análisis estadísticos y de diversidad alfa y beta.	37
Ensamblaje de secuencias limpias.	38
Filtrado de plásmidos, identificación de genes de movilidad e identificación de plásmido completo.	38
Asignación taxonómica de los plásmidos ensamblados.	38
Anotación funcional de plásmidos ensamblados.	39
Anotación estructural de los plásmidos ensamblados.	39
Resultados.	40
Anotación taxonómica del plasmidoma de <i>A. germinans</i> y <i>R. mangle</i> .	40
Análisis de diversidad alfa del plasmidoma de <i>A. germinans</i> y <i>R. mangle</i> .	41
Análisis estadísticos y de diversidad beta del plasmidoma de <i>A. germinans</i> y <i>R. mangle</i> .	42
Anotación taxonómica de los plásmidos en muestras de <i>A. germinans</i> y <i>R. mangle</i> .	43
Anotación funcional de plásmidos en muestras de <i>A. germinans</i> y <i>R. mangle</i> .	44
Anotación estructural y funcional del plásmido pZW5A de <i>Lactiplantibacillus plantarum</i> de muestras foliares de <i>R. mangle</i> .	46
Discusión.	46
Análisis taxonómico y de diversidad del plasmidoma de <i>A. germinans</i> y <i>R. mangle</i> .	46
Análisis funcional de los plásmidos parcialmente reconstruidos en muestras de <i>A. germinans</i> y <i>R. mangle</i> .	48
Análisis funcional y estructural del plásmido reconstruido pZW5A de <i>Lactiplantibacillus</i> .	49
Capítulo 4: Discusión y conclusiones generales.	52
Discusión general.	52
Conclusiones generales.	53

Perspectivas y recomendaciones.	54
Aspectos éticos.	54
Literatura citada.	55
Anexos	82

Viomas y plasmidomas asociados a tejido foliar y radicular de *Rhizophora mangle* L. (*Rhizophoraceae*) y *Avicennia germinans* L. (*Acanthaceae*) en la Bahía de Buenaventura, Colombia.

Resumen general.

El ecosistema del manglar, caracterizado por el predominio de vegetación arbórea, se encuentra expuesto a condiciones fisicoquímicas de alta variabilidad espacio-temporal, propias de los estuarios. Este ecosistema se distingue por la provisión de servicios ecosistémicos y su considerable impacto ecológico, que contribuye a la regulación de los ciclos biogeoquímicos, por lo cual, resulta relevante comprender las dinámicas ecológicas que gobiernan su establecimiento y supervivencia. La interacción planta-microorganismo desempeña un papel crucial en la modulación de diversos parámetros de la capacidad adaptativa vegetal. Dichas interacciones se pueden ver afectadas por elementos genéticos móviles (MGEs), tales como virus y plásmidos, que participan en la regulación de la composición del microbioma y le confieren genes que amplían su capacidad metabólica. No obstante, hasta ahora, no se han documentado estudios sobre estos MGEs en microorganismos endófitos de manglar. Esta investigación tuvo como objetivo caracterizar el viroma y plasmidoma en tejidos foliares y radiculares de *Rhizophora mangle* y *Avicennia germinans*, dos especies de mangle muy importantes en la Bahía de Buenaventura, Colombia, utilizando un enfoque metagenómico.

El análisis de la comunidad viral reveló 2137 especies virales, pertenecientes a 166 familias. Los géneros *Andhravirus* y *Valbivirus* fueron predominantes en la comunidad viral de ambos tejidos y especies, mientras que el género *Mementomorivirus* fue exclusivo de los tejidos foliares de *A. germinans*. Los tejidos radiculares mostraron mayor riqueza y diversidad viral en comparación con los tejidos foliares, lo que sugiere un efecto significativo del tipo de tejido en la composición viral. Por otra parte, se identificaron genomas virales reconstruidos con genes auxiliares metabólicos (AMGs) asociados a funciones como colonización vegetal, exclusión competitiva y respuesta al estrés abiótico. Estos AMGs codifican enzimas relacionadas con el metabolismo de carbohidratos (CAZimas), proteínas de choque térmico, sistemas toxina-antitoxina y enzimas asociadas al metabolismo de nucleósidos, lo que indica un potencial rol de los virus en la modulación de la capacidad adaptativa de comunidades bacterianas endófitas.

El plasmidoma mostró una alta abundancia de plásmidos asociados a géneros bacterianos como *Chlamydia*, *Bacillus*, *Lactiplantibacillus* y *Streptomyces*. Los plásmidos de *Azospirillum* y *Rhizobium* se detectaron exclusivamente en tejidos radiculares, lo que refleja el papel de los plásmidos en la interacción planta-microorganismo en entornos rizosféricos. Se reconstruyeron 5 plásmidos parcialmente

reconstruidos y un plásmido completo. El plásmido completo pZW5A de *Lactiplantibacillus*, contenía genes funcionales como integrasas, sistemas toxina-antitoxina y enzimas asociadas al metabolismo de carbohidratos. Por su parte, los plásmidos reconstruidos portaban genes relacionados con la homeostasis del hierro, halotolerancia y biosíntesis de vitaminas, lo que sugiere su contribución a la adaptación bacteriana y a la interacción planta-bacteria. La diversidad del plasmidoma también estuvo influenciada por el tipo de tejido, destacando la importancia de las condiciones ambientales locales en la composición de estos MGEs.

Este estudio demostró que el tipo de tejido es un factor clave en la composición y diversidad del viroma y plasmidoma endófitos en *R. mangle* y *A. germinans*, superando incluso las diferencias entre especies. Los virus y plásmidos identificados actúan como reservorios de genes funcionales que facilitan la adaptación de las plantas y su microbiota a condiciones estresantes. Los hallazgos resaltan el potencial biotecnológico de estos MGEs, especialmente en aplicaciones agrícolas sostenibles, como el desarrollo de inoculantes microbianos y estrategias de control biológico.

Palabras clave: elementos genéticos móviles, manglar, metagenómica, microorganismos endófitos, plásmidos, transferencia horizontal de genes, virus.

General abstract.

The mangrove ecosystem, characterized by the predominance of tree vegetation, is exposed to highly variable spatiotemporal physicochemical conditions typical of estuaries. This ecosystem is distinguished by the provision of ecosystem services and its considerable ecological impact, which contributes to the regulation of biogeochemical cycles. Therefore, it is important to understand the ecological dynamics that govern its establishment and survival. Plant-microorganism interactions play a crucial role in modulating various parameters of plant adaptive capacity. These interactions can be affected by mobile genetic elements (MGEs), such as viruses and plasmids, which participate in regulating the composition of the microbiome and confer genes that expand its metabolic capacity. However, to date, no studies on these MGEs in mangrove endophytic microorganisms have been documented. This research aimed to characterize the virome and plasmidome in leaf and root tissues of *Rhizophora mangle* and *Avicennia germinans*, two very important mangrove species in Buenaventura Bay, Colombia, using a metagenomic approach.

Analysis of the viral community revealed 2,137 viral species belonging to 166 families. The genera *Andhravirus* and *Valbivirus* were predominant in the viral community of both tissues and species, while the genus *Mementomorivirus* was exclusive to the leaf tissues of *A. germinans*. Root tissues showed greater viral richness and diversity compared to leaf tissues, suggesting a significant effect of tissue type on viral composition. Furthermore, reconstructed viral genomes with metabolic auxiliary genes (MAGs) associated with functions such as plant colonization, competitive exclusion, and abiotic stress response were identified. These MAGs encode enzymes related to carbohydrate metabolism (CAZymes), heat shock proteins, toxin-antitoxin systems, and enzymes associated with nucleoside metabolism, indicating a potential role for viruses in modulating the adaptive capacity of endophytic bacterial communities.

The plasmidome showed a high abundance of plasmids associated with bacterial genera such as *Chlamydia*, *Bacillus*, *Lactiplantibacillus*, and *Streptomyces*. *Azospirillum* and *Rhizobium* plasmids were detected exclusively in root tissues, reflecting the role of plasmids in plant-microorganism interaction in rhizospheric environments. Five partially reconstructed plasmids and one complete plasmid were reconstructed. The complete *Lactiplantibacillus* plasmid pZW5A contained functional genes such as integrases, toxin-antitoxin systems, and enzymes associated with carbohydrate metabolism. The reconstructed plasmids carried genes related to iron homeostasis, halotolerance, and vitamin biosynthesis, suggesting their contribution to bacterial adaptation and plant-bacteria interaction. Plasmidome diversity

was also influenced by tissue type, highlighting the importance of local environmental conditions in the composition of these MGEs.

This study demonstrated that tissue type is a key factor in the composition and diversity of the endophytic virome and plasmidome in *R. mangle* and *A. germinans*, even surpassing differences between species. The identified viruses and plasmids act as reservoirs of functional genes that facilitate the adaptation of plants and their microbiota to stressful conditions. The findings highlight the biotechnological potential of these MGEs, especially in sustainable agricultural applications, such as the development of microbial inoculants and biological control strategies.

Key words: mobile genetic elements, mangrove, metagenomics, endophytic microorganisms, plasmids, horizontal gene transfer, viruses.

Capítulo 1: Introducción y métodos generales.

Introducción general.

Los manglares son ecosistemas costeros de gran relevancia ecológica, que se extienden a lo largo de las zonas litorales tropicales y subtropicales, cubriendo aproximadamente una cuarta parte de estas regiones a nivel global (Ranjan *et al.*, 2012). Estos ecosistemas estuarinos, caracterizados por una alta variabilidad espaciotemporal en sus condiciones fisicoquímicas, como salinidad, pH, disponibilidad de nutrientes, inundación, elevación y luz solar (Joshi & Ghose, 2003; Friess, 2016), albergan una vegetación altamente adaptada, dominada por árboles de mangle pertenecientes a diversas familias de plantas (Kathiresan, 2012; Selvam, 2007).

La importancia de los manglares radica en su provisión de múltiples servicios ecosistémicos, esenciales para el sustento de las comunidades humanas circundantes, incluyendo servicios de aprovisionamiento, regulación, apoyo y culturales (Kathiresan, 2012; Friess, 2016). Igualmente, importante, este ecosistema genera un fuerte impacto ecológico en el mantenimiento de los ciclos biogeoquímicos del carbono, nitrógeno y fósforo, lo que lo convierte en un sumidero de carbono en entornos terrestres y acuáticos (Palit *et al.*, 2022).

Un componente crucial, en la regulación de estos ciclos biogeoquímicos en un bosque de manglar, es el microbioma, definido como la comunidad de microorganismos (bacterias, arqueas, hongos, protozoos, algas y virus), sus genomas, sus interacciones, sus funciones y los productos asociados a un ambiente u organismo específico (Berg *et al.*, 2020). En particular, el microbioma asociado a plantas juega un papel fundamental en la modulación de diversos aspectos de la capacidad adaptativa de las plantas, asociados a la germinación, el crecimiento, la floración, la resistencia a patógenos y la adaptación a condiciones ambientales adversas (Rosenblueth & Martínez-Romero, 2006; Reinhold-Hurek & Hurek, 2011).

Dentro del microbioma, las interacciones entre los microorganismos pueden estar influenciadas por diversos factores, incluyendo la presencia de múltiples elementos genéticos móviles (MGEs, por sus siglas en inglés), como los virus y los plásmidos, que facilitan la transferencia horizontal de genes (HGT, por sus siglas en inglés) al interior de las comunidades microbianas (Tokuda & Shintani, 2024). La HGT es un proceso que permite la propagación de regiones genéticas de manera horizontal en una población o comunidad (Niehus *et al.*, 2015). En comunidades microbianas, especialmente bacterianas, la HGT desempeña un papel fundamental en el aumento de la velocidad de adaptación de estos microorganismos a su entorno, influyendo directamente en las tasas de evolución y en la capacidad

metabólica de cepas y especies bacterianas (Arnold *et al.*, 2022; Tokuda & Shintani, 2024). Un ejemplo de MGE, son los virus (Dolja, Krupovic & Koonin, 2020), los cuales, a través de sus fases lítica y lisogénica, pueden regular la composición y función de las comunidades microbianas, al aportar genes metabólicos auxiliares (AMGs, por sus siglas en inglés) que modulan la respuesta de los microorganismos a su entorno, como también, efectúan el control biológico sobre comunidades bacterianas (Luo *et al.*, 2022; Pratama *et al.*, 2020).

Además de los virus, otros MGE que juegan un papel importante en la modulación de las interacciones microbianas, son los plásmidos, estos son elementos genéticos extracromosómicos presentes en las bacterias (Cai *et al.*, 2023). Los plásmidos, al portar genes de resistencia a antibióticos, virulencia y otras funciones adaptativas, pueden conferir ventajas selectivas a las bacterias, alterando la diversidad y función del microbioma (Smalla & Sobecky, 2002; Newbury *et al.*, 2022; Fondi *et al.*, 2010).

A pesar de la importancia ecológica y socioeconómica de los manglares, y del papel crucial del viroma y plasmidoma en las dinámicas microbianas y en la salud de las plantas bajo estrés, el conocimiento sobre estos componentes microbianos en este ecosistema es limitado. Hasta la fecha, alrededor de tres estudios han explorado el viroma del suelo y sedimentos de manglar, mostrando su alta diversidad y potencial impacto en los ciclos biogeoquímicos del carbono, fósforo y nitrógeno, como su efecto modulador sobre comunidades microbianas (Jin *et al.*, 2019; Su *et al.*, 2024; Zhu *et al.*, 2024).

En este contexto, el presente estudio tiene como objetivo caracterizar y comparar el viroma de ADN y plasmidoma del tejido foliar y radicular de *Rhizophora mangle* L. y *Avicennia germinans* L., dos especies de mangle predominantes en el bosque de manglar ubicado en la desembocadura del río Dagua, Bahía de Buenaventura, Colombia. A nivel general, esta investigación busca comprobar que la composición, diversidad y funciones génicas putativas del viroma y plasmidoma endófito difieren significativamente entre tipos de tejido y especies vegetales analizadas. Esta investigación se enmarca en el macroproyecto “Investigación para el aprovechamiento sostenible de la biodiversidad de los microorganismos asociados al ecosistema de manglar y la exploración de su potencial biotecnológico en la costa pacífica del Valle del Cauca”, financiado por el Sistema General de Regalías (SGR) y la Universidad del Valle.

Objetivos.

Objetivo general.

Caracterizar y comparar el viroma de ADN y plasmidoma del tejido foliar y radicular de *Rhizophora mangle* L. y *Avicennia germinans* L., dos especies de mangle predominantes en el bosque de manglar ubicado en la desembocadura del río Dagua, Bahía de Buenaventura, Colombia..

Objetivos específicos.

- Caracterizar la composición, diversidad y funciones genéticas putativas de la comunidad viral y del viroma endófito presentes en el tejido foliar y radicular de *R. mangle* y *A. germinans*.
- Dilucidar la composición, diversidad y funciones genéticas putativas del plasmidoma endófito presente en tejido foliar y radicular de *R. mangle* y *A. germinans*, mediante un enfoque metagenómico.

Hipótesis.

La composición, diversidad y funciones génicas putativas del viroma y plasmidoma endófito difieren significativamente entre tipos de tejido y especies vegetales analizadas.

Métodos generales.

Toma de muestras.

El tejido foliar y radicular de *Rhizophora mangle* y *Avicennia germinans* se recolectó en el manglar ubicado en la desembocadura del río Dagua, en la bahía de Buenaventura (Colombia), con coordenadas 3°52'44.387" N, 77°3'44.577" W, el día 14 de junio de 2023 (Anexo 1). Se identificaron tres plantas de cada especie, de las cuales se tomaron hojas y raíces sanas, sin daños ni enfermedades aparentes. Los tejidos fueron tratados con un protocolo de lavado de alcohol etílico al 70 % durante un minuto, seguido de cinco minutos con hipoclorito de sodio al 3 %, esto, con el fin de eliminar los microorganismos de la superficie de los tejidos y capturar, solamente los endófitos (Sahu *et al.*, 2022). En el caso del mangle de la especie *A. germinans*, la muestra de tejido radicular se obtuvo a partir de los neumatóforos. Inmediatamente después de ser recolectadas, las muestras fueron transportadas en nitrógeno líquido hasta el Laboratorio de Biología Molecular de la Universidad del Valle en Cali, donde se preservaron hasta su uso a -80 °C.

Por otra parte, para conocer las características fisicoquímicas del entorno, se tomaron muestras de agua intersticial y del suelo en la zona de muestreo, cuando la marea se encontraba en su nivel más bajo (1,39 m), entre las 8:00 A.M. y las 12:00 M. En el agua intersticial se midieron el pH, el potencial de óxido-reducción, la saturación de oxígeno, el oxígeno disuelto, la conductividad, la resistividad, los sólidos totales disueltos, la salinidad (PSU), la turbidez y la temperatura mediante una sonda multiparamétrica. Las muestras de suelo fueron enviadas al Laboratorio de Aguas y Suelos Agrícolas (LASA) de la Universidad del Valle para realizar las mediciones de pH, conductividad eléctrica, porcentaje de materia orgánica, nitrógeno total, potasio, calcio, magnesio, sodio, fósforo y azufre (Anexo 2).

Extracción de ácidos nucleicos.

La extracción del ADN, a partir de los tejidos conservados a -80 °C, se realizó empleando el kit “InviSorb® Spin Plant Mini Kit” (Invitex), con el protocolo sugerido por el fabricante modificado de la siguiente manera: Se emplearon 100 mg de tejido inicial y, al adicionar los 400 µL de buffer de lisis, se realizó una segunda maceración utilizando un DREMEL con un pistilo esterilizado. El ADN obtenido se cuantificó mediante el fluorómetro Qubit 4 (Thermo Fisher Scientific), para lo cual se empleó el kit “dsDNA BR Assay Kits” (Thermo Fisher Scientific). Posteriormente, se realizó una electroforesis en gel de agarosa al 0.8% para probar la integridad del ADN extraído.

Identificación molecular de las plantas hospederas.

Con el fin de confirmar que se tratara de *A. germinans* y *R. mangle*, se realizó una identificación molecular de las plantas de las que provenía cada tejido. Para ello, se utilizaron los cebadores descritos por Cerón-Souza et al. (2012), y se amplificó el sistema psbJ-petA para *A. germinans* y los sistemas atpI-atpH y psbJ-petA para *R. mangle*. A continuación, los amplicones obtenidos fueron enviados a la empresa Macrogen, en Corea del Sur, para realizar la secuenciación Sanger. Los productos amplificados se mapearon contra la base de datos del NCBI para confirmar que los individuos colectados, seis en total, pertenecían a la especie correspondiente.

Construcción de librerías y secuenciación.

Las muestras de ADN extraídas se enviaron a Macrogen (Corea del Sur). Se tuvo en cuenta que la cantidad de ADN de cada muestra fuera de al menos 1,0 µg. A continuación, se procedió a la construcción de las librerías, para lo cual, se empleó el kit “TruSeq DNA nano kit”. Finalmente, se utilizó la plataforma NovaSeq6000 (Illumina) para realizar una secuenciación de fragmentos de 150 pb, que generó al menos 20 GB de información por cada una de las muestras.

Análisis bioinformático.

- Limpieza de secuencias de baja calidad.

A las librerías de cada muestra se les realizó un análisis de calidad con el software FastQC v.0.12.0 (2010), y mediante Trimmomatic v.0.39 (Bolger, Lohse & Usadel, 2014), se eliminaron las lecturas con un puntaje PHRED <30.

- Eliminación del genoma de la planta hospedera.

Para obtener únicamente las secuencias asociadas a la microbiota, se sustrajeron las lecturas del genoma del hospedero vegetal correspondiente (*Rhizophora mangle* o *Avicennia germinans*) empleando como referencia los genomas ensamblados de ambas especies y obtenidos en esta investigación. Para este proceso se realizó un alineamiento local con Bowtie2 v.2.5.4 (Langmead & Salzberg, 2012), y se procedió con la recuperación de los archivos SAM empleando Samtools v.1.21 (Li *et al.*, 2009).

Las lecturas limpias se encuentran consignadas en la base de datos del NCBI, en el Bioproject PRJNA1215819 (Anexo 4).

Capítulo 2: Comunidad viral y viomas endófitos de ADN del tejido foliar y radicular de *R. mangle* y *A. germinans*.

Introducción.

El estudio de los viomas se refiere a la caracterización de las poblaciones virales presentes en una muestra, organismo o ecosistema, así como su contenido genómico y las interacciones biológicas que los involucran (McDaniel *et al.*, 2008). Tradicionalmente, los estudios virales han empleado técnicas dependientes de cultivo *in vitro*, que requieren de un huésped que favorezca la reproducción de los virus, seguido por el aislamiento de los viriones y una detección por anticuerpos específicos para la identificación del virus cultivado (Hierholzer & Killington, 1996; Van Oirschot, 1999). Sin embargo, los métodos tradicionales limitan la detección de virus nuevos y no cultivables o de difícil cultivo en organismos modelo (Smits *et al.*, 2015, Plyusnin *et al.*, 2020).

Hacia la década de los 2000, con la aparición de la metagenómica, se planteó el uso de esta nueva metodología para el estudio de viomas, lo cual fue posible con diversos avances en el desarrollo de las técnicas de secuenciación, en los algoritmos para el tratamiento de secuencias y con el fortalecimiento de las bases de datos genómicas; lo que permitió un óptimo filtrado, análisis e identificación de las muestras virales (Plyusnin *et al.*, 2020). Entre los estudios pioneros en viomas metagenómicos, se encuentra el realizado por Breitbart *et al.* (2002), en muestras oceánicas, donde se identificaron en total 1934 secuencias, de las cuales al menos el 65% no habían sido reportadas hasta ese momento, demostrando el potencial que tiene el enfoque metagenómico independiente de cultivo, para comprender con mayor detalle, la diversidad de las comunidades virales.

Guiados por los resultados sobresalientes de los estudios de viomas oceánicos, hacia el 2003 se formuló el proyecto “Viroma humano”, con el fin de aislar, caracterizar y secuenciar los agentes virales asociados al ser humano alrededor del mundo, dando prioridad a especies patogénicas (Anderson *et al.*, 2003). No obstante, solo hasta el año 2005, se retomaron los estudios en viomas humanos, y se propuso a los métodos moleculares de detección de virus como viables (Allander *et al.*, 2005).

De forma paralela, la investigación en viomas metagenómicos vegetales, se inició con el estudio de Coetzee *et al.* (2010), en viñedos sudafricanos enfermos, donde identificaron cuatro especies virales conocidas, asociadas a esta condición y al menos tres virus de los cuales no había reportes anteriores de su infección a la vid. En la siguiente década, se presentó un auge en el número de estudios en viomas vegetales con enfoques metagenómicos y metatranscriptómicos, en los que se caracterizó el viroma de

plantas modelo y de interés comercial pertenecientes a los géneros *Solanum* (Gutiérrez-Sánchez *et al.*, 2014), *Capsicum* (Alabi *et al.*, 2015), *Ipomoea* (Gu *et al.*, 2014), *Arachis* (Sánchez *et al.*, 2016), *Carica* (Alabi *et al.*, 2016), y *Theobroma* (Chingandu *et al.*, 2017).

En los últimos años, se ha evidenciado un creciente interés en el estudio de viomas y virósferas asociados a ecosistemas como el manglar, donde sobresalen, principalmente, los estudios sobre muestras de sedimentos, en los cuales se ha mostrado el papel de los virus como moduladores de las comunidades bacterianas, bajo condiciones ambientales cambiantes (Zhu *et al.*, 2024), y los estudios sobre efectos de la cobertura vegetal y la acuicultura sobre el papel ecológico desempeñado por los virus (Su *et al.*, 2024). Sin embargo, aunque se ha establecido la importancia de los virus en los suelos de manglar para la regulación de las comunidades bacterianas, aún no se ha caracterizado el viroma asociado a las plantas de mangle ni explorado su potencial para modificar las interacciones planta-microbioma. Esta brecha en el conocimiento resulta particularmente relevante para comprender la ecología microbiana de plantas que se desarrollan bajo condiciones de estrés.

Este capítulo tiene como objetivo mostrar y discutir los resultados de la caracterización del viroma endófito presente en el tejido foliar y radicular de dos especies dominantes de mangle del Pacífico colombiano, *Rhizophora mangle* y *Avicennia germinans*, utilizando un enfoque metagenómico. Con análisis bioinformáticos específicos para virus, se determinó la composición de la comunidad viral, el ensamblaje de genomas virales parcialmente reconstruidos y sus funciones putativas y, así, analizar el papel ecológico que tiene esta comunidad en el desarrollo del microbioma en los tejidos señalados y su posible rol en la fisiología y adaptación de estas plantas a su entorno.

Objetivo específico.

Caracterizar la composición, diversidad y funciones genéticas putativas de la comunidad viral y del viroma endófito presentes en el tejido foliar y radicular de *R. mangle* y *A. germinans*.

Métodos.

Sumados a los métodos generales descritos en el capítulo 1, en el capítulo 2 se procede con los métodos específicos para el análisis de las comunidades y genomas virales parcialmente reconstruidos.

Ensamblaje de secuencias limpias.

Luego, de efectuar la limpieza de las librerías de secuencias de baja calidad y de lecturas del genoma del hospedero vegetal (*A. germinans* o *R. mangle*), se realizó el ensamblaje de las lecturas por muestra. El ensamblaje de las lecturas limpias se efectúa con el pipeline SPAdes v.4.0.0, empleando el programa metaSPADES (Nurk *et al.*, 2017), para ensamblar las librerías metagenómicas.

Anotación taxonómica del viroma asociado a cada tejido y a cada especie de mangle.

Para la anotación taxonómica de las especies virales asociadas a las librerías de ADN, se realizó la identificación a partir de las lecturas limpias. Para este fin se empleó la suite de Kraken2 v.2.1.4 (Wood *et al.*, 2019), junto con la base de datos de virus asociada a esta herramienta bioinformática. Una vez obtenida la anotación taxonómica, se procedió a construir un archivo BIOM con los reportes producidos, empleando Kraken-biom v.1.0.1 (Dabdou, 2016). La visualización se hizo mediante los paquetes vegan v.2.7-0 (Oksanen *et al.*, 2025) y ggplot2 v. 3.5.1 (Wickham, 2016) del repositorio CRAN v.2.6-10 (Hornik, 2012) de R v.4.4.2 (R Core Team, 2025)

Análisis estadísticos y de diversidad alfa y beta.

Una vez obtenida la anotación taxonómica, se importaron los archivos BIOM en el software R v.4.4.2 (R Core Team, 2025) y, con la ayuda del paquete phyloseq v.3.20 (McMurdie & Holmes, 2013), se realizaron análisis comparativos de la diversidad alfa mediante los índices de Shannon-Wiener, Chao1, ACE y Simpson. A continuación, se evaluó la diversidad beta mediante análisis estadísticos como el análisis de coordenadas principales (PCoA) y el análisis de varianza multivariado permutacional (PERMANOVA) del paquete vegan v.2.7-0 (Oksanen *et al.*, 2025) (bajo el supuesto de homogeneidad de varianzas evaluado en el Anexo 18).

Detección, filtrado y anotación de genomas virales reconstruidos por especie de planta.

Para la identificación de los genomas virales ensamblados a partir de metagenomas (vMAG, por sus siglas en inglés), se empleó el pipeline VIRify, en su versión 0.4.0 (Rangel-Pineros *et al.*, 2023). Este pipeline fue útil para detectar, filtrar y anotar los vMAGs parcialmente reconstruidos provenientes de las

secuencias metagenómicas ensambladas con metaSPADES (Nurk *et al.*, 2017). La predicción y detección inicial de los vMAGs parcialmente reconstruidos se realizó integrando herramientas de anotación viral, como VirSorter (Roux *et al.*, 2015), VirFinder (Ren *et al.*, 2017) y PPRMeta (Fang *et al.*, 2019), para luego finalizar con el análisis de completación y contaminación con CheckV (Nayfach *et al.*, 2021). Este procedimiento permitió obtener vMAGs parcialmente reconstruidos de media y baja calidad, los cuales fueron clasificados dependiendo de parámetros de calidad (completación y contaminación) obtenidos a partir de CheckV. Los genomas virales parcialmente reconstruidos con valores de completación entre 50-90 % y con porcentajes de contaminación de hasta el 20 %, fueron categorizados como vMAGs parcialmente reconstruidos de calidad media. Por su parte, los genomas virales parcialmente reconstruidos con un porcentaje de completación menor al 50 % o con un porcentaje de contaminación superior al 20 % fueron clasificados como vMAGs parcialmente reconstruidos de calidad baja.

La anotación taxonómica de los vMAGs parcialmente reconstruidos, se efectuó con el pipeline VIRify v.0.4.0 (Rangel-Pineros *et al.*, 2023) y la suite de Kraken2 v.2.1.4 (Wood *et al.*, 2019). La anotación de los genomas virales parcialmente reconstruidos con VIRify, se realizó con el filtrado de proteínas virales usando Prodigal v.2.6.3 (Hyatt *et al.* 2010), BLASTn (Altschul *et al.*, 1990) y hmmscan v.3.2.1 (Finn *et al.*, 2011), y la base de datos ViPhOG (Moreno-Gallego & Reyes, 2021). Y la anotación con la suite de Kraken2, se realizó con el software Kraken2 v.2.1.4 (Wood *et al.*, 2019) y la base de datos de virus asociada a esta herramienta bioinformática.

Anotación de genes auxiliares metabólicos en los vMAGs reconstruidos y ensamblados por planta.

Para la anotación de genes auxiliares metabólicos en los vMAGs parcialmente reconstruidos ensamblados por planta, se emplearon dos herramientas bioinformáticas, el pipeline de anotación de Rasgos de Promoción del Crecimiento Vegetal (PGPT), PGPg finder v.1.1.0 (Pellegrinetti *et al.*, 2024) y el software run_dbcan v.4.1.4 (Zheng, J., *et al.*, 2023).

El pipeline PGPg finder v.1.1.0 (Pellegrinetti *et al.*, 2024), con la base de datos PLant-associated BActeria web resource (PLaBAse) v.1.02 (Patz *et al.*, 2021), se emplearon para la identificación de genes auxiliares metabólicos asociados al metabolismo bacteriano, en especial al metabolismo de compuestos nitrogenados, carbohidratos, lípidos y metales pesados, también incluyendo genes involucrados en la respuesta a estrés biótico y abiótico.

Por otra parte, el software run_dbcan v.4.1.4 (Zheng, J., *et al.*, 2023), asociado a la base de datos dbCAN3 (Zheng, J., *et al.*, 2023), fue implementado para la identificación de genes auxiliares metabólicos que expresan enzimas activas en el metabolismo de carbohidratos (CAZimas) en los genomas virales parcialmente reconstruidos ensamblados por planta.

Para la visualización de los resultados de ambas herramientas bioinformáticas empleadas para la identificación de genes auxiliares metabólicos, se empleó software R v.4.4.2 (R Core Team, 2025), con los paquetes ggplot2 v. 3.5.1 (Wickham, 2016) y pheatmap v.1.0.12 (Kolde, 2015), del repositorio CRAN v.2.6-10 (Hornik, 2012).

Anotación estructural y funcional de los vMAGs reconstruidos.

Para realizar la anotación estructural y funcional de los vMAGs parcialmente reconstruidos, y su visualización, se empleó el portal Proksee v.1.0.1 (Grant *et al.*, 2023). Primero, se realizó la anotación de los CDS presentes en cada contig viral con el software Prokka v.1.2.0 (Seemann, 2014). Posteriormente, se predijeron las regiones susceptibles de transferencia horizontal (HTG, por sus siglas en inglés) con el software Alien Hunter v.1.3.0 (Vernikos & Parkhill, 2006) y se identificaron grupos de ortólogos móviles asociados a elementos genéticos móviles (MGEs, por sus siglas en inglés) con mobileOG-db v.1.1.3 (Brown *et al.*, 2022), para finalmente identificar, anotar y visualizar secuencias asociadas a fagos con la herramienta PHASTEST v.3.0 (DB: 22/12/2020) (Wishart *et al.*, 2023).

Resultados.

Anotación taxonómica de la comunidad viral.

La anotación taxonómica de las librerías metagenómicas reveló la presencia de 2137 especies virales, agrupadas en 1245 géneros y 166 familias (Figura 1). A nivel de género, los clados predominantes en todas las muestras, con abundancias relativas superiores al 10 %, fueron *Andhravirus* y *Valbivirus*. Entre ellos, *Andhravirus* presentó las mayores abundancias, superando el 50 % en las muestras foliares de *Rhizophora mangle* y *Avicennia germinans*. Por otro lado, *Oryzopoxvirus* se detectó principalmente en muestras radiculares de *A. germinans*, así como en ambos tipos de tejido de *R. mangle*. Además, *Betabaculovirus* se encontró presente en las comunidades virales de las muestras radiculares de ambas

especies vegetales. Por otra parte, se identificaron lecturas relacionadas al género *Mementomorivirus*, asociadas a muestras foliares de *A. germinans*, de manera selectiva.

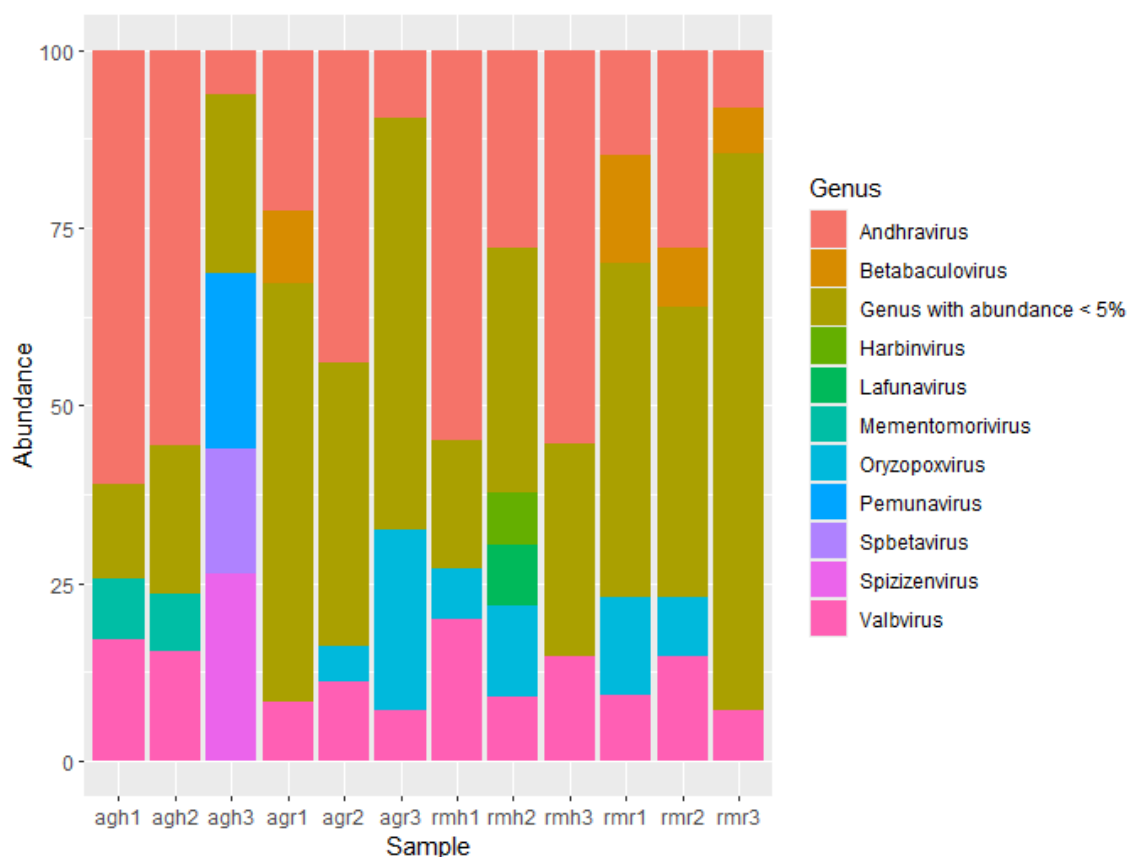


Figura 1. Abundancia relativa de los géneros de virus de ADN más representativos, en tejido de hojas y raíces de *A. germinans* y *R. mangle*. Cada barra representa la composición y la abundancia relativa (%) de virus identificados en las muestras de hojas (h) y raíces (r) de *A. germinans* (ag) y *R. mangle* (rm). Los géneros con una abundancia menor al 5 % fueron agrupados en la categoría "Genus with abundance < 5 %"

Análisis de diversidad alfa.

Se evaluó la diversidad alfa de las comunidades virales presentes en muestras de tejido foliar y radicular de *Avicennia germinans* y *Rhizophora mangle* mediante los índices Chao1, ACE, Shannon-Wiener y Simpson (Figura 2).

En el análisis de librerías metagenómicas (Figura 2), se evidenció una marcada diferencia entre los tejidos. Las muestras de raíz presentaron una riqueza considerablemente mayor (Chao1: 499–2500; ACE: 523–2497) en comparación con las muestras foliares (Chao1 y ACE entre 40 y 100). Asimismo, los valores de diversidad y equidad fueron más elevados en las muestras radiculares (Shannon-Wiener: 3.1–5.5; Simpson: 0.84–0.98).

No obstante, las curvas de rarefacción no llegaron a la meseta (Anexo 5), lo que demuestra que es necesario ampliar la profundidad de secuenciación para analizar, en su totalidad, la diversidad de la comunidad viral asociada a tejidos foliares y radiculares de *A. germinans* y *R. mangle*.

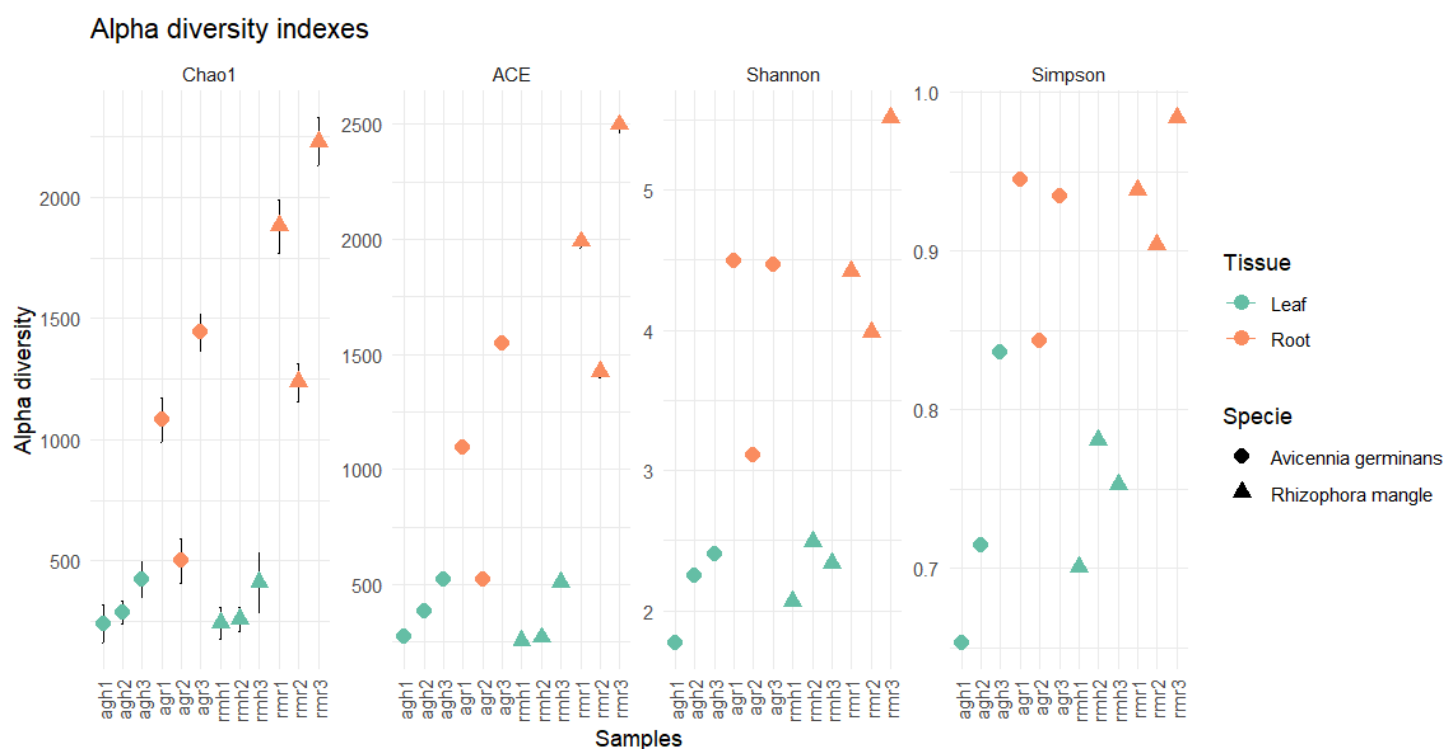


Figura 2. Diversidad alfa asociada a la comunidad viral de tejido foliar y radicular de *A. germinans* y *R. mangle*. Los índices Chao1, ACE, Shannon y Simpson fueron calculados a partir de secuencias de virus. Los colores indican el tipo de tejido (hoja o raíz) y las formas, la especie vegetal (*A. germinans* o *R. mangle*).

Análisis estadísticos y de diversidad beta.

Para analizar la similitud entre las muestras endófitas respecto a su composición taxonómica viral, se aplicó un Análisis de Coordenadas Principales (PCoA). En este análisis de las librerías metagenómicas (Figura 4), se evidenció un único patrón de agrupación a lo largo del eje horizontal, correspondiente al tipo de tejido de origen. Esta tendencia fue respaldada por el análisis PERMANOVA (Anexo 6), el cual mostró un efecto significativo únicamente para el factor tejido ($p = 0.002$), sin diferencias significativas atribuibles a la especie vegetal ni a su interacción con el tejido.

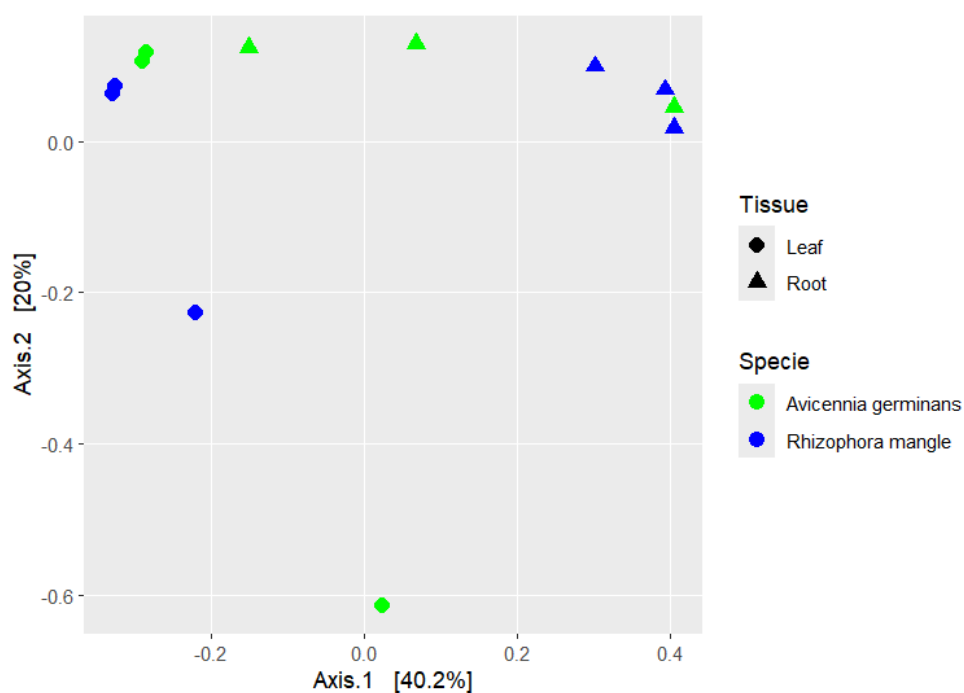


Figura 4. Análisis de coordenadas principales (PCoA) de la comunidad de virus de ADN asociados a tejido foliar y radicular de *A. germinans* y *R. mangle*. Se evidencia el efecto del tipo de tejido (foliar y radicular), y de la especie de planta (*A. germinans* y *R. mangle*) en la composición de las comunidades virales asociadas a cada muestra.

Anotación taxonómica y análisis de calidad de los vMAGs reconstruidos y filtrados con VIRify

Luego del ensamblaje de las secuencias metagenómicas, se empleó el pipeline VIRify para la detección, filtrado, anotación y análisis de calidad de los genomas virales parcialmente reconstruidos por especie de planta. Este procedimiento resultó en la identificación de 6 vMAGs parcialmente reconstruidos en muestras de *A. germinans* (Tabla 1) y 14 vMAGs parcialmente reconstruidos en muestras de *R. mangle* (Tabla 1), los cuales fueron clasificados taxonómicamente con los pipelines VIRify y Kraken2 (Tabla 1).

Como resultado de los análisis de calidad del pipeline VIRify, los vMAGs reconstruidos identificados en muestras de *A. germinans*, se clasificaron en 5 vMAGs parcialmente reconstruidos de calidad media y un contig de calidad baja (Tabla 1). Por su lado, los genomas virales reconstruidos e identificados en muestras de *R. mangle*, se clasificaron en 9 vMAGs parcialmente reconstruidos de calidad media y 5 vMAGs parcialmente reconstruidos de calidad baja (Tabla 1).

Tabla 1. Anotación taxonómica y medidas de calidad de los vMAGs parcialmente reconstruidos a partir de librerías metagenómicas, y anotados a través de VIRify y Kraken2.

Especie	Contig ID	Categoría taxonómica asignada	CheckV quality	MIUViG quality	CC	GV/GT
<i>A. germinans</i>	NODE_92_length_638 22_cov_20.496479	<i>Howeivirus</i>	Medium-quality	Genome-fragment	76.26 / 0.0	15 / 106
<i>A. germinans</i>	NODE_244_length_42 930_cov_4.574671	<i>Tunavirinae</i>	Medium-quality	Genome-fragment	80.56 / 0.0	9 / 47
<i>A. germinans</i>	NODE_338_length_37 207_cov_10.523175	<i>Aresaunavirus</i> RSA1	Medium-quality	Genome-fragment	78.59 / 0.0	26 / 58
<i>A. germinans</i>	NODE_341_length_37 118_cov_4.696112	<i>Klebsiella</i> phage LASTA	Medium-quality	Genome-fragment	58.0 / 0.0	30 / 57
<i>A. germinans</i>	NODE_80_length_681 56_cov_33.310451	<i>Bacillus</i> phage PM1*	Medium-quality	Genome-fragment	81.36 / 19.07	29 / 95
<i>A. germinans</i>	NODE_2570_length_1 4003_cov_32.072268	<i>Arquatrovirinae</i>	Low-quality	Genome-fragment	30.77 / 0.0	11 / 16
<i>R. mangle</i>	NODE_29_length_989 59_cov_6.234975	<i>Rhodococcus</i> phage Weasels2	Medium-quality	Genome-fragment	78.76 / 0.0	28 / 135
<i>R. mangle</i>	NODE_39_length_867 05_cov_32.529856_11 2	Viruses	Medium-quality	Genome-fragment	89.13 / 0.0	19 / 161
<i>R. mangle</i>	NODE_74_length_719 96_cov_6.814181	Viruses	Medium-quality	Genome-fragment	86.16 / 0.0	21 / 96
<i>R. mangle</i>	NODE_107_length_63	<i>Caudoviricetes</i>	Medium-qu	Genome-fra	76.95 / 0.0	19 / 96

	048_cov_8.605337		ality	gment		
<i>R. mangle</i>	NODE_138_length_56 770_cov_12.767363	<i>Caudoviricetes</i>	Medium-quality	Genome-fragment	74.89 / 0.0	18 / 83
<i>R. mangle</i>	NODE_360_length_41 263_cov_46.921544	<i>Caudoviricetes</i>	Medium-quality	Genome-fragment	73.31 / 0.0	14 / 62
<i>R. mangle</i>	NODE_377_length_40 637_cov_4.285521	<i>Queuovirinae</i>	Medium-quality	Genome-fragment	61.3 / 0.0	17 / 66
<i>R. mangle</i>	NODE_378_length_40 626_cov_8.627936	Viruses	Medium-quality	Genome-fragment	63.32 / 0.0	7 / 61
<i>R. mangle</i>	NODE_440_length_38 268_cov_6.664983	<i>Tokiovirus A1</i>	Medium-quality	Genome-fragment	54.88 / 0.0	10 / 61
<i>R. mangle</i>	NODE_1281_length_2 4149_cov_5.527932	<i>Lactobacillus</i> phage Sha1	Low-quality	Genome-fragment	46.96 / 0.0	19 / 34
<i>R. mangle</i>	NODE_1454_length_2 2705_cov_4.604901	<i>Autographiviridae</i>	Low-quality	Genome-fragment	29.99 / 0.0	6 / 29
<i>R. mangle</i>	NODE_1497_length_2 2346_cov_4.789377	<i>Pseudomonas</i> phage Kopi	Low-quality	Genome-fragment	29.33 / 0.0	9 / 37
<i>R. mangle</i>	NODE_2675_length_1 6321_cov_13.360261	<i>Caudoviricetes</i>	Low-quality	Genome-fragment	24.45 / 0.0	9 / 19
<i>R. mangle</i>	NODE_13975_length_5934 5934_cov_5.845042	<i>Lactobacillus</i> phage phigle	Low-quality	Genome-fragment	11.12 / 0.0	3 / 7

*: Provirus. CC: Completación/Contaminación. GV/GT: Conteo de genes virales / Conteo de genes totales.

Anotación de genes auxiliares metabólicos en los vMAGs reconstruidos obtenidos por planta.

Después de identificar, anotar taxonómicamente y evaluar la calidad de los genomas virales reconstruidos de *A. germinans* y *R. mangle* (Tabla 1), se procede a identificar los genes auxiliares metabólicos (AMGs) presentes en ellos, para ello se emplearon dos herramientas bioinformáticas, PGPg-finder v1.1.0 y run_dbcan v.4.1.4.

- Anotación funcional con la base de datos de PLaBase de AMG de los genomas virales reconstruidos.

La anotación de AMGs que intervienen en la interacción planta-bacteria y la identificación de sus funciones putativas, se realizó con PGPg-finder v1.1.0 y la base de datos PLaBase. La realización de este procedimiento resultó en la identificación de AMGs relacionados con la interacción planta-bacteria

en 6 / 6 vMAGs parcialmente reconstruidos identificados en muestras de *A. germinans* (Figura 5A), y 12/ 14 vMAGs parcialmente reconstruidos identificados en muestras de *R. mangle* (Figura 5B).

Para la identificación de funciones putativas, se clasificaron los AMGs identificados, dentro las categorías funcionales de la base de datos PLaBase, estas categorías están organizadas dentro de 6 niveles de detalle, siendo el nivel 1 el más general (clasifica según el tipo de efecto: directo o indirecto) y el nivel 6 el más específico (clasifica según el gen) (Pellegrinetti *et al.*, 2024). Como resultado de la identificación de las funciones putativas de los AMGs identificados dentro de los vMAGs reconstruidos de muestras de ambas especies vegetales, se obtuvo la clasificación de los AMGs en 5 categorías de nivel 2 (Figura 5): *biofertilización*, *sistemas de colonización de plantas*, *exclusión competitiva*, *producción de señales/fitohormonas*, y *control de estrés/biocontrol*.

- Enriquecimiento funcional con PLaBase de los AMGs de los genomas virales reconstruidos de *A. germinans*.

Para los AMGs de los 6 vMAGs reconstruidos de *A. germinans* (Figura 5A), la categoría funcional de nivel 2 más enriquecida fue el *sistema de colonización de plantas*, seguido por la *exclusión competitiva* y el *control de estrés o biocontrol*.

Entre los 6 vMAGs parcialmente reconstruidos de *A. germinans* (Figura 5A), el vMAG del fago PM1 de *Bacillus* fue el que presentó un mayor número de AMGs asociados a las categorías funcionales de nivel 2, presentando cuatro categorías anotadas (*sistemas de colonización de plantas*, *exclusión competitiva*, *producción de señales/fitohormonas*, y *control de estrés/biocontrol*), de las cuales la categoría funcional *sistemas de colonización de plantas*, fue la categoría con mayores anotaciones de AMGs.

Por otro lado, el vMAG *Howevirus* (Figura 5A), se distingue por presentar AMGs anotados a las categorías funcionales de *biofertilización*, *sistemas de colonización de plantas*, *exclusión competitiva* y *control del estrés o biocontrol*, entre las cuales la categoría *sistemas de colonización de plantas* fue la que presentó una mayor proporción de AMGs anotados para este genoma viral (Figura 5A).

Además, los vMAGs parcialmente reconstruidos fago LASTA de *Klebsiella* y *Aresaunavirus* RSA1 (Figura 5A), junto al contig del fago PM1 de *Bacillus*, comprenden a los vMAGs parcialmente reconstruidos con mayores anotaciones de AMGs relacionados con la categoría de *exclusión competitiva*, la cual es la segunda categoría funcional de nivel 2 más enriquecida entre los genomas virales parcialmente reconstruidos de muestras de *A. germinans* (Figura 5A).

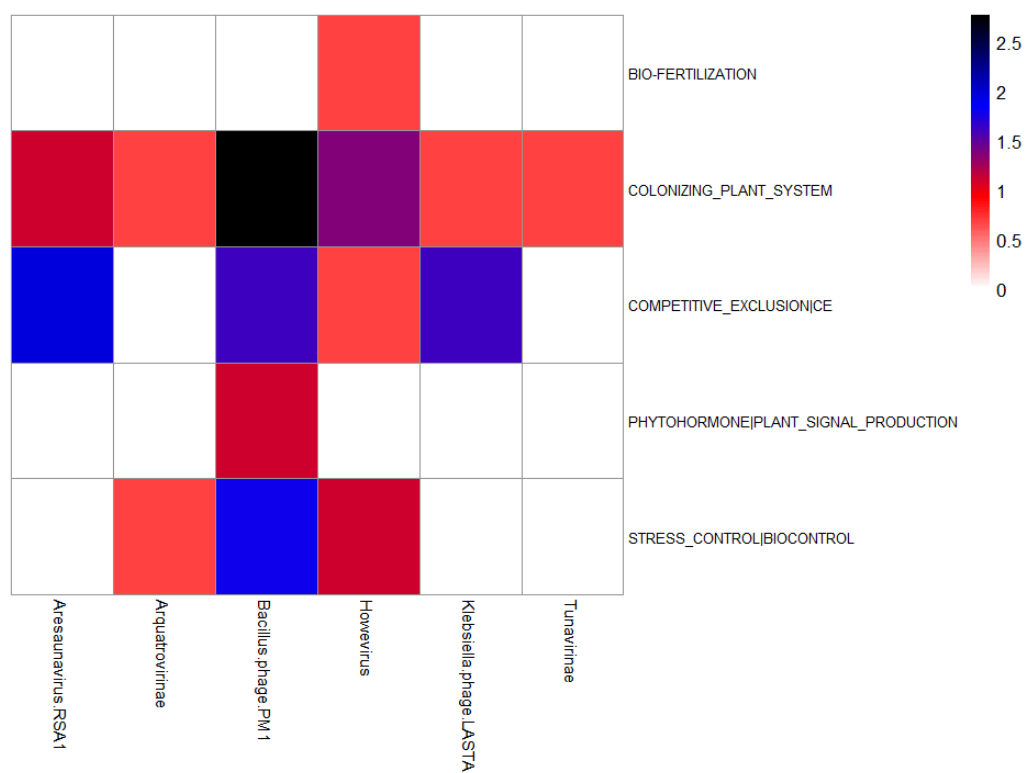
- ***Enriquecimiento funcional con PLaBase de los AMGs de los vMAGs reconstruidos de R. mangle.***

Para las muestras de *R. mangle*, se observó que los AMGs de los 12 vMAGs reconstruidos y enriquecidos con la base de datos PLaBase, en su mayoría se encuentran anotados en las categorías funcionales de nivel 2 de *control de estrés o biocontrol*, seguido por *sistemas de colonización de plantas y exclusión competitiva* (Figura 5B).

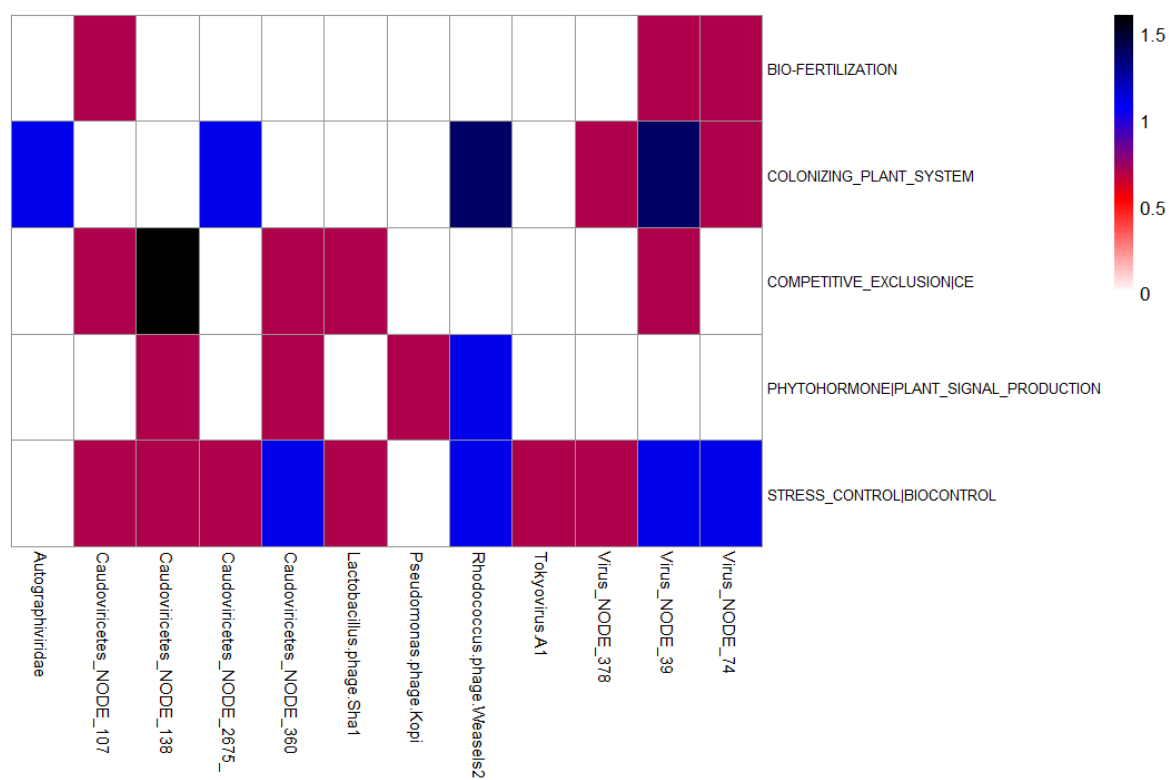
Entre los 12 vMAGs parcialmente reconstruidos anotados, el vMAG del fago Weasels2 de *Rhodococcus* fue el que presentó el mayor número de AMGs anotados (Figura 5B), los cuales se distribuyeron en las tres categorías funcionales de *sistemas de colonización de plantas, fitohormonas o producción de señales vegetales y control de estrés o biocontrol* (Figura 5B).

Por otro lado, la categoría funcional de nivel 2, *control de estrés o biocontrol* (Figura 5B), la cual fue la categoría funcional con mayor cantidad de AMG anotados, también fue la categoría con más vMAGs anotados, con 10 vMAGs reconstruidos, entre los que se encuentran: *Caudoviricetes_NODE_107*, *Caudoviricetes_NODE_138*, *Caudoviricetes_NODE_2675*, *Caudoviricetes_NODE_360*, el fago Sha1 de *Lactobacillus*, el fago Weasels2 de *Rhodococcus*, Tokyovirus A1, *Virus_NODE_378*, *Virus_NODE_39* y *Virus_NODE_74*.

Además, los vMAGs reconstruidos *Virus_NODE_39*, *Caudoviricetes_NODE_2675* y *Autographiviridae*, junto al vMAG del fago Weasels2 de *Rhodococcus*, fueron identificados como los vMAGs con mayor número de AMGs anotados para la categoría funcional de nivel 2 *sistemas de colonización de plantas*, la cual fue la segunda categoría funcional con mayor cantidad de AMG anotados en muestras de *R. mangle*.



A.



B.

Figura 5. Anotación de categorías génicas involucradas en la relación microbiota-planta, dentro de los vMAGs reconstruidos de tejidos de *A. germinans* (A.) y *R. mangle* (B.). En vMAGs reconstruidos, se evidencia la abundancia de CDS, clasificados en las categorías del nivel 2 (Lv2) de la base de datos de PLaBase, los cuales están relacionados con funciones metabólicas implicadas en la interacción planta-microorganismo. En la escala ubicada a la derecha de la figura, el color negro simboliza una alta abundancia de secuencias y el color blanco representa la ausencia de secuencias anotadas para esa función.

- ***Anotación funcional de CAZimas en los vMAGs reconstruidos de muestras de A. germinans y R. mangle.***

Igualmente, se realizó la anotación de AMGs asociados a enzimas activas en el metabolismo de carbohidratos (CAZimas) dentro de los vMAGs reconstruidos anotados por cada especie de mangle, con el software run_dbcan v.4.1.4, con la base de datos dbCAN3 (Figura 6). Tras la realización de este procedimiento, los AMGs anotados fueron categorizados en 12 familias de CAZimas, las cuales se relacionan con módulos de unión a carbohidratos (CBM, por sus siglas en inglés), glicosil hidrolasas (GH) y con glicosil transferasas (GT). En los vMAGs parcialmente reconstruidos de muestras endófitas de *A. germinans*, se identificaron un total de 15 AMGs asociados a CAZimas, dentro de los cuales, se encuentran 12 AMGs asociados a 4 familias de CAZimas GH (GH23, GH73, GH108 y GH133), 2 AMGs asociados a 2 familias de CAZimas CBM (CBM32 y CBM50) y un AMG asociado a la familia de CAZimas GT2. En el caso de las muestras de *R. mangle*, se identificaron 11 AMGs asociados a CAZimas, de los cuales 7 están asociados a 6 familias de CAZimas GH (GH13, GH19, GH23, GH24, GH25 y GH108), y 4 se asocian a 2 familias de CAZimas GT (GT2 y GT101).

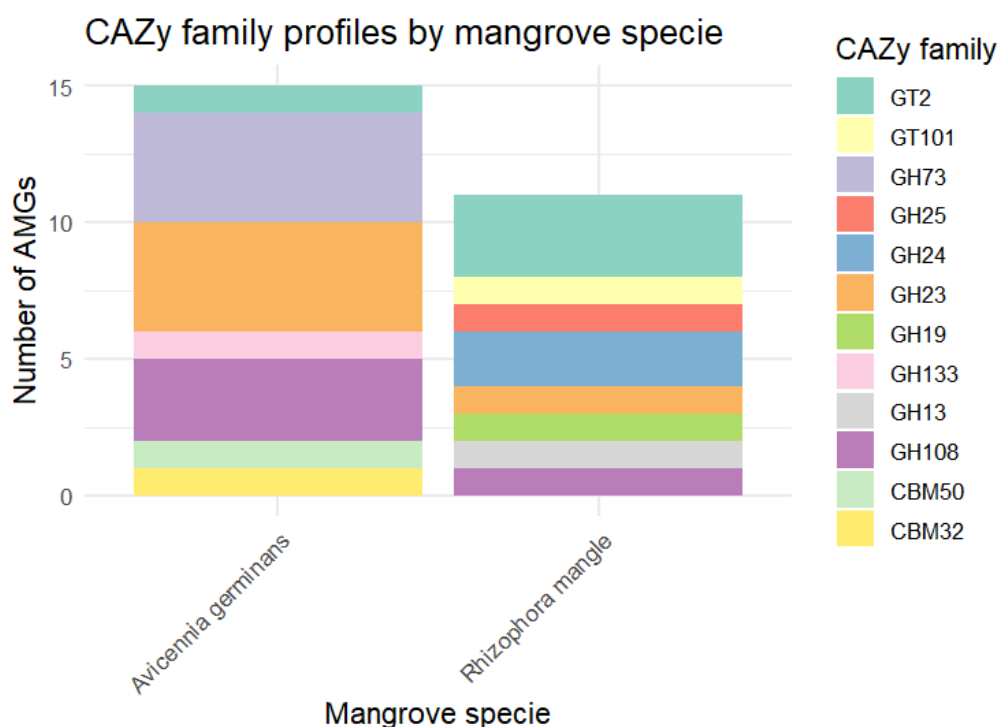


Figura 6. Perfil de familias de CAZimas identificadas en vMAGs reconstruidos de muestras de tejidos de *A. germinans* y *R. mangle*. Los AMGs que fueron anotados con la Carbohydrate-Active enZYmes Database (CAZY database) y clasificados en 12 familias de CAZimas (CBM32, CBM50, GH108, GH13, GH133, GH19, GH23, GH24, GH25, GH73, GT101, GT2). Cada barra representa la abundancia absoluta de los AMGs de los genomas virales parcialmente reconstruidos ensamblados, que fueron anotados como CAZimas, en muestras de tejidos de las dos especies de mangle, *A. germinans* y *R. mangle*.

Discusión.

En este capítulo se realizó el análisis taxonómico y funcional de las comunidades virales de librerías metagenómicas de muestras de hojas y raíces de *A. germinans* y *R. mangle*

Análisis taxonómico de comunidades virales asociadas a tejidos foliares y radiculares de A. germinans y R. mangle.

En los análisis taxonómicos, se identificó una alta abundancia de secuencias de ADN asociadas a los géneros virales *Andhravirus* y *Valbivirus*, tanto en *A. germinans* como en *R. mangle* (Figura 1). Los virus del género *Andhravirus*, representados por fagos de *Staphylococcus* (Cater *et al.*, 2017), generan un

control sobre la población de bacterias del género *Staphylococcus*. Este clado bacteriano se ha caracterizado por ser endófito y por mostrar un gran potencial como promotor del crecimiento vegetal en tallos (Olayemi & Odedara, 2017; Zhang, Y., *et al.*, 2023), hojas (Olayemi & Odedara, 2017; Zhang, Y., *et al.*, 2023), raíces (Olayemi & Odedara, 2017; Zhang, Y., *et al.*, 2023) y semillas (Alibrandi *et al.*, 2018; Zhang, Y., *et al.*, 2023) de especies vegetales como el arroz, la vid, el ajo y *Anadenanthera colubrina*. Por su parte, los *Valbivirus*, son virus con genoma de doble cadena de ADN (dsDNA), representados por el bacteriófago ValB1MD-2, presente en el género bacteriano *Vibrio*. Se ha encontrado que estos virus pueden controlar poblaciones de *Vibrio alginolyticus* (Mihara *et al.*, 2016), una especie bacteriana característica de muestras oceánicas y de ecosistemas costeros, como el manglar. Como característica interesante, a *Vibrio alginolyticus* se le asocia con un gran potencial biotecnológico ya que, posee un reservorio de genes de resistencia (Tan *et al.*, 2022), además, cuenta con potencial para la transferencia horizontal de genes (HGT, por sus siglas en inglés) mediada por MGEs (Chibani *et al.*, 2020), lo cual la identifica como una especie bacteriana con importancia ecológica y evolutiva en estudios microbianos de ambientes marinos (Luo *et al.*, 2012; Hernández-Robles *et al.*, 2016; Chibani *et al.*, 2020).

Por su parte, los *Mementomorivirus*, identificados únicamente en muestras foliares de *A. germinans* (Figura 1), es un clado representado por bacteriófagos que parasitan a *Microbacterium paraoxydans*, *Microbacterium oxydans* y *Microbacterium foliorum* (Mihara *et al.*, 2016). Estas tres bacterias, han demostrado potencial en la promoción del crecimiento vegetal (Alka *et al.*, 2021; Alka *et al.*, 2023), tolerancia a metales pesados (Kaushik *et al.*, 2012; Alka *et al.*, 2021; Alka *et al.*, 2023) y formación de sideróforos (Alka *et al.*, 2023), funciones ecológicas de gran impacto para las plantas en entornos rizosféricos. Sin embargo, estas bacterias, especialmente *Microbacterium foliorum*, han sido aisladas de muestras foliares (Behrendt, Ulrich & Schumann, 2001). Probablemente, los fagos del género *Mementomorivirus*, identificados en este estudio, se encuentren regulando el crecimiento de las *Microbacterium* en un entorno endófito en muestras foliares de *A. germinans*.

Con respecto a las raíces, se encontraron virus del género *Betabaculovirus* en abundancias superiores al 5% en muestras radiculares de *A. germinans* y *R. mangle* (Figura 1). Estos virus han sido encontrados, principalmente, en insectos lepidópteros (también conocidos como granulovirus infectantes de lepidópteros) (Reid *et al.*, 2023; Rohrmann, 2019). Sin embargo, Gómez-Bonilla y coautores (2012), reportan el aislamiento y la identificación de cepas virales provenientes del suelo, con un gran potencial para usarse en actividades de control biológico de insectos plaga. Se puede hipotetizar que la presencia de *Betabaculovirus*, como endófito de raíz de ambas especies de mangle, podría controlar y evitar efectos dañinos en las raíces causadas por larvas de lepidópteros, sin embargo, es evidente que se

requieren más estudios sobre este clado, donde se compruebe experimentalmente su papel ecológico como endófito en plantas de este bioma.

Otros clados virales, poco estudiados, con abundancias relativas significativas altas como *Oryzopoxvirus*, fueron detectados. Este género viral ha sido identificado inicialmente en tejidos de mamíferos (Mihara *et al.*, 2016; Ibrahim *et al.*, 2014), pero se cuenta con limitada información sobre su ecología.

Análisis de diversidad de comunidades virales asociadas a tejidos foliares y radicales de A. germinans y R. mangle.

Al analizar el viroma de ambas especies de mangle y sus tejidos se identificó una mayor diversidad en muestras radicales, contrastada con las muestras foliares de ambas especies (Figura 2), afirmación apoyada en el PCoA (Figura 4) y la PERMANOVA (Anexo 6), donde se observó una diferenciación significativa en la composición viral teniendo en cuenta el factor tejido. Se puede plantear que las diferencias entre las condiciones, a las cuales se exponen ambos tipos de tejidos, prevalecen sobre el factor especie, lo cual lleva a la modificación de la composición del viroma. Condiciones variables y contrastantes entre ambos tipos de tejidos, como la disponibilidad de agua y oxígeno, cantidad de materia orgánica, salinidad y nivel de radiación lumínica, se han identificado como factores ambientales que modulan la composición de especies bacterianas del microbioma (Williamson *et al.*, 2005; Fierer & Jackson, 2006; Orland *et al.*, 2019; Underwood *et al.*, 2022; Mészárosóvá *et al.*, 2024), de igual manera, las comunidades virales se ven afectadas por estas condiciones ambientales, dado que se ha identificado que los virus, en su mayoría, presentan una alta especificidad al hábitat (Paez-Espino *et al.*, 2016)

Por otra parte, al realizar el análisis de las curvas de rarefacción de cada muestra (Anexo 5), se observó que, en el caso de las secuencias virales, en ninguna se llegó a la meseta, lo cual indica que los esfuerzos de secuenciación, aunque muy productivos en términos de información, no fueron suficientes para tener una cobertura óptima de la comunidad viral endófito de tejidos foliares y radicales de *A. germinans* y *R. mangle* (Rodríguez-R & Konstantinidis, 2014). Basado en estos análisis de rarefacción, se recomienda para próximas investigaciones, con enfoques de caracterización del viroma endófito, establecer una profundidad de secuenciación del metagenoma por muestra mayor a la empleada en este estudio (>20 GB), o implementar metodologías de enriquecimiento viral antes de la extracción de ácidos nucleicos (Sbardellati & Vannette, 2024), para obtener una caracterización más amplia del viroma.

Análisis de calidad de los genomas virales reconstruidos y filtrados con VIRify.

El análisis de calidad de los genomas virales parcial y totalmente reconstruidos de las muestras de *A. germinans* y *R. mangle* (Tabla 1), mostró la presencia de 5 vMAGs de *A. germinans* y 9 vMAGs de *R. mangle*, clasificados como genomas virales parcialmente reconstruidos de calidad media, bajo umbrales de completación y contaminación, del >50% y <20%, respectivamente (Tabla 1).

Aunque el umbral de contaminación empleado en este estudio para la identificación de vMAGs de calidad media (<20%), difiere del planteado por Bowers y colaboradores (2017) (<10%), se nota que la mayoría de los vMAGs de calidad media, aquí analizados, evidenciaron niveles de completación, superiores al 50%, y de contaminación, inferiores al 10%. La única excepción fue el fago de *Bacillus* PM1 de *A. germinans*, que supera el umbral máximo de contaminación planteado por Bowers y colaboradores (2017) (Tabla 1).

En particular, el vMAG del fago de *Bacillus* PM1 de *A. germinans* (Tabla 1), fue el único genoma viral identificado como provirus por el pipeline VIRify, que también fue categorizado como vMAG de calidad media, a pesar de superar el umbral del 10% contaminación (19.07%) (Tabla 1). El genoma viral del fago de *Bacillus*, puede ser considerado vMAG (Meziti *et al.*, 2021), dada la alta completación de este genoma viral de calidad media, la presencia de genes estructurales para fagos (Anexo 11), y la clasificación inicial de este vMAG como contig viral de alta confianza por VirSorter, VirFinder y PPRMeta, a través del pipeline VIRify. Además, a través de los análisis realizados con VIRify, en este vMAG, se identificaron 6 genes provenientes de hospederos bacterianos, y 29 genes virales, la presencia de estos genes bacterianos es la principal razón del aumento de la contaminación y la clasificación como provirus que recibió este genoma viral.

Análisis estructurales y funcionales de los genomas virales reconstruidos y anotados por especie de mangle.

- Análisis estructural de genomas virales de A. germinans y R. mangle.

En el análisis estructural de los genomas virales reconstruidos a partir de muestras de *A. germinans* (Anexo 11), se identificaron los fagos *Aresaunavirus* RSA1, *Klebsiella* fago LASTA y *Bacillus* fago PM1. En estos se evidenció un repertorio genético que incluye múltiples genes estructurales (cabeza, cola, portal, plato, fibras de la cola), proteínas de recombinación e integración (terminasa, endonucleasa, integrasa), enzimas asociadas a la replicación (DUT, PolA, DNA helicase), así como holinas, proteínas

antirrepresoras y proteínas similares a fagos. La presencia de este conjunto de genes sugiere que estos vMAGs tienen la capacidad de generar proteínas estructurales e insertarse de forma autónoma en el genoma bacteriano (Umene & Shiraishi, 2013). Las proteínas de recombinación/integración, en particular las integrasas, se asocian con la transferencia horizontal de genes mediada por integrones (Qi et al., 2023). Estas características confieren a estos fagos el potencial para movilizar casetes génicos entre el cromosoma y los plásmidos de la bacteria huésped, así como para transferir genes de resistencia a antimicrobianos (AMR) y genes metabólicos auxiliares (AMGs).

Por otra parte, en el análisis de los vMAGs de *R. mangle* (Anexo 12), destacan los genomas de los virus *Caudoviricetes_NODE_107*, *Caudoviricetes_NODE_138*, *Caudoviricetes_NODE_377* y el fago Sha1 de *Lactobacillus*. Estos presentaron múltiples genes estructurales (cola, cabeza, portal, fibras de cola, plato), proteínas de recombinación/integración (terminasa, endonucleasa, integrasa), proteasas virales y genes de ARN de transferencia (tRNA). Respecto a los genes de tRNA, estos han sido reportados principalmente en virus de ADN con genomas extensos (Dreher, 2010), donde pueden complementar la capacidad de transcripción de la célula huésped para favorecer la replicación viral.

- Análisis funcional de AMGs de vMAGs de *A. germinans* y *R. mangle*.

En el análisis funcional con PLaBase de los AMGs de los vMAGs de *A. germinans* y *R. mangle*, se enriquecieron las categorías del *sistema de colonización de plantas, exclusión competitiva y control de estrés o biocontrol*, para las muestras de ambas especies (Figura 5).

En la categoría funcional de *control de estrés o biocontrol*, se identificaron principalmente genes relacionados con estrés térmico (Anexo 7), es decir, genes que codifican proteínas de choque térmico (dnaJ, groes, mopB, clpP) (Anexo 8) y proteínas de ensamblaje de lipopolisacáridos (lapA) (Anexo 8). Estas funciones putativas, se relacionan directamente a las condiciones de alta temperatura del manglar, que alcanzaban los 30.2°C, al momento de la toma de muestras (Anexo 2). Estas condiciones ambientales propias del bioma de manglar pueden explicar la presencia de AMGs que codifican a proteínas de choque térmico y de ensamblaje de lipopolisacáridos (Anexo 7) (Luo et al., 2022). Entre los genomas virales analizados en *R. mangle* (*Tokyovirus A1*, fago Sha1 de *Lactobacillus*, *Virus_NODE_378*, *Caudoviricetes_NODE_2675*, fago Weasels2 de *Rhodococcus*, *Virus_NODE_39*) y 2 genomas virales de *A. germinans* (fago de *Bacillus* PM1 y *Howevirus*), se encontraron AMGs que codifican proteínas de choque térmico, lo que puede indicar una diferenciación en las estrategias empleadas por la microbiota endófito de ambas especies para adquirir y transferir AMG asociados con funciones de estrés térmico.

Por otra parte, en la categoría funcional *exclusión competitiva*, se anotaron AMGs relacionados con la exclusión competitiva en bacterias, como los genes asociados a sistemas toxina-antitoxina (mqsR, YdaS, mqsA, higA, yobL y darT) (Anexo 8) y a sistemas de restricción- modificación (hsdM, dam, mcrA y mod) (Anexo 8). Los sistemas toxina-antitoxina en bacterias (TA, por sus siglas en inglés) están compuestos por una toxina estable y una antitoxina inestable, estos sistemas cumplen diversas funciones biológicas, como el mantenimiento de la estabilidad de MGEs, la virulencia y la resistencia al estrés (Qiu *et al.*, 2022; Guglielmini & Van Melder, 2011). La presencia de sistemas toxina-antitoxina dentro de MGEs, puede suponer el mantenimiento de estos MGEs en la célula bacteriana, aunque en este momento, hay pocos reportes de estos sistemas TA en profagos (Yao *et al.*, 2018; Peltier *et al.*, 2020). No obstante, la presencia de genes asociados a estos sistemas toxina-antitoxina, podrían suponer que la regulación de funciones biológicas que mejoran la aptitud bacteriana también ocurre mediada por fagos (Qui *et al.*, 2022). También la presencia de AMGs asociados a los sistemas toxina-antitoxina YodL/YobK, YdaS/YdaT, HigB/HigA y MqsR/MqsA, en los vMAGs del fago LASTA de *Klebsiella*, fago PM1 de *Bacillus* y *Aresaunavirus* RSA1, podrían suponer un evento de transferencia horizontal de genes entre el cromosoma bacteriano y el genoma viral (Sonika *et al.*, 2023), y se pueden relacionar con mejoras en la capacidad adaptativa bacteriana, al favorecer la colonización de raíces (Chagnot *et al.*, 2013; Liu *et al.*, 2024; Zboralski & Fillion, 2020). Sin embargo, se requieren más estudios para explicar cada mecanismo de transferencia y asegurar la función ecológica de estos genes en las dinámicas de la relación planta-microorganismo.

Para la categoría funcional *sistema de colonización de plantas*, se anotaron AMGs relacionados con el metabolismo de macromoléculas provenientes de plantas (y ácidos nucleicos principalmente) (Anexo 7), y asociados a mecanismos bacterianos de adhesión a superficies (Anexo 7). Entre los AMGs relacionados con el metabolismo de macromoléculas provenientes de plantas, se identificaron genes que codifican para proteínas glicohidrolasas (yubE, rrrD y acm) (Anexo 8), proteínas de degradación de celulosa y hemicelulosa (celJ, eglA y eglS) (Anexo 8), y proteínas asociadas al metabolismo de purinas y pirimidinas (nrdA, nrdE, tadA, nrdB, nrdF, dut y rdgB) (Anexo 8). Estas proteínas, asociadas con el metabolismo de carbohidratos y nucleósidos provenientes de plantas, cumplen funciones vitales en la capacidad adaptativa y establecimiento de las bacterias endófitas (Liu *et al.*, 2025). El metabolismo de nucleósidos (pirimidinas y purinas), es una de las vías metabólicas de mayor impacto en la relación planta-bacteria, pues se ha relacionado con la producción de compuestos nitrogenados que actúan como promotores del crecimiento vegetal (Zheng, Y., *et al.*, 2024; Zhang, M., *et al.*, 2023), y la presencia de AMGs asociados a vMAGs que codifican proteínas que intervienen en esta ruta metabólica, suponen una modulación del metabolismo de nucleósidos mediada por virus, especialmente por profagos.

- ***Anotación funcional de CAZimas en los genomas virales reconstruidos de muestras de A. germinans y R. mangle.***

La anotación de CAZimas en los genomas virales de muestras de ambas especies, reveló la presencia mayoritaria de proteínas con actividad hidrolasa, superando el 80% y el 60% de la cantidad de genes de CAZimas anotados en genomas virales parcialmente reconstruidos de muestras de virus de *A. germinans* y *R. mangle*, respectivamente (Figura 6). La presencia de hidrolasas en genomas virales, especialmente en bacteriófagos, resulta funcionalmente relevante, ya que muchas de estas enzimas actúan como despolimerasas de polisacáridos (Wang *et al.*, 2024). Esta actividad permite la degradación de estructuras bacterianas de protección, como la cápsula, matriz de polisacáridos extracelular, la membrana externa de lipopolisacáridos, y la pared externa de peptidoglicano, favoreciendo el contacto del fago con la membrana celular bacteriana e iniciando el proceso de infección viral (Latka *et al.*, 2017). Entre las glicosil hidrolasas (GH) identificadas, se resalta la actividad de la familia GH23, en la cual se clasifican transglicosilasas líticas, que degradan el peptidoglicano (Martinez-Bond *et al.*, 2022; Papayova *et al.*, 2025), ampliando de esta forma la capacidad infecciosa de estos virus a bacterias gram positivas.

Las enzimas virales activas sobre carbohidratos del tipo GH, pueden encontrarse asociadas a módulos de unión a carbohidratos (CBMs, por sus siglas en inglés) (Ficko-Blean *et al.*, 2009). Estos módulos son dominios proteicos no catalíticos que incrementan la especificidad y afinidad en el reconocimiento de sustratos de polisacáridos, facilitando así la acción enzimática sobre estructuras complejas (Squeglia *et al.*, 2020). En el caso de la familia CBM32, Ficko-Blean y colaboradores (2009) reportaron una notable diversidad de módulos con una amplia gama de especificidades de unión a carbohidratos, lo cual sugiere una alta versatilidad funcional en contextos ecológicos diversos, incluyendo unidades virales.

Capítulo 3: Plasmidoma endófito del tejido foliar y radicular de *R. mangle* y *A. germinans*

Introducción.

El plasmidoma se define como el conjunto de plásmidos presentes en una muestra, cuya caracterización se realiza principalmente mediante herramientas metagenómicas (Kav *et al.*, 2012; Walker, 2012).

El estudio de los plasmidomas es relativamente reciente, puesto que el primer estudio de este tipo se remonta a 2010, cuando Fondi *et al.* (2010) caracterizaron el pan-plasmidoma del género *Acinetobacter*. Mediante un análisis filogenético de distancias de secuencias plasmídicas y análisis funcionales. Fondi *et al.* (2010), reportaron genes plasmídicos que codifican transposasas, confieren resistencia al mercurio y desempeñan un papel fundamental en la supervivencia del género bacteriano *Acinetobacter*.

Sin embargo, la investigación de los plasmidomas con enfoque metagenómico comenzó con el estudio del rumen bovino realizado por Kav *et al.* (2012). En este trabajo, además de caracterizar el plasmidoma ruminal, se comparó su perfil funcional con el de muestras metagenómicas, identificando una importante contribución metabólica, genética y funcional proporcionada únicamente por el plasmidoma.

En dicha década continuaron las investigaciones sobre plasmidomas de bacterias de importancia médica y biotecnológica (Song *et al.*, 2013; Ainsworth *et al.*, 2014; Van Mastrigt, 2018), además de la aparición de estudios sobre el plasmidoma asociado a ambientes con condiciones extremas, como bajas temperaturas (Ciok *et al.*, 2018) y presencia de fármacos y metales pesados (Dziewit *et al.*, 2015; Vincent *et al.*, 2016).

No obstante, a pesar de los avances bioinformáticos logrados hasta la fecha, aún existen dificultades significativas para la caracterización de los plasmidomas metagenómicos. En el estudio de Hilpert y colaboradores (2021) se analizaron las limitaciones en la capacidad de análisis de secuencias de ADN ambiental para caracterizar el plasmidoma, siendo la dificultad principal, el poder diferenciar el ADN cromosómico del plasmídico en el ensamblaje de secuencias. Como solución, se planteó el uso de ensambladores de novo como cBar y Recycler, y de referencia, como metaplasmidSPADES, así como la utilización conjunta de técnicas de secuenciación de lecturas largas y cortas (Hilpert *et al.*, 2021).

Actualmente, el interés en el estudio del plasmidoma en medicina ha aumentado, especialmente en investigaciones con muestras intestinales y sanguíneas (He *et al.*, 2024; Lipworth *et al.*, 2024; Stockdale & Hill, 2023), en las que se destaca el papel de los plásmidos como repositorio de genes funcionales que mejoran la capacidad de supervivencia de sus hospedadores bacterianos (He *et al.*, 2024), y se ha

identificado que hay plásmidos con potencial para transportar genes de resistencia a antibióticos (ARGs, por sus siglas en inglés) dada su estructura similar a la de plásmidos que sí los tienen, lo que otorga una gran importancia a los estudios de plásmidos y elementos móviles que no presentan ARG, debido a su potencial como vectores de transferencia de información genética (Lipworth *et al.*, 2024).

Por otra parte, en la revisión de literatura realizada, no se encontraron estudios de plasmidomas asociados a tejidos vegetales aislados de metagenomas de ADN ambiental. Sin embargo, hacia la segunda mitad de la década de 2010, Lanciano *et al.* (2017) investigaron sobre elementos genéticos móviles en plantas e introdujeron la técnica mobilome-seq para caracterizar el moviloma activo en plantas mediante la secuenciación de ADN circular extracromosómico (eccDNA). Aplicando esta metodología, identificaron retrotransposones activos, como PopRice, que genera eccDNA durante el desarrollo del endospermo en *Arabidopsis thaliana* y *Oryza sativa*. Este trabajo demuestra que los elementos genéticos móviles, incluidos los retrotransposones, están activos en los tejidos vegetales y pueden influir en la plasticidad y adaptación de las plantas.

En el contexto de las interacciones planta-microbio, los plásmidos desempeñan un papel central como elementos genéticos móviles que permiten la rápida adaptación de las bacterias a las condiciones ambientales cambiantes (Schierstaedt *et al.*, 2019). Se ha evidenciado que estos elementos móviles encontrados en la rizosfera promueven la motilidad (Scharf *et al.*, 2016), la competencia (Molinatto *et al.*, 2016) y la proliferación y establecimiento bacteriano (Wang *et al.*, 2007). Además, los plásmidos presentes en bacterias endófitas pueden determinar su capacidad para persistir en los tejidos, interactuar con el huésped y responder a las presiones ambientales en cada órgano de la planta (Taghavi *et al.*, 2005; Wang *et al.*, 2007), razón por la cual es de gran importancia el estudio de estos elementos móviles para comprender a detalle las dinámicas presentes en la interacción planta-microorganismo.

En ambientes extremos y altamente variables, como los manglares, donde los microorganismos endófitos asociados a las hojas y raíces están expuestos a alta radiación, concentraciones de salinidad y óxido-reducción variables (Dargiri *et al.*, 2025; Nie *et al.*, 2024), es vital comprender cómo se estructura el plasmidoma endófito, junto con los roles funcionales que este puede cumplir en la interacción planta-microorganismo. Este capítulo se centra en dilucidar el plasmidoma endófito presente en tejido foliar y radicular de dos especies dominantes de mangle del Pacífico colombiano (*Rhizophora mangle* y *Avicennia germinans*) mediante un enfoque metagenómico. Basándose en análisis bioinformáticos, busca caracterizar la diversidad taxonómica y funcional del plasmidoma, y mostrar el impacto ecológico y evolutivo de estos elementos genéticos móviles (MGEs, por sus siglas en inglés) en bacterias endófitas de *R. mangle* y *A. germinans*.

Objetivo específico.

Dilucidar la composición, diversidad y funciones genéticas putativas del plasmidoma endófito presente en tejido foliar y radicular de *R. mangle* y *A. germinans*, mediante un enfoque metagenómico.

Métodos.

Teniendo en cuenta y dando continuación a los métodos generales expuestos, se procede con los métodos específicos para el análisis de plasmidomas.

Anotación taxonómica del plasmidoma asociado a cada tejido y a cada especie de mangle .

Para la anotación taxonómica del plasmidoma asociado a las muestras foliares y radicales de *A. germinans* y *R. mangle*, se realizó la identificación a partir de las lecturas limpias. Para este fin se empleó la suite de Kraken2 v.2.1.4 (Wood *et al.*, 2019), junto con la base de datos de plásmidos asociada a esta herramienta bioinformática. Una vez obtenida la anotación taxonómica, se procedió a construir un archivo BIOM con los reportes producidos, empleando Kraken-biom v.1.0.1 (Dabdoub, 2016). La visualización se hizo mediante los paquetes vegan v.2.7-0 (Oksanen *et al.*, 2025) y ggplot2 v. 3.5.1 (Wickham, 2016) del repositorio CRAN v.2.6-10 (Hornik, 2012) de R v.4.4.2 (R Core Team, 2025)

Análisis estadísticos y de diversidad alfa y beta.

Una vez obtenida la anotación taxonómica del plasmidoma, se importaron los archivos BIOM en el software R v4.4.2 (R Core Team, 2025), y con la ayuda del paquete phyloseq v3.20 (McMurdie y Holmes, 2013), se realizaron análisis comparativos de la diversidad alfa mediante los índices de Shannon-Wiener, Chao1, ACE y Simpson. A continuación, se evaluó la diversidad beta mediante análisis estadísticos como el Análisis de Coordenadas Principales (PCoA) y el análisis de varianza multivariado permutacional (PERMANOVA) del paquete vegan v2.7-0 (Oksanen *et al.*, 2025) (bajo el supuesto de homogeneidad de varianzas evaluado en el Anexo 19).

Ensamblaje de secuencias limpias.

El ensamblaje de contigs plasmídicos a partir de lecturas limpias de ADN se realizó mediante el pipeline SPAdes v.4.0.0, empleando el programa metaplasmidSPADES (Antipov *et al.*, 2019).

Filtrado de plásmidos, identificación de genes de movilidad e identificación de plásmido completo.

Para el filtrado e identificación de plásmidos a partir de los contigs plasmídicos ensamblados con metaplasmidSPADES, se empleó el programa MOB-typer, de la caja de herramientas MOB-suite v.3.1.9 (Robertson & Nash, 2018). Con el software MOB-typer, se realizó identificación y tipificación de los plásmidos parcial y totalmente reconstruidos, teniendo en cuenta un umbral máximo de distancia MASH de 0.30.

Como apoyo al filtrado de los plásmidos parcial y totalmente reconstruidos, realizado con MOB-typer, se realizó un filtrado con la base de datos de proteínas plasmídicas del RefSeq del NCBI, para este proceso se empleó el programa DIAMOND v2.1.11 (Buchfink *et al.*, 2015), y se confirmó la identidad de los plásmidos parcial y totalmente reconstruidos que alcanzaran un E-value inferior a $1e-20$, y un porcentaje de identidad superior o igual al 50%.

Por su parte, para el filtrado y anotación de plásmidos totalmente reconstruidos, se empleó la herramienta MOB-recon, de la caja de herramientas MOB-suite v.3.1.9 (Robertson & Nash, 2018). Con esta herramienta, a partir de los plásmidos filtrados con MOB-typer, se identificaron como plásmidos completos, a los plásmidos reconstruidos que incluyen en su estructura genes MPF (mating pair formation genes), CDS para replicasas, relaxasas y orígenes de transferencia (oriT), además, se seleccionaron los plásmidos que presentan una distancia MASH <0.1 .

Para confirmar la identidad de plásmido completo, se empleó el software BLASTn v2.16.0 (Camacho & Madden, 2013), con la base de datos del NCBI, y se tuvo en cuenta los alineamientos con E-value inferior al $1e-20$, identidad superior al 90% y cobertura superior al 80%.

Asignación taxonómica de los plásmidos ensamblados.

La anotación taxonómica de los plásmidos ensamblados (plásmidos parcial y completamente reconstruidos) se realizó con la herramienta MASH, incluida en los análisis realizados por MOB-typer

(Robertson & Nash, 2018), empleando la base de datos del NCBI. Para la asignación taxonómica se tuvo en cuenta un valor umbral de distancia al vecino más cercano menor a 0.30, y se asignó el clado estimado en la columna “mash_neighbor_identification” por la herramienta MASH.

Anotación funcional de plásmidos ensamblados.

Para la anotación de genes auxiliares metabólicos en los plásmidos ensamblados (plásmidos parcial y completamente reconstruidos) por planta, se emplearon dos herramientas bioinformáticas, el pipeline de anotación de Rasgos de Promoción del Crecimiento Vegetal (PGPT), PGPg finder v.1.1.0 (Pellegrinetti *et al.*, 2024) y el software run_dbcan v.4.1.4 (Zheng, J., *et al.*, 2023).

El pipeline PGPg finder v.1.1.0 (Pellegrinetti *et al.*, 2024), con la base de datos PLant-associated BActeria web resource (PLaBAse) v.1.02 (Patz *et al.*, 2021), se empleó para la identificación de genes auxiliares metabólicos asociados al metabolismo bacteriano, en especial al metabolismo de compuestos nitrogenados, carbohidratos, lípidos y metales pesados, también incluyendo genes involucrados en la respuesta a estrés biótico y abiótico.

Por otra parte, el software AMRFinderPlus v.4.0.23 (Feldgarden *et al.*, 2021), asociado a la base de datos del NCBI, fue implementado para la identificación de genes de resistencia antimicrobiana (AMR) en los plásmidos ensamblados por muestra vegetal.

Para la visualización de los resultados de anotación funcional de genes auxiliares metabólicos en los plásmidos ensamblados, se empleó software R v.4.4.2 (R Core Team, 2025), con los paquetes ggplot2 v.3.5.1 (Wickham, 2016) y pheatmap v.1.0.12 (Kolde, 2015), del repositorio CRAN v.2.6-10 (Hornik, 2012).

Anotación estructural de los plásmidos ensamblados.

Por otra parte, para realizar la anotación estructural de plásmidos parcial y completamente reconstruidos, se utilizó el portal Proksee v.1.0.1 (Grant *et al.*, 2023). Primero, se realizó la anotación de los CDS presentes en cada fragmento plasmídico con el software Prokka v.1.2.0 (Seemann, 2014). Posteriormente, se predijeron las regiones susceptibles de transferencia horizontal (HTG) con el software Alien Hunter v.1.3.0 (Vernikos & Parkhill, 2006) y se identificaron grupos de ortólogos móviles asociados a elementos genéticos móviles (MGEs) con mobileOG-db v.1.1.3 (Brown *et al.*,

2022), por último, se anotan y ubican los CDS asociados a plásmidos con la herramienta pLannotate v.1.2.0 (Wishart *et al.*, 2023).

Resultados.

Anotación taxonómica del plasmidoma de *A. germinans* y *R. mangle*.

La anotación taxonómica de las secuencias plasmídicas ensambladas y filtradas reveló una alta abundancia relativa (superior al 25%) de plásmidos asociados a los géneros bacterianos *Chlamydia*, *Bacillus*, *Lactiplantibacillus* y *Streptomyces* (Figura 7). En particular, los plásmidos de *Streptomyces* se encuentran ampliamente distribuidos en las muestras de tejido foliar de *A. germinans* y *R. mangle* (Figura 7), mientras que los de *Chlamydia* están presentes en 10 de las 12 muestras analizadas, independientemente del tipo de tejido o especie vegetal (Figura 7). Por otro lado, los plásmidos asociados a los géneros *Azospirillum* y *Rhizobium* se detectaron exclusivamente en muestras de endófitos radiculares de ambas especies (Figura 7), los plásmidos de *Pseudomonas* se limitaron a las raíces de *A. germinans* (Figura 7), y los de *Nonomuraea* por su parte se encontraron únicamente en raíces de *R. mangle* (Figura 7).

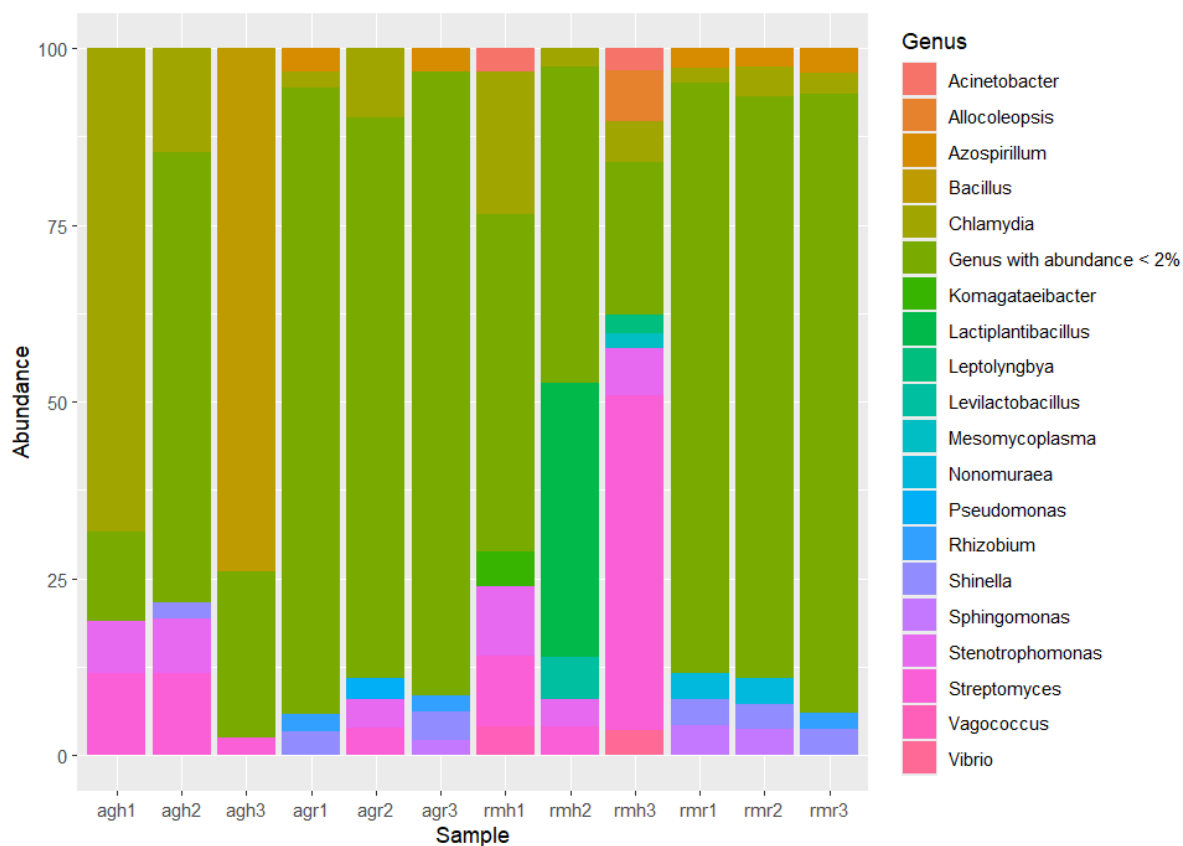


Figura 7. Abundancia relativa de plásmidos asociados a los géneros bacterianos más representativos de hojas y raíces de *A. germinans* y *R. mangle*. Cada barra representa la composición y la abundancia relativa de plásmidos asignados a diferentes géneros bacterianos en las muestras de hojas (h) y raíces (r) de *A. germinans* (ag) y *R. mangle* (rm). Los géneros con una abundancia menor al 2 % fueron agrupados en la categoría "Genus with abundance < 2 %"

Análisis de diversidad alfa del plasmidoma de A. germinans y R. mangle.

Se evaluó la diversidad alfa del plasmidoma de muestras de tejido foliar y radicular de *Avicennia germinans* y *Rhizophora mangle* mediante los índices Chao1, ACE, Shannon-Wiener y Simpson (Figura 8).

En el análisis del plasmidoma de librerías metagenómicas (Figura 8), se evidenció una marcada diferencia entre los tejidos. Las muestras de tejido radicular presentaron niveles de diversidad y riqueza mayores en muestras de *A. germinans* y *R. mangle* (Figura 8). En particular, los índices de Chao1 y ACE, que estiman la riqueza de especies, superaron las 2500 unidades operativas taxonómicas (OTUs,

por sus siglas en inglés), en la mayoría de las muestras del plasmidoma radicular de ambas especies, con la excepción de la réplica 2 de la muestra radicular de *A. germinans* (Figura 8). Asimismo, los valores de diversidad y equidad fueron más elevados en las muestras radiculares (Shannon-Wiener: 5.3–6.0; Simpson: 9.7–10.0) (Figura 8).

Por otra parte, las curvas de rarefacción de las muestras foliares no llegaron a la meseta (Anexo 9A), lo que muestra que es necesario ampliar la profundidad de secuenciación de este tipo de muestras para analizar la diversidad del plasmidoma en su totalidad.

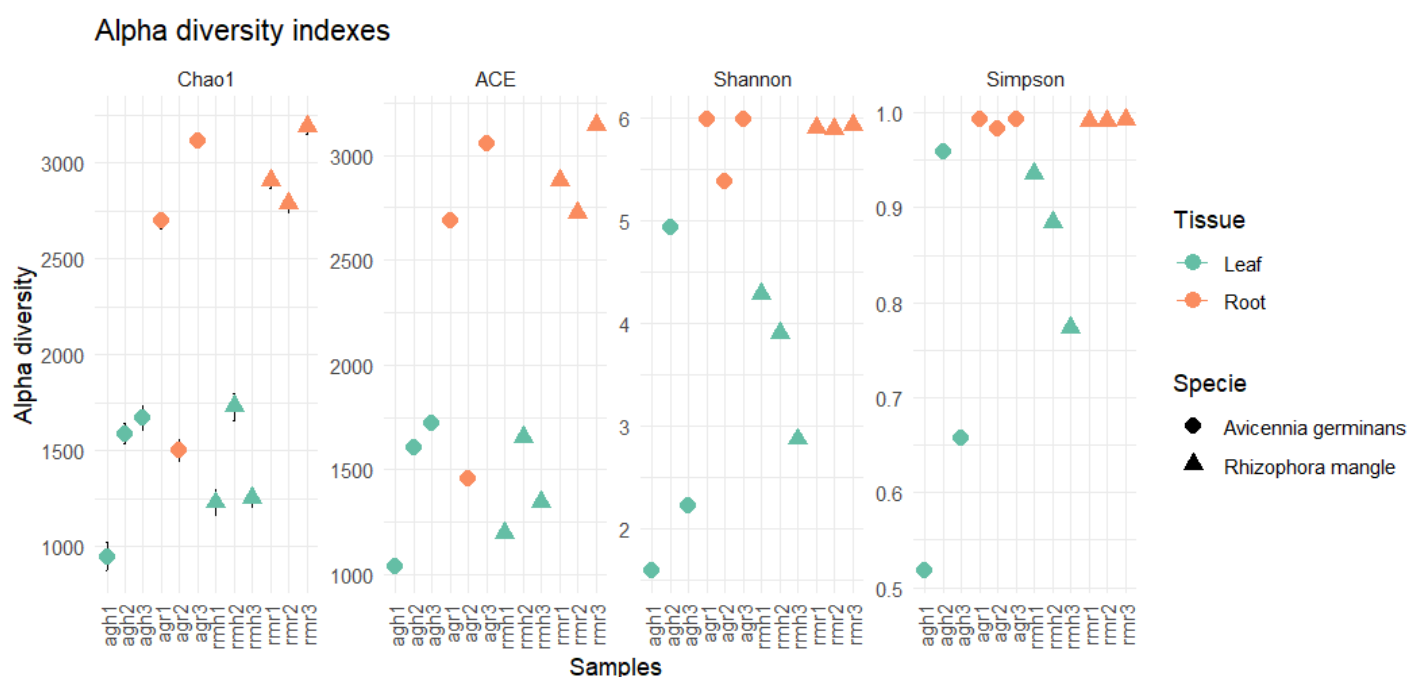


Figura 8. Diversidad alfa asociada al plasmidoma de tejido foliar y radicular de *A. germinans* y *R. mangle*. Los índices Chao1, ACE, Shannon y Simpson fueron calculados a partir de secuencias de plásmidos. Los colores indican el tipo de tejido (hoja o raíz) y las formas, la especie vegetal (*Avicennia germinans* o *Rhizophora mangle*).

Análisis estadísticos y de diversidad beta del plasmidoma de A. germinans y R. mangle.

Para evaluar la diversidad beta, se aplicaron un Análisis de Coordenadas Principales (PCoA) y un PERMANOVA a las muestras foliares y radiculares de *A. germinans* y *R. mangle*. En el PCoA (Figura 9), el eje horizontal explicó el 46,3 % de la variación entre muestras, y el eje vertical, el 14,4 %. Las

muestras se agruparon principalmente según el tipo de tejido (Figura 9), indicando una diferenciación en la composición del plasmidoma según el tejido de origen de la muestra.

Esta agrupación fue respaldada por el PERMANOVA (Anexo 10), que mostró un efecto significativo únicamente del factor “tejido” ($p = 0,004$), mientras que “especie” y la interacción “especie x tejido” no fueron significativos. Esto confirma que la variación del plasmidoma está influenciada principalmente por el tipo de tejido.

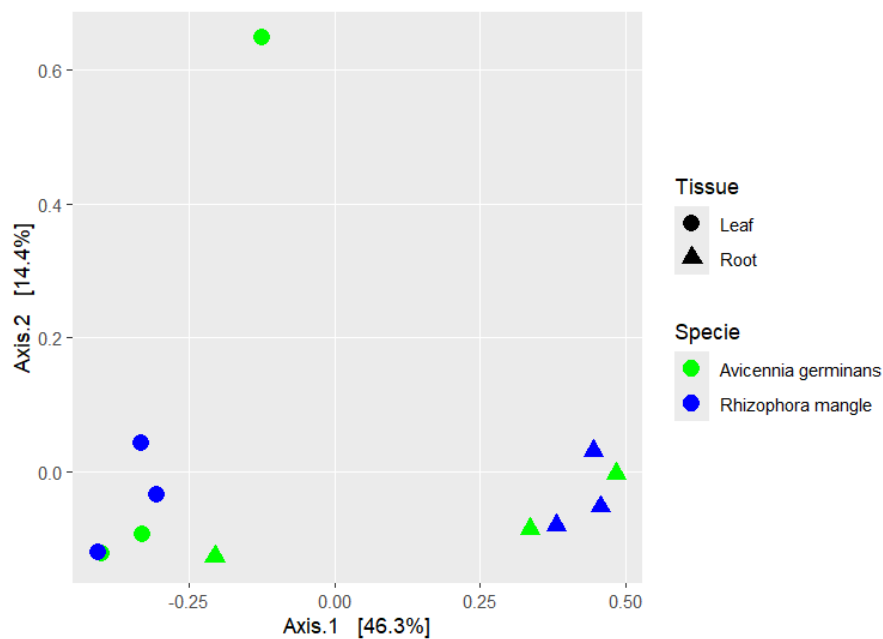


Figura 9. Análisis de Coordenadas Principales (PCoA) del conjunto de plásmidos asociados a tejido foliar y radicular de *A. germinans* y *R. mangle*. Se evidencia el efecto del tipo de tejido (foliar y radicular), y de la especie de planta (*A. germinans* y *R. mangle*) en la composición del plasmidoma identificado en cada muestra.

Anotación taxonómica de los plásmidos en muestras de A. germinans y R. mangle.

La identificación taxonómica de plásmidos parcialmente reconstruidos y plásmidos completos por planta se realizó luego del filtrado con MOB-typer y DIAMOND, con la base de datos del RefSeq del NCBI. Para la anotación taxonómica se empleó la herramienta MASH de MOB-typer (Robertson & Nash, 2018), junto a la base de datos del NCBI. Los resultados obtenidos se presentan en el Tabla 2, donde se identificó un plásmido completo, proveniente de muestras foliares de *R. mangle* (plásmido de

Lactiplantibacillus plantarum), y un total de 5 plásmidos parcialmente reconstruidos (Tabla 2), de los cuales 2 provienen de muestras de *A. germinans* (plásmidos de *Lactobacillus* sp. y *Paeniclostridium sordellii*), y 3 de las muestras de *R. mangle* (plásmidos de *Clostridium botulinum*, *Streptomyces albidoflavus* y *Paeniclostridium sordellii*).

Tabla 2. Anotación taxonómica de los plásmidos parcial y totalmente reconstruidos, que fueron anotados taxonómicamente con la herramienta MASH de MOB-typer. Para su anotación se empleó un umbral de valor MASH de 0,3.

Especie	Tejido	Contig ID	Categoría taxonómica asignada	Valor MASH
<i>A. germinans</i>	Raíz	NODE_1_length_134371_cov_13.866080	<i>Lactobacillus</i> sp.	0,296
<i>A. germinans</i>	Raíz	NODE_2_length_24580_cov_16.684333	<i>Paeniclostridium sordellii</i>	0,296
<i>R. mangle</i>	Raíz	NODE_1_length_105738_cov_15.535721	<i>Clostridium botulinum</i>	0,296
<i>R. mangle</i>	Raíz	NODE_3_length_12325_cov_1.434252	<i>Streptomyces albidoflavus</i>	0,296
<i>R. mangle</i>	Raíz	NODE_1_length_24581_cov_28.658870	<i>Paeniclostridium sordellii</i>	0,296
<i>R. mangle</i>	Hoja	NODE_1_length_13664_cov_8.493093	<i>Lactiplantibacillus plantarum</i>	0,056

Anotación funcional de plásmidos en muestras de A. germinans y R. mangle.

La anotación funcional de los plásmidos parcialmente reconstruidos ensamblados y filtrados por especie y tejido, se realizó mediante el software PGPg finder v1.1.0 (Pellegrinetti *et al.*, 2024), utilizando la base de datos PLaBAs. Como resultado, se identificaron siete plásmidos parcialmente reconstruidos con genes asociados a funciones bacterianas relacionadas con plantas: tres provenientes de raíces de *A. germinans* y cuatro de *R. mangle* (tres radicales y uno foliar) (Figura 10).

En total, se anotaron doce categorías funcionales de nivel 5 (Lv5) de la base de datos PLaBAs, distribuidas en seis de los siete fragmentos identificados (Figura 10). Cinco categorías se encontraron en *A. germinans* y once en *R. mangle*, lo que indica una mayor diversidad funcional en esta última especie (Figura 10).

Las categorías más representativas fueron *respiración aeróbica/fosforilación oxidativa*, *homeostasis del hierro*, y *halotolerancia mediada por enzimas* (Figura 10). Estas categorías estuvieron presentes en cuatro plásmidos parcialmente reconstruidos (de *Lactobacillus*-AgR, *Paeniclostridium*-AgR,

Paeniclostridium-RmR y *Clostridium*-RmR) y presentaron una alta abundancia de AMGs asociados (Figura 10).

También se identificaron otras funciones relevantes, como el metabolismo de fungicidas (compuesto antifúngico Soraphen) en un plásmido de *Streptomyces* de raíces de *R. mangle*, y la biosíntesis de biotina (vitamina B7) en un plásmido de *Lactobacillus* de raíces de *A. germinans*.

Por último, se evaluó la presencia de genes de resistencia antimicrobiana mediante AMRFinderPlus v4.0.23 (Feldgarden *et al.*, 2021) y la base de datos del NCBI. Sin embargo, no se detectaron genes de resistencia en ninguno de los plásmidos parcialmente reconstruidos analizados en este estudio.

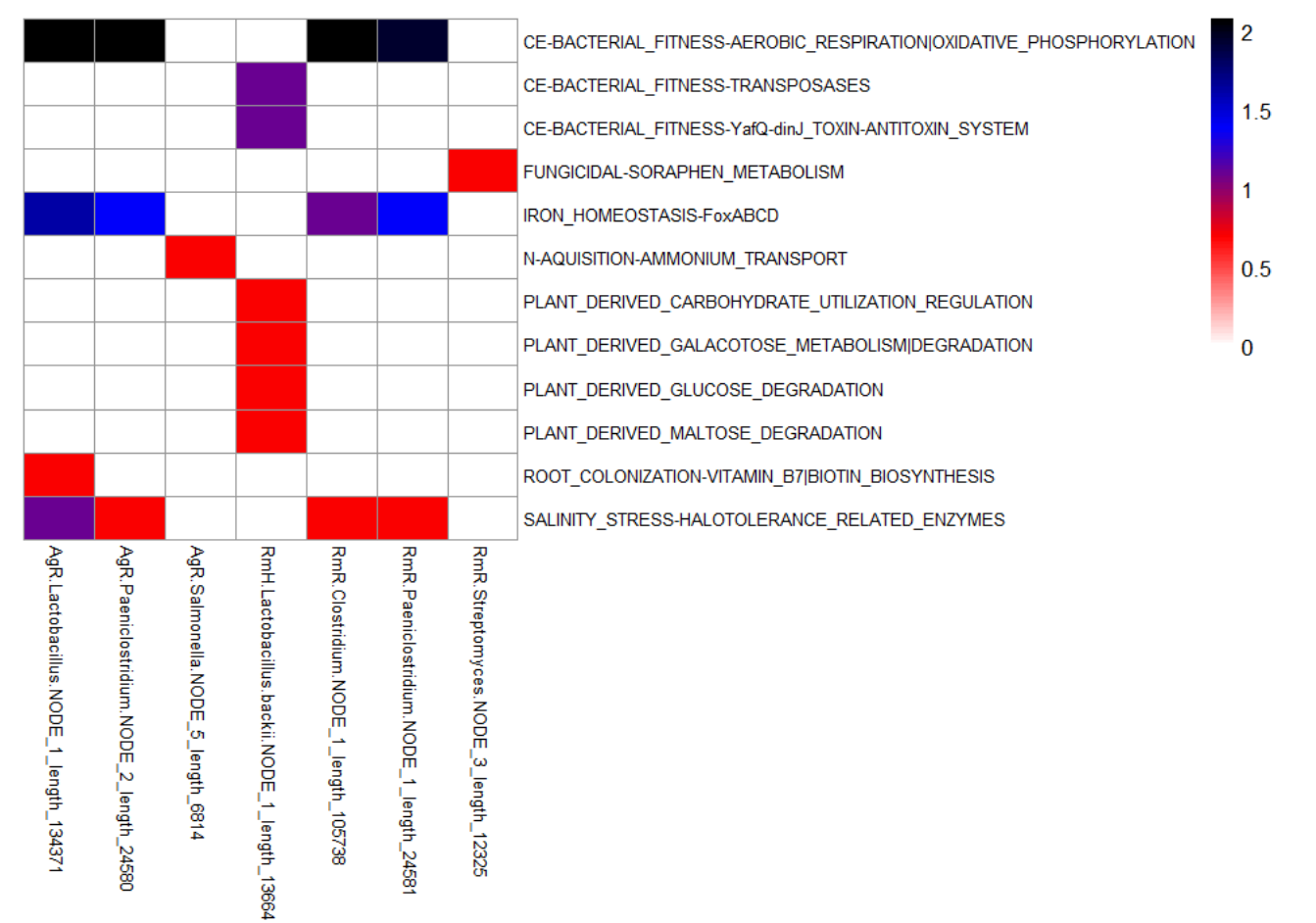


Figura 10. Análisis funcional de los plásmidos reconstruidos que presentan genes metabólicamente relevantes para la comunidad microbiana de tejidos foliares (H) y radiculares (R) de *A. germinans* (Ag) y *R. mangle* (Rm). Se evidencia la abundancia de los CDS relacionados con funciones metabólicas implicadas en la interacción planta-microorganismo. En la escala, ubicada a la derecha de la figura, el color negro simboliza una alta abundancia de las secuencias codificantes asociadas a cada categoría de

la base de datos PLaBase, y por su parte, el color blanco representa la ausencia de CDS dentro de las categorías.

Anotación estructural y funcional del plásmido pZW5A de *Lactiplantibacillus plantarum* de muestras foliares de *R. mangle*.

Entre los plásmidos reconstruidos, resaltó el plásmido de *Lactiplantibacillus* de muestras foliares de *R. mangle*, este plásmido fue el único categorizado por MOB-recon como un plásmido completo, al presentar un gen plasmídico de replicasas reportado en el plásmido pCIS4. Además, este plásmido totalmente reconstruido, se identificó como plásmido pZW5A, de *Lactiplantibacillus plantarum* cepa ZW5, al alinearse completamente con las secuencias reportadas en la base de datos del NCBI para el plásmido pZW5A, obteniendo un E-value de 0.0, un valor de cobertura de 100% y un porcentaje de identidad del 99.99%.

En términos de la anotación funcional (Figura 10), el plásmido reconstruido pZW5A, se anotó con funciones involucradas en el fitness bacteriano, como genes para transposasas y el sistema toxina-antitoxina YafQ-dinJ, como también se anotaron CDS asociados con el metabolismo de carbohidratos de origen vegetal, como galactosa, glucosa y maltosa. Por su parte, el análisis estructural del plásmido reconstruido pZW5A revela la presencia de genes relacionados con funciones de integración/excisión (tISRso8a, IS30, int1 y ISL3), proteínas de movilización de plásmidos, proteínas asociadas al sistema toxina-antitoxina tipo II (toxina de la familia YafQ) y replicasas (repA1 y repX) (Anexo 13).

Discusión.

En este capítulo, se filtraron, anotaron e identificaron a partir de secuencias metagenómicas de muestras foliares y radicales de *A. germinans* y *R. mangle*, los plásmidos y plásmidos parcialmente reconstruidos presentes.

Análisis taxonómico y de diversidad del plasmidoma de *A. germinans* y *R. mangle*.

La asignación taxonómica de las secuencias plasmídicas (Figura 7) reveló una amplia distribución de plásmidos asignados al género *Chlamydia*, detectados en 10 de las 12 muestras analizadas,

independientemente de la especie o tejido (Figura 7). *Chlamydia* es un clado tradicionalmente caracterizado por su papel como patógeno en humanos y otros animales (Geisler, 2010; Taylor-Brown & Polkinghorne, 2017), sin embargo, se han identificado también en entornos naturales infectando algunos protistas presentes en suelos, aguas dulces y saladas (Collingro *et al.*, 2020, Taylor-Brown *et al.*, 2015). En este caso, la presencia de plásmidos relacionados con esta bacteria en la endosfera de raíces y hojas de *A. germinans* y *R. mangle*, puede explicarse por dos hipótesis: A) La presencia de protistas al interior de estos tejidos los cuales albergan bacterias del género *Chlamydia*, que contienen estos plásmidos (Collingro *et al.*, 2020). B) Otra posible razón, es la presencia de eventos de transferencia horizontal mediada por plásmidos, entre bacterias del género *Chlamydia* a bacterias de otros géneros de la familia *Chlamydiaceae* (Marti *et al.*, 2022).

El análisis de diversidad alfa (Figura 8) mostró una mayor riqueza y diversidad en las muestras radiculares de ambas especies, mientras que el análisis de diversidad beta (Figura 9) reveló una diferenciación significativa guiada por el tipo de tejido, lo cual fue confirmado mediante la PERMANOVA (Anexo 10). Esta diferenciación puede deberse a las diferencias ambientales entre tejidos, que afectan la persistencia de plásmidos y sus bacterias portadoras (Trivedi *et al.*, 2022).

Entre los datos obtenidos (Figura 7), destaca la presencia en abundancias relativas altas (mayores al 5%) de plásmidos de bacterias de los géneros *Bacillus*, *Lactiplantibacillus* y *Streptomyces*, principalmente en muestras foliares (Figura 7). Las bacterias del género *Bacillus*, han sido caracterizadas ampliamente como simbiote vegetal, con capacidades como promotoras del crecimiento vegetal y biocontroladores de insectos e infecciones virales (Balderas-Ruiz *et al.*, 2020; Sorokan *et al.*, 2020; Wang *et al.* 2014), con representantes como *Bacillus thuringiensis*, *Bacillus velezensis* y *Bacillus subtilis*, que han sido aislados y caracterizados como parte de la filosfera y endosfera. Las bacterias de este género cuentan plásmidos, como el plásmido pBV71 de *Bacillus velezensis*, que ha demostrado tener un papel primordial en la formación de biopelículas, la represión de catabolitos y la biosíntesis de antibióticos, funciones que incrementan el potencial de promotoras del crecimiento vegetal en esta bacteria (Choi *et al.*, 2021). Por su parte, las bacterias del género *Lactiplantibacillus*, categorizadas como bacterias ácido-lácticas, con hábitos epífitos (como *Lactiplantibacillus pentosus*) y endófitos (como *Lactiplantibacillus plantarum*), que han albergado plásmidos conjugativos de gran importancia funcional, como pWCFS103, con genes que le confieren resistencia arsenito y el arseniato (Van Kranenburg *et al.*, 2005), y el plásmido pLU4, que puede proveer funciones asociadas a la adhesión y colonización (Kim *et al.*, 2017). Por otro lado, las bacterias del género *Streptomyces*, se caracterizan por su potencial como protectores y biocontroladores de enfermedades vegetales (Vergnes *et al.*, 2020), con capacidad para albergar plásmidos con genes para síntesis de antibióticos (como lankacidina C,

lankamicina y eritromicina) y regiones del cluster CRISPR, como el plásmido pSLA2-L (Mochizuki *et al.*, 2003). En resumen, los géneros *Bacillus*, *Lactiplantibacillus* y *Streptomyces* han sido caracterizados como simbioses vegetales con funciones promotoras del crecimiento vegetal, biocontrol de patógenos, y producción de antibacterianos y antifúngicos.

Los plásmidos asociados a los géneros *Azospirillum* y *Rhizobium*, se detectaron exclusivamente en muestras de endófitos radiculares de ambas especies vegetales, *A. germinans* y *R. mangle*, concuerda con estudios previos que reportan la asociación de estos géneros al tejido radicular, a la rizósfera y al suelo (Bashan & de-Bashan, 2010; Sessitsch *et al.*, 2002). En estos entornos las bacterias de los géneros *Azospirillum* y *Rhizobium*, actúan como promotoras del crecimiento vegetal, al mejorar la fijación de nitrógeno, actuar en la producción de fitohormonas, la reducción del estrés oxidativo y la defensa contra patógenos (Fukami *et al.*, 2018; Zahran, 2017), estas funciones metabólicas se ven apoyadas por plásmidos, que codifican la producción de lipopolisacáridos, exopolisacáridos, polisacáridos capsulares y metabolitos secundarios que favorecen la interacción con la planta (Katsy, 2011; DeJong *et al.*, 1981; Ding & Hynes, 2009; Hooykaas *et al.*, 1981).

Análisis funcional de los plásmidos parcialmente reconstruidos en muestras de A. germinans y R. mangle.

Entre los plásmidos parcialmente reconstruidos en muestras radiculares de *A. germinans* y *R. mangle*, resalta la anotación de cuatro funciones relacionadas a bacterias asociadas a plantas (Figura 10). Entre estas funciones, se encuentra la ruta de mantenimiento de la homeostasis del hierro, a través de los genes *foxA* y *foxB*, en los cuales se ha documentado un papel activo en la captación del hierro (Ó Cuív *et al.*, 2007). El gen *foxA*, se ha visto implicado en la formación de transportadores dependientes de TonB (TBDT), los cuales reconocen complejos ferrisideróforos, como la nocardamina, la bisucaberina y la ferrioxamina B, a las que le permiten el ingreso a la célula a través de la membrana celular empleando transporte activo (Will *et al.*, 2024; Josts *et al.*, 2021). Al ingresar estos compuestos a la célula, son transformados por una reductasa codificada por el gen *foxB*, lo que favorece la captación del hierro al estar reducido como Fe^{2+} (Josts *et al.*, 2021). La presencia de CDS de plásmidos bacterianos, asociados a esta ruta en endófitos de raíz de plantas de *A. germinans* y *R. mangle*, pueden incrementar la biodisponibilidad de este mineral en su estado reducido, lo que mejora significativamente el crecimiento de plantas de los géneros *Avicennia* y *Rhizophora* en ambientes estuarinos (Alongi, 2010).

La producción de enzimas relacionadas con la halotolerancia, es otra función anotada entre 4 de los 6 plásmidos parcialmente reconstruidos (Figura 10), la cual se relacionó con la identificación de CDS asociados a los genes *atpE* y *atpB* en estos contigs, que codifican a subunidades de la ATP sintasa. En el estudio de Soontharapirakkul y coautores (2011), se identificó la presencia de ATP sintasa dependiente de Na⁺ en cianobacterias halotolerantes. A pesar de que, los genes *atpE* y *atpB* generalmente se encuentran en el cromosoma bacteriano, tienen la capacidad de transferirse a plásmidos como parte del operón *atp* (Das & Ljungdahl, 2003). Teniendo en cuenta, los niveles de salinidad cambiantes y potencialmente elevados en las aguas y suelos del manglar de la zona muestreada (Anexo 3), es consistente que, entre las bacterias asociadas a suelos y raíces, existan estrategias de adaptación que faciliten la producción de ATP en entornos salinos.

Entre las categorías anotadas en plásmidos parcialmente reconstruidos, la única relacionada de manera directa con el control biológico fue la identificada en el fragmento plásmídico de *Streptomyces* de *R. mangle*, relacionada con la biosíntesis de sorafeno, asociada al gen *sorB* (Figura 10). El hallazgo de este gen en el fragmento plásmídico de *Streptomyces*, puede conferirles ventajas adaptativas a las bacterias pertenecientes a este género, además de proveer protección frente a infecciones fúngicas, a las plantas de *R. mangle* (Naini *et al.*, 2019)

Análisis funcional y estructural del plásmido reconstruido pZW5A de Lactiplantibacillus.

El plásmido pZW5A de *Lactiplantibacillus* ensamblado de muestras foliares de *R. mangle* (Anexo 13), presenta en su estructura genes relacionados con funciones de integración/excisión (*tISRso8a*, *IS30*, *int1* e *ISL3*), proteínas de movilización de plásmidos, proteínas asociadas al sistema toxina-antitoxina tipo II (toxina de la familia *YafQ*), replicasas (*repA1* y *repX*) y CDS asociados con el metabolismo de carbohidratos de origen vegetal, como galactosa (*ganR*), glucosa (*yvdM*) y maltosa (*malP*, *malR*, maltosa epimerasa) (Anexo 13).

Los genes anotados en el plásmido pZW5A de *Lactiplantibacillus*, con funciones de integración/excisión (*tISRso8a*, *IS30*, *int1* e *ISL3*), codifican para enzimas transposasas e integrasas. Entre estos genes, los genes *IS30* y *ISL3*, codifican para transposasas relacionadas con familias de secuencias de inserción (IS, por sus siglas en inglés) (Holmes *et al.*, 2001), las cuales cuentan con diversos mecanismos de transposición. En el caso de la familia de secuencias de inserción *IS30*, esta cuenta con un mecanismo de transposición descrito como de copiar-pegar (Olasz *et al.*, 2022), y la familia de secuencias de inserción *ISL3*, no cuenta con un mecanismo de transposición claramente

descrito (Siguier *et al.*, 2014). Por su parte, el gen *int1* codifica una proteína integrasa tirosina-dependiente, que hace parte de integrones con capacidad de capturar y transferir genes exógenos dentro de casetes génicos (Gillings, 2014). La presencia de genes vinculados a funciones de integración/excisión, como *tISRso8a*, *IS30*, *int1* e *ISL3* (Anexo 13), en el plásmido pZW5A de *Lactiplantibacillus*, demuestra el potencial de transferencia de este plásmido, como también posibilita la ocurrencia de eventos de transferencia horizontal en bacterias del género *Lactiplantibacillus*, presentes en las muestras foliares de *R. mangle*.

Por su parte, el sistema toxina-antitoxina tipo II de la familia YafQ/dinJ, identificado en el plásmido pZW5A de *Lactiplantibacillus* ensamblado en este estudio, es un mecanismo de regulación de la expresión génica (Prysak *et al.*, 2009). La toxina YafQ es una endorribonucleasa dependiente del ribosoma, que degrada ARNm en ausencia de la antitoxina dinJ (Prysak *et al.*, 2009; Singh *et al.*, 2021). Este sistema puede conferir a las bacterias que cargan este plásmido, la capacidad de modular la expresión génica en condiciones de estrés, con el fin de entrar en estado de latencia, lo cual puede alargar la supervivencia de estas poblaciones bacterianas (Hu *et al.*, 2015; Singh *et al.*, 2021).

La anotación de genes que codifican a replicasas (*repA1* y *repX*), en el plásmido pZW5A ensamblado, aseguran que es una unidad autoreplicante. Puesto que el gen *repA1*, traduce a la proteína de replicación A1, la cual tiene la función de marcar el sitio de inicio de la replicación e iniciar este proceso en el plásmido (Erjavec & Žgur-Bertok, 2006), y el gen *repX*, por su lado controla la replicación de los plásmidos (Markowska-Barkic *et al.*, 2022).

Además, se anotaron CDS relacionados con el metabolismo de carbohidratos de origen vegetal, como galactosa (*ganR*), glucosa (*yvdM*) y maltosa (*malP*, *malR*, maltosa epimerasa), en este plásmido. Estos genes se diferencian de los demás anotados en este plásmido por generar efectos directos sobre la planta, por ejemplo, la activación del gen *ganR*, inhibe la transcripción de los genes pertenecientes al operon *gan*, caracterizados por ejecutar el proceso de degradación del galactano, un polisacárido de origen vegetal (Habib *et al.*, 2017). También, la presencia del gen *yvdM*, que codifica a la betafosfoglucomutasa, una enzima con la capacidad de transformar glucosa-1-fosfato (G1P) en glucosa-6-fosfato (G6P) (Yamamoto *et al.*, 2001), y con ello activar rutas del metabolismo de carbohidratos, como la glucólisis. Finalmente, la anotación de genes asociados al metabolismo de la maltosa, entre ellos *malP*, *malR* y maltosa epimerasa, apoya la importancia funcional de este plásmido. El gen *malR*, cumple una función regulatoria sobre el operón de maltosa, a nivel transcripcional, al afectar el transporte de maltosa (Andersson & Rådström, 2002). El gen *malP*, codifica a la proteína maltodextrina fosforilasa, la cual interviene en la degradación de las maltodextrinas y la formación de la

G1P (Geremia *et al.*, 2002; Palm *et al.*, 1987), este G1P que podría ser transformado en G6P por la acción de la betafosfoglucomutasa del gen yvdM, anotado en este plásmido.

En resumen, el plásmido pZW5A de *Lactiplantibacillus*, ensamblado a partir de muestras de tejido foliar de *R. mangle*, corresponde a un plásmido funcional, con potencial metabólico y de transferencia, que tiene la capacidad de proveer adaptaciones relacionadas con el metabolismo de la maltosa a bacterias presentes en este tejido.

Capítulo 4: Discusión y conclusiones generales.

Discusión general.

La caracterización del viroma de ADN y del plasmidoma asociado a tejidos foliares y radicales de *A. germinans* y *R. mangle* en la bahía de Buenaventura revela que el tipo de tejido constituye un factor modulador clave en la composición de ambos conjuntos de elementos genéticos móviles endófitos, por encima de la especie vegetal (Anexo 6 y 8). Aunque estas dos especies pertenecen a familias distintas (*Acanthaceae* y *Rhizophoraceae*, respectivamente) y presentan adaptaciones ecológicas diferenciales que afectan la estructura y fisiología de sus órganos vegetativos (como los neumatóforos en raíces de *A. germinans* y las raíces en forma de zancos en *R. mangle*) (Srikanth *et al.*, 2016), se esperaba que la divergencia de condiciones dado el factor especie, fuera determinante en la composición de virus y plásmidos. Sin embargo, independientemente de la especie de planta, cada órgano vegetal se encuentra bajo presiones selectivas distintas, asociadas a la función, componentes anatómicos y el ambiente circundante que lo afecta (Dastogeer *et al.*, 2020; Turner *et al.*, 2013), lo cual podría explicar la convergencia en el patrón de diferenciación en la composición de los elementos móviles analizados. En el caso de las raíces en un entorno de manglar, estas se enfrentan a condiciones de inundación, baja disponibilidad de oxígeno, concentraciones de salinidad altas en suelo y agua (reflejado en la alta conductividad eléctrica en el suelo (30.2 mS/cm) y en el agua (41.4 mS/cm)) y altas temperaturas del agua (30.4 °C) (Anexo 3) (Sudhir *et al.*, 2022; Srikanth *et al.*, 2016). Estas condiciones contrastan con las que afectan a las hojas, como son alta exposición a radiación solar, humedad relativa intermitente, altas temperaturas, y altas concentraciones de taninos y fenoles en el tejido (estos últimos pueden tener una actividad antimicrobiana) (Madhavan *et al.*, 2024).

A nivel funcional, tanto los viromas como los plasmidomas revelaron la presencia de genes relacionados a la adaptación al estrés ambiental, lo cual, teniendo en cuenta las condiciones ambientales a las que enfrenta cada tejido, era congruente la anotación de CDS relacionados con condiciones de choque térmico, sistemas toxina-antitoxina, halotolerancia y resistencia a estrés oxidativo en los genomas virales parcialmente reconstruidos y los plásmidos parcialmente reconstruidos ensamblados. Por su parte, los genomas virales parcialmente reconstruidos mostraron una batería funcional amplia y diversa, lo que sugiere que estos cuentan con la capacidad para modular comunidades bacterianas endófitas y sus interacciones, a través de la transferencia de AMGs, integrasas, recombinasas y enzimas de acción sobre carbohidratos (Enebe & Erasmus, 2024; Braga *et al.*, 2020). Además, como señalan Siyanbola y colaboradores

(2024), la acción de los fagos como agentes de biocontrol puede representar una estrategia valiosa para la agricultura sostenible. Por su parte, los plásmidos parcialmente reconstruidos ensamblados, incluyendo el plásmido completo pZW5A de *Lactiplantibacillus*, reportaron funciones que complementan el metabolismo vegetal, como la regulación del metabolismo de carbohidratos de origen vegetal (maltosa, glucosa, galactosa) y biosíntesis de vitaminas (vitamina B7).

La presencia de CDS en MGEs asociados a funciones adaptativas bacterianas, con capacidad de transferencia horizontal, sugiere que estos genes y funciones podrían transmitirse entre virus, plásmidos y el cromosoma bacteriano. La identificación de sistemas toxina-antitoxina e integrasas, tanto en fagos como en plásmidos, abre la posibilidad de la existencia de red de transferencia horizontal de genes entre estos elementos móviles (Pfeifer & Rocha, 2024; Szczepankowska & Łobocka, 2024), y de ellos con el cromosoma bacteriano (Rodríguez-Beltrán *et al.*, 2021; Colavecchio *et al.*, 2017).

Por otra parte, el conjunto de AMGs identificados en los genomas virales y plásmidos reconstruidos, podría sugerir su aplicación en procesos biotecnológicos diversos. Entre estos AMGs, se encuentran genes asociados a la halotolerancia, el estrés térmico y oxidativo, que podrían ser aprovechados en el diseño de inoculantes microbianos orientados a mejorar la productividad agrícola en suelos salinos, áridos o degradados (AbuQamar *et al.*, 2024). Además, los plásmidos que codifican para rutas de movilización de hierro y biosíntesis de vitaminas vegetales podrían ser de utilidad en la agricultura, con la creación de inóculos bacterianos con efectos bioestimulantes (Woo & Pepe, 2018; AbuQamar *et al.*, 2024). De igual forma, los vMAGs con un alto porcentaje de reconstrucción y con capacidad de transferencia horizontal de genes, podrían explorarse como posibles vectores con genes funcionales de interés y con aplicación de estrategias de biocontrol microbiano (Villalpando-Aguilar *et al.*, 2022), contribuyendo al desarrollo de prácticas agrícolas más sostenibles y menos dependientes de agroquímicos (Holtappels *et al.*, 2021).

Conclusiones generales.

Este estudio proporciona la primera caracterización de la composición del viroma de ADN y del plasmidoma de hojas y raíces de *A. germinans* y *R. mangle*, mostrando el impacto del tipo de tejido en la diversidad y funciones metabólicas asociadas a estos elementos móviles. Los

resultados destacan la relevancia ecológica y funcional de los virus y plásmidos reconstruidos en los microbiomas endófitos de manglar, al identificar funciones adaptativas vinculadas a la halotolerancia, regulación térmica, estrés oxidativo y homeostasis del hierro.

Esta investigación, evidenció el papel de los virus y plásmidos como reservorios de AMGs en entornos endófitos de manglar, y resaltó el potencial que tienen estos MGEs de participar en los procesos de transferencia horizontal de información genética, sugiriendo posibles interacciones entre virus, plásmidos y cromosomas bacterianos. Estos hallazgos refuerzan la importancia de investigar los MGEs en ecosistemas con condiciones variables y, en algunos casos, extremas, como en los manglares, y abren nuevas perspectivas para su aplicación biotecnológica en escenarios de agricultura sostenible y restauración ecológica.

Perspectivas y recomendaciones.

Durante el desarrollo de esta investigación, se puso de manifiesto la escasez de estudios sobre viomas y plasmidomas en muestras ambientales, especialmente en biomas de manglar. Como perspectiva general, se plantea ampliar la línea de investigación de los elementos genéticos móviles (MGEs) en biomas con condiciones fisicoquímicas altamente variables. Dichas investigaciones deberían analizar la identidad taxonómica, la función génica, el impacto y la función ecológica, y plantear y ejecutar aplicaciones biotecnológicas asociadas.

Los futuros trabajos deberían incluir muestreos en distintas estaciones del año y en diversos microambientes del manglar (filosfera, rizosfera, flores, frutos, suelo y agua). Asimismo, es fundamental emplear una mayor profundidad de secuenciación que la utilizada en este macroproyecto (>20 GB para metagenomas y >70X para metatranscriptomas), complementar con tecnologías de secuenciación de lectura larga e implementar técnicas de filtrado durante la extracción de ácidos nucleicos. Estas mejoras permitirían aumentar la cobertura y la calidad de los análisis de diversidad y el ensamblaje de genomas, así como caracterizar en profundidad el papel de virus y plásmidos en las interacciones microbianas de estos ambientes dinámicos y adversos.

Por otro lado, los análisis funcionales de los plásmidos y genomas virales reconstruidos revelaron la presencia de numerosos genes metabólicos auxiliares (AMGs) y genes asociados a la transferencia horizontal. Una perspectiva futura clave es la validación *in vitro* de la actividad de estos genes, esto requeriría ensayos con viriones aislados y con plásmidos introducidos en bacterias modelo, con el fin de

confirmar su función y evaluar su capacidad para alterar la estructura y el funcionamiento del microbioma vegetal endófito.

Aspectos éticos.

Entre los aspectos éticos de esta investigación se encuentran los permisos para la colecta del tejido vegetal y para el uso del material genético.

En el caso de la colecta de tejido foliar y radicular de *Rhizophora mangle* y *Avicennia germinans*, esta se encuentra soportada bajo el “Permiso Marco de Recolección de Especímenes de Especies Silvestres de la Diversidad Biológica con Fines de Investigación Científica No Comercial”, concedido a la Universidad del Valle por la Autoridad Nacional de Licencias Ambientales (ANLA) mediante Resolución 1070 del 28 de agosto de 2015.

Por otra parte, según lo consultado formalmente al Ministerio de Ambiente y Desarrollo Sostenible de Colombia, el macroproyecto del cual hace parte esta investigación no requiere un contrato de acceso a recursos genéticos. Obteniendo la siguiente respuesta por parte de esta entidad:

“Una vez analizado el caso particular objeto de consulta, y a partir de la información proporcionada por el peticionario, se tiene que las actividades descritas NO configuran acceso a recursos genéticos y sus productos derivados en Colombia, dado que se menciona la extracción de ADN vegetal y microbiano asociado al tejido foliar y radicular de las especies *Rhizophora mangle* y *Avicennia germinans* con el fin de realizar análisis de secuenciación genética, para la caracterización de la microbiota y el microbioma que se encuentra asociado a estas especies, las cuales de acuerdo a lo descrito en la metodología aportada, son actividades que corresponden a sistemática molecular y, por lo tanto, este proyecto NO requiere de un contrato de acceso a recursos genéticos”.

En resumen, este estudio cumple con los permisos requeridos asociados a la metodología planteada para resolver la pregunta de investigación y desarrollar cada uno de los objetivos.

Literatura citada.

AbuQamar, S. F., El-Saadony, M. T., Saad, A. M., Desoky, E. S. M., Elrys, A. S., Abd El-Mageed, T. A., ... & El-Tarabily, K. A. (2024). Halotolerant plant growth-promoting rhizobacteria improve soil fertility and plant salinity tolerance for sustainable agriculture—A review. *Plant Stress*, 12, 100482.

Ainsworth, S., Stockdale, S., Bottacini, F., Mahony, J., & van Sinderen, D. (2014). The *Lactococcus lactis* plasmidome: much learnt, yet still lots to discover. *FEMS microbiology reviews*, 38(5), 1066-1088.

Alabi, O. J., Al Rwahnih, M., Brown, J. K., Idris, A. M., Gregg, L., Kmiecik, E., ... & Jifon, J. L. (2016). First report of papaya (*Carica papaya*) naturally infected with the introduced Tomato yellow leaf curl virus-Israel. *Plant Disease*, 100(9), 1959-1959.

Alabi, O. J., Al Rwahnih, M., Jifon, J. L., Gregg, L., Crosby, K. M., & Mirkov, T. E. (2015). First report of Pepper vein yellows virus infecting pepper (*Capsicum* spp.) in the United States. *Plant disease*, 99(11), 1656-1656.

Alibrandi, P., Cardinale, M., Rahman, M. M., Strati, F., Ciná, P., De Viana, M. L., ... & Puglia, A. M. (2018). The seed endosphere of *Anadenanthera colubrina* is inhabited by a complex microbiota, including *Methylobacterium* spp. and *Staphylococcus* spp. with potential plant-growth promoting activities. *Plant and Soil*, 422, 81-99.

Alka, S., Shahir, S., Ibrahim, N., Rahmad, N., Haliba, N., & Abd Manan, F. (2021). Histological and proteome analyses of *Microbacterium foliorum*-mediated decrease in arsenic toxicity in *Melastoma malabathricum*. *3 Biotech*, 11(7), 336.

- Alka, S., Shahir, S., Ibrahim, N., Vo, D. V. N., & Abd Manan, F. (2023). Assessment of plant growth promotion properties and impact of *Microbacterium foliorum* for arsenic removal in *Melastoma malabathricum*. *Bioremediation Journal*, 27(3), 251-262.
- Allander, T., Tammi, M. T., Eriksson, M., Bjerkner, A., Tiveljung-Lindell, A., & Andersson, B. (2005). Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proceedings of the National Academy of Sciences*, 102(36), 12891-12896.
- Alongi, D. M. (2010). Dissolved iron supply limits early growth of estuarine mangroves. *Ecology*, 91(11), 3229-3241.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3), 403-410.
- Anderson, N. G., Gerin, J. L., & Anderson, N. L. (2003). Global screening for human viral pathogens. *Emerging infectious diseases*, 9(7), 768.
- Andersson, U., & Rådström, P. (2002). Physiological function of the maltose operon regulator, MalR, in *Lactococcus lactis*. *BMC microbiology*, 2, 1-7.
- Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Antipov, D., Raiko, M., Lapidus, A., & Pevzner, P. A. (2019). Plasmid detection and assembly in genomic and metagenomic data sets. *Genome research*, 29(6), 961-968.

Antipov, D., Raiko, M., Lapidus, A., & Pevzner, P. A. (2019). Plasmid detection and assembly in genomic and metagenomic data sets. *Genome research*, 29(6), 961-968.

Antipov, D., Raiko, M., Lapidus, A., & Pevzner, P. A. (2020). Metaviral SPAdes: assembly of viruses from metagenomic data. *Bioinformatics*, 36(14), 4126-4129.

Arnold, B. J., Huang, I. T., & Hanage, W. P. (2022). Horizontal gene transfer and adaptive evolution in bacteria. *Nature Reviews Microbiology*, 20(4), 206-218.

Bashan, Y., & de-Bashan, L. E. (2010). How the plant growth-promoting bacterium *Azospirillum* promotes plant growth—a critical assessment. *Advances in agronomy*, 108, 77-136.

Balderas-Ruiz, K. A., Bustos, P., Santamaria, R. I., González, V., Cristiano-Fajardo, S. A., Barrera-Ortiz, S., ... & Serrano-Carreón, L. (2020). *Bacillus velezensis* 83 a bacterial strain from mango phyllosphere, useful for biological control and plant growth promotion. *Amb Express*, 10, 1-19.

Behrendt, U., Ulrich, A., & Schumann, P. (2001). Description of *Microbacterium foliorum* sp. nov. and *Microbacterium phyllosphaerae* sp. nov., isolated from the phyllosphere of grasses and the surface litter after mulching the sward, and reclassification of *Aureobacterium resistens* (Funke et al. 1998) as *Microbacterium resistens* comb. nov.. *International journal of systematic and evolutionary microbiology*, 51(4), 1267-1276.

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M. C. C., Charles, T., ... & Schlöter, M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, 8(1), 1-22.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120.

Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., ... & Woyke, T. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature biotechnology*, 35(8), 725-731.

Braga, L. P., Spor, A., Kot, W., Breuil, M. C., Hansen, L. H., Setubal, J. C., & Philippot, L. (2020). Impact of phages on soil bacterial communities and nitrogen availability under different assembly scenarios. *Microbiome*, 8, 1-14.

Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J. M., Segall, A. M., Mead, D., ... & Rohwer, F. (2002). Genomic analysis of uncultured marine viral communities. *Proceedings of the National Academy of Sciences*, 99(22), 14250-14255.

Breitwieser, F. P., & Salzberg, S. L. (2020). Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics*, 36(4), 1303-1304.

Brown, C. L., Mullet, J., Hindi, F., Stoll, J. E., Gupta, S., Choi, M., Keenum, I., Vikesland, P., Pruden, A., & Zhang, L. (2022). mobileOG-db: a manually curated database of protein families mediating the life cycle of bacterial mobile genetic elements. *Applied and environmental microbiology*, 88(18), e00991-22.

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature methods*, 12(1), 59-60.

Bushmanova, E., Antipov, D., Lapidus, A., & Prjibelski, A. D. (2019). rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience*, 8(9), giz100.

Cai, Z., Li, P., Zhu, W., Wei, J., Lu, J., Song, X., ... & Li, M. (2023). Metagenomic analysis reveals gut plasmids as diagnosis markers for colorectal cancer. *Frontiers in Microbiology*, *14*, 1130446.

Camacho, C., & Madden, T. (2013). BLAST+ release notes. *BLAST® Help [Internet]*.

Cater, K., Dandu, V. S., Bari, S. N., Lackey, K., Everett, G. F., & Hatoum-Aslan, A. (2017). A novel *Staphylococcus podophage* encodes a unique lysin with unusual modular design. *MSphere*, *2*(2), 10-1128.

Cerón-Souza, I., Bermingham, E., McMillan, W. O., & Jones, F. A. (2012). Comparative genetic structure of two mangrove species in Caribbean and Pacific estuaries of Panama. *BMC Evolutionary Biology*, *12*, 1-15.

Chagnot, C., Zorgani, M. A., Astruc, T., & Desvaux, M. (2013). Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective. *Frontiers in microbiology*, *4*, 303.

Chingandu, N., Kouakou, K., Aka, R., Ameyaw, G., Gutierrez, O. A., Herrmann, H. W., & Brown, J. K. (2017). The proposed new species, cacao red vein virus, and three previously recognized badnavirus species are associated with cacao swollen shoot disease. *Virology journal*, *14*(1), 1-14.

Chibani, C. M., Roth, O., Liesegang, H., & Wendling, C. C. (2020). Genomic variation among closely related *Vibrio alginolyticus* strains is located on mobile genetic elements. *Bmc Genomics*, *21*(1), 354.

Choi, Y., Pham, H., Nguyen, M. P., Tran, L. V. H., Kim, J., Kim, S., ... & Kim, Y. H. (2021). A native conjugative plasmid confers potential selective advantages to plant growth-promoting *Bacillus velezensis* strain GH1-13. *Communications Biology*, *4*(1), 582.

- Ciok, A., Budzik, K., Zdanowski, M. K., Gawor, J., Grzesiak, J., Decewicz, P., ... & Dziewit, L. (2018). Plasmids of psychrotolerant *Polaromonas* spp. isolated from arctic and antarctic glaciers—diversity and role in adaptation to polar environments. *Frontiers in Microbiology*, 9, 1285.
- Coetzee, B., Freeborough, M. J., Maree, H. J., Celton, J. M., Rees, D. J. G., & Burger, J. T. (2010). Deep sequencing analysis of viruses infecting grapevines: virome of a vineyard. *Virology*, 400(2), 157-163.
- Colavecchio, A., Cadieux, B., Lo, A., & Goodridge, L. D. (2017). Bacteriophages contribute to the spread of antibiotic resistance genes among foodborne pathogens of the Enterobacteriaceae family—a review. *Frontiers in microbiology*, 8, 1108.
- Collingro, A., Köstlbacher, S., & Horn, M. (2020). Chlamydiae in the environment. *Trends in Microbiology*, 28(11), 877-888.
- Dabdoub, S. (2016). *kraken-biom: Enabling interoperative format conversion for Kraken results (Version 1.2)[Software]*.
- Dargiri, S. A., Naeimi, S., & Nekouei, M. K. (2025). Enhancing wheat resilience to salinity: the role of endophytic *Penicillium chrysogenum* as a biological agent for improved crop performance. *BMC Plant Biology*, 25(1), 354.
- Das, A., & Ljungdahl, L. G. (2003). *Clostridium pasteurianum* F₁F_o ATP synthase: operon, composition, and some properties. *Journal of bacteriology*, 185(18), 5527-5535.
- Dastogeer, K. M., Tumpa, F. H., Sultana, A., Akter, M. A., & Chakraborty, A. (2020). Plant microbiome—an account of the factors that shape community composition and diversity. *Current Plant Biology*, 23, 100161.

DeJong, T. M., Brewin, N. J., & Phillips, D. A. (1981). Effects of plasmid content in *Rhizobium leguminosarum* on pea nodule activity and plant growth. *Microbiology*, 124(1), 1-7.

Ding, H., & Hynes, M. F. (2009). Plasmid transfer systems in the rhizobia. *Canadian Journal of Microbiology*, 55(8), 917-927.

Dreher, T. W. (2010). Viral tRNAs and tRNA-like structures. *Wiley Interdisciplinary Reviews: RNA*, 1(3), 402-414.

Dziewit, L., Pyzik, A., Szuplewska, M., Matlakowska, R., Mielnicki, S., Wibberg, D., ... & Bartosik, D. (2015). Diversity and role of plasmids in adaptation of bacteria inhabiting the Lubin copper mine in Poland, an environment rich in heavy metals. *Frontiers in Microbiology*, 6, 152.

Enebe, M. C., & Erasmus, M. (2024). Disease suppression in plants: A perspective on the impact of phage-directed soil fertilization in enhancing soil health and sustainability. *Biological Control*, 188, 105429.

Erjavec, M. S., & Žgur-Bertok, D. (2006). The RepFIIA replicon of the natural *Escherichia coli* plasmid pRK100. *Acta Biologica Slovenica*, 49(2).

Fang, Z., Tan, J., Wu, S., Li, M., Xu, C., Xie, Z., & Zhu, H. (2019). PPR-Meta: a tool for identifying phages and plasmids from metagenomic fragments using deep learning. *Gigascience*, 8(6), giz066.

Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., Hoffmann, M., Pettengill, J. B., Prasad, A. B., Tillman, G. E., Tyson, G. H. & Klimke, W. (2021). AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Scientific reports*, 11(1), 12728.

Ficko-Blean, E., & Boraston, A. B. (2009). N-acetylglucosamine recognition by a family 32 carbohydrate-binding module from *Clostridium perfringens* NagH. *Journal of molecular biology*, 390(2), 208-220.

Fondi, M., Bacci, G., Brilli, M., Papaleo, M. C., Mengoni, A., Vaneechoutte, M., ... & Fani, R. (2010). Exploring the evolutionary dynamics of plasmids: the *Acinetobacter* pan-plasmidome. *BMC evolutionary biology*, 10, 1-15.

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences*, 103(3), 626-631.

Friess, D. A. (2016). Ecosystem services and disservices of mangrove forests: insights from historical colonial observations. *Forests*, 7(9), 183.

Fukami, J., Cerezini, P., & Hungria, M. (2018). *Azospirillum*: benefits that go far beyond biological nitrogen fixation. *Amb Express*, 8(1), 73.

Geisler, W. M. (2010). Duration of untreated, uncomplicated *Chlamydia trachomatis* genital infection and factors associated with chlamydia resolution: a review of human studies. *The Journal of infectious diseases*, 201, S104-S113.

Geremia, S., Campagnolo, M., Schinzel, R., & Johnson, L. N. (2002). Enzymatic catalysis in crystals of *Escherichia coli* maltodextrin phosphorylase. *Journal of molecular biology*, 322(2), 413-423.

Gómez-Bonilla, Y., López-Ferber, M., Caballero, P., Léry, X., & Muñoz, D. (2012). Costa Rican soils contain highly insecticidal granulovirus strains against *Phthorimaea operculella* and *Tecia solanivora*. *Journal of Applied Entomology*, 136(7), 530-538.

Grant, J. R., Enns, E., Marinier, E., Mandal, A., Herman, E. K., Chen, C. Y., Graham, M., Van Domselaar, G., & Stothard, P. (2023). Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic acids research*, 51(W1), W484–W492. <https://doi.org/10.1093/nar/gkad326>

Gu, Y. H., Tao, X., Lai, X. J., Wang, H. Y., & Zhang, Y. Z. (2014). Exploring the polyadenylated RNA virome of sweet potato through high-throughput sequencing. *PloS one*, 9(6), e98884.

Guglielmini, J., & Van Melderren, L. (2011). Bacterial toxin-antitoxin systems: Translation inhibitors everywhere. *Mobile genetic elements*, 1(4), 283-306.

Gutiérrez-Sánchez, P., Alzate-Restrepo, J., & Marín-Montoya, M. (2014). Characterization of the root RNA-virome in *Solanum phureja* using 454 GS-FLX pyrosequencing. *BioAgro*, 26(2), 89-98.

Habib, C., Yu, Y., Gozzi, K., Ching, C., Shemesh, M., & Chai, Y. (2017). Characterization of the regulation of a plant polysaccharide utilization operon and its role in biofilm formation in *Bacillus subtilis*. *PLoS One*, 12(6), e0179761.

He, W., Russel, J., Klincke, F., Nesme, J., & Sørensen, S. J. (2024). Insights into the ecology of the infant gut plasmidome. *Nature Communications*, 15(1), 6924.

Hernández-Robles, M. F., Álvarez-Contreras, A. K., Juárez-García, P., Natividad-Bonifacio, I., Curiel-Quesada, E., Vázquez-Salinas, C., & Quiñones-Ramírez, E. I. (2016). Virulence factors and antimicrobial resistance in environmental strains of *Vibrio alginolyticus*. *Int Microbiol*, 19(4), 191-198.

Hierholzer, J. C., & Killington, R. A. (1996). Virus isolation and quantitation. In *Virology methods manual* (pp. 25-46). Academic Press.

Hilpert, C., Bricheux, G., & Debroas, D. (2021). Reconstruction of plasmids by shotgun sequencing from environmental DNA: which bioinformatic workflow?. *Briefings in Bioinformatics*, 22(3), bbaa059.

Holmes, D. S., Zhao, H. L., Levican, G., Ratouchniak, J., Bonnefoy, V., Varela, P., & Jedlicki, E. (2001). ISAFE1, an ISL3 family insertion sequence from *Acidithiobacillus ferrooxidans* ATCC 19859. *Journal of bacteriology*, 183(14), 4323-4329.

Holtappels, D., Fortuna, K., Lavigne, R., & Wagemans, J. (2021). The future of phage biocontrol in integrated plant protection for sustainable crop production. *Current Opinion in Biotechnology*, 68, 60-71.

Hooykaas, P. J. J., Van Brussel, A. A. N., den Dulk-Ras, H., Van Slogteren, G. M. S., & Schilperoort, R. A. (1981). Sym plasmid of *Rhizobium trifolii* expressed in different rhizobial species and *Agrobacterium tumefaciens*. *Nature*, 291(5813), 351-353.

Hornik, K. (2012). The comprehensive R archive network. *Wiley interdisciplinary reviews: Computational statistics*, 4(4), 394-398.

Hu, Y., Kwan, B. W., Osbourne, D. O., Benedik, M. J., & Wood, T. K. (2015). Toxin YafQ increases persister cell formation by reducing indole signalling. *Environmental microbiology*, 17(4), 1275-1285.

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, 11, 1-11.

Ibrahim, M. S., Antwerpen, M., Georgi, E., Vette, P., Zoeller, G., & Meyer, H. (2014). Complete genome sequence of the Embu virus strain SPAn880. *Genome Announcements*, 2(6), 10-1128.

Jin, M., Guo, X., Zhang, R., Qu, W., Gao, B., & Zeng, R. (2019). Diversities and potential biogeochemical impacts of mangrove soil viruses. *Microbiome*, 7, 1-15.

Jo, Y., Back, C. G., Kim, K. H., Chu, H., Lee, J. H., Moh, S. H., & Cho, W. K. (2021). Using RNA-Sequencing Data to Examine Tissue-Specific Garlic Microbiomes. *International journal of molecular sciences*, 22(13), 6791. <https://doi.org/10.3390/ijms22136791>

Joshi, H., & Ghose, M. (2003). Forest structure and species distribution along soil salinity and pH gradient in mangrove swamps of the Sundarbans. *Tropical Ecology*, 44(2), 195-204.

Josts, I., Veith, K., Normant, V., Schalk, I. J., & Tidow, H. (2021). Structural insights into a novel family of integral membrane siderophore reductases. *Proceedings of the National Academy of Sciences*, 118(34), e2101952118.

Kandel, S. L., Joubert, P. M., & Doty, S. L. (2017). Bacterial endophyte colonization and distribution within plants. *Microorganisms*, 5(4), 77.

Katsy, E. I. (2011). Plasmid plasticity in the plant-associated bacteria of the genus *Azospirillum*. *Bacteria in agrobiolgy: plant growth responses*, 139-157.

Kaushik, P., Rawat, N., Mathur, M., Raghuvanshi, P., Bhatnagar, P., Swarnkar, H., & Flora, S. (2012). Arsenic hyper-tolerance in four *Microbacterium* species isolated from soil contaminated with textile effluent. *Toxicology International*, 19(2), 188.

Kav, A. B., Sasson, G., Jami, E., Doron-Faigenboim, A., Benhar, I., & Mizrahi, I. (2012). Insights into the bovine rumen plasmidome. *Proceedings of the National Academy of Sciences*, 109(14), 5452-5457.

- Kim, D. H., Jeon, Y. J., Chung, M. J., Seo, J. G., & Ro, Y. T. (2017). Complete sequence and gene analysis of a cryptic plasmid pLU4 in *Lactobacillus reuteri* strain LU4 (KCTC 12397BP). *Applied Biological Chemistry*, 60, 145-153.
- Klein, G., Kobylak, N., Lindner, B., Stupak, A., & Raina, S. (2014). Assembly of lipopolysaccharide in *Escherichia coli* requires the essential LapB heat shock protein. *Journal of Biological Chemistry*, 289(21), 14829-14853.
- Kolde, R. (2015). Package 'pheatmap'. *R package*, 1(7), 790.
- Lagares, A., Sanjuán, J., & Pistorio, M. (2014). The plasmid mobilome of the model plant-symbiont *Sinorhizobium meliloti*: coming up with new questions and answers. *Microbiology spectrum*, 2(5), 10-1128.
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9(4), 357-359.
- Latka, A., Maciejewska, B., Majkowska-Skrobek, G., Briers, Y., & Drulis-Kawa, Z. (2017). Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. *Applied microbiology and biotechnology*, 101, 3103-3119.
- LeRoux, M., & Laub, M. T. (2022). Toxin-antitoxin systems as phage defense elements. *Annual Review of Microbiology*, 76(1), 21-43.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *bioinformatics*, 25(16), 2078-2079.

Lipworth, S., Matlock, W., Shaw, L., Vihta, K. D., Rodger, G., Chau, K., ... & Stoesser, N. (2024). The plasmidome associated with Gram-negative bloodstream infections: A large-scale observational study using complete plasmid assemblies. *Nature Communications*, 15(1), 1612.

Liu, H., Dou, W., Zhang, W., Li, Y., Shi, G., Fu, L., & Wang, L. (2025). Analysis of the Synergistic Effect of Endophytic Bacteria and Metabolites on Yield in Soybean Roots Under Biochar with Nitrogen Fertilization. *Agronomy*, 15(1), 177.

Liu, Y., Xu, Z., Chen, L., Xun, W., Shu, X., Chen, Y., ... & Zhang, R. (2024). Root colonization by beneficial rhizobacteria. *FEMS Microbiology Reviews*, 48(1).

Luo, P., Jiang, H., Wang, Y., Su, T., Hu, C., Ren, C., & Jiang, X. (2012). Prevalence of mobile genetic elements and transposase genes in *Vibrio alginolyticus* from the southern coastal region of China and their role in horizontal gene transfer. *Int Microbiol*, 15(201), 10.

Luo, X. Q., Wang, P., Li, J. L., Ahmad, M., Duan, L., Yin, L. Z., ... & Li, W. J. (2022). Viral community-wide auxiliary metabolic genes differ by lifestyles, habitats, and hosts. *Microbiome*, 10(1), 1-18.

Madhavan, C., Meera, S. P., & Kumar, A. (2024). Anatomical adaptations of mangroves to the intertidal environment and their dynamic responses to various stresses. *Biological Reviews*.

Markowska-Barkic, A., Lewicka, E., Czeredys, M., Mitura, M., & Jagura-Burdzy, G. (2022). Deciphering the Regulatory Circuits of RA3 Replication Module-Mechanisms of the Copy Number Control. *International Journal of Molecular Sciences*, 23(17), 9964.

Marti, H., Suchland, R. J., & Rockey, D. D. (2022). The impact of lateral gene transfer in *Chlamydia*. *Frontiers in cellular and infection microbiology*, 12, 861899.

- Martinez-Bond, E. A., Soriano, B. M., & Williams, A. H. (2022). The mechanistic landscape of lytic transglycosylase as targets for antibacterial therapy. *Current Opinion in Structural Biology*, 77, 102480.
- McDaniel, L., Breitbart, M., Mobberley, J., Long, A., Haynes, M., Rohwer, F., & Paul, J. H. (2008). Metagenomic analysis of lysogeny in Tampa Bay: implications for prophage gene expression. *PLoS One*, 3(9), e3263.
- 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.
- Mészárošová, L., Kuřáková, E., Kohout, P., Münzbergová, Z., & Baldrian, P. (2024). Plant effects on microbiome composition are constrained by environmental conditions in a successional grassland. *Environmental Microbiome*, 19(1), 8.
- Meziti, A., Rodriguez-R, L. M., Hatt, J. K., Peña-Gonzalez, A., Levy, K., & Konstantinidis, K. T. (2021). The reliability of metagenome-assembled genomes (MAGs) in representing natural populations: insights from comparing MAGs against isolate genomes derived from the same fecal sample. *Applied and environmental microbiology*, 87(6), e02593-20.
- Mihara, T., Nishimura, Y., Shimizu, Y., Nishiyama, H., Yoshikawa, G., Uehara, H., ... & Ogata, H. (2016). Linking virus genomes with host taxonomy. *Viruses*, 8(3), 66.
- Mochizuki, S., Hiratsu, K., Suwa, M., Ishii, T., Sugino, F., Yamada, K., & Kinashi, H. (2003). The large linear plasmid pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary metabolism. *Molecular microbiology*, 48(6), 1501-1510.

- Molinatto, G., Puopolo, G., Sonogo, P., Moretto, M., Engelen, K., Viti, C., ... & Pertot, I. (2016). Complete genome sequence of *Bacillus amyloliquefaciens* subsp. *plantarum* S499, a rhizobacterium that triggers plant defences and inhibits fungal phytopathogens. *Journal of biotechnology*, 238, 56-59.
- Moreno-Gallego, J. L., & Reyes, A. (2021). Informative regions in viral genomes. *Viruses*, 13(6), 1164.
- Naini, A., Sasse, F., & Brönstrup, M. (2019). The intriguing chemistry and biology of soraphens. *Natural Product Reports*, 36(10), 1394-1411.
- Nayfach, S., Camargo, A. P., Schulz, F., Eloie-Fadrosh, E., Roux, S., & Kyrpides, N. C. (2021). CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nature biotechnology*, 39(5), 578-585.
- Newbury, A., Dawson, B., Klümper, U., Hesse, E., Castledine, M., Fontaine, C., ... & Sanders, D. (2022). Fitness effects of plasmids shape the structure of bacteria–plasmid interaction networks. *Proceedings of the National Academy of Sciences*, 119(22), e2118361119.
- Nie, X., Zhao, Z., Zhang, X., Bastías, D. A., Nan, Z., & Li, C. (2024). Endophytes alleviate drought-derived oxidative damage in *Achnatherum inebrians* plants through increasing antioxidants and regulating host stress responses. *Microbial Ecology*, 87(1), 73.
- Niehus, R., Mitri, S., Fletcher, A. G., & Foster, K. R. (2015). Migration and horizontal gene transfer divide microbial genomes into multiple niches. *Nature communications*, 6(1), 8924.
- Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome research*, 27(5), 824-834.

Ó Cuív, P., Keogh, D., Clarke, P., & O'Connell, M. (2007). FoxB of *Pseudomonas aeruginosa* functions in the utilization of the xenosiderophores ferrichrome, ferrioxamine B, and schizokinen: evidence for transport redundancy at the inner membrane. *Journal of bacteriology*, 189(1), 284-287.

Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Borman T, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2025). *vegan: Community Ecology Package*. R package version 2.7-0, <https://github.com/vegandevs/vegan>, <https://vegandevs.github.io/vegan/>.

Olasz, F., Szabó, M., Veress, A., Bibó, M., & Kiss, J. (2022). The dynamic network of IS 30 transposition pathways. *Plos one*, 17(7), e0271414.

Olayemi, O. P., & Odedara, O. O. (2017). Screening of endophytic plant growth-promoting bacteria isolated from two Nigerian rice varieties. *Nigerian Journal of Biotechnology*, 33, 1-10.

Orland, C., Emilson, E. J., Basiliko, N., Mykytczuk, N. C., Gunn, J. M., & Tanentzap, A. J. (2019). Microbiome functioning depends on individual and interactive effects of the environment and community structure. *The ISME Journal*, 13(1), 1-11.

Paez-Espino, D., Eloie-Fadrosch, E. A., Pavlopoulos, G. A., Thomas, A. D., Huntemann, M., Mikhailova, N., ... & Kyrpides, N. C. (2016). Uncovering Earth's virome. *Nature*, 536(7617), 425-430.

Palit, K., Rath, S., Chatterjee, S., & Das, S. (2022). Microbial diversity and ecological interactions of microorganisms in the mangrove ecosystem: Threats, vulnerability, and adaptations. *Environmental Science and Pollution Research*, 29(22), 32467-32512.

Palm, D., Goerl, R., Weidinger, G., Zeier, R., & Fischer, B. (1987). E. coli maltodextrin phosphorylase: primary structure and deletion mapping of the C-terminal site. *Zeitschrift für Naturforschung C*, 42(4), 394-400.

Papayova, K., Bocanova, L., Bauerova, V., Bauer, J., Halgasova, N., Kajsikova, M., & Bukovska, G. (2025). From sequence to function: Exploring biophysical properties of bacteriophage BFK20 lytic transglycosylase domain from the minor tail protein gp15. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1873(1), 141044.

Patz, S., Gautam, A., Becker, M., Ruppel, S., Rodríguez-Palenzuela, P., & Huson, D. H. (2021). PLaBAs: A comprehensive web resource for analyzing the plant growth-promoting potential of plant-associated bacteria. *BioRxiv*, 2021-12.

Pellegrinetti, T. A., Monteiro, G. G. T. N., Lemos, L. N., dos Santos, R. A. C., Barros, A. G., & Mendes, L. W. (2024). PGPg_finder: a comprehensive and user-friendly pipeline for identifying plant growth-promoting genes in genomic and metagenomic data. *Rhizosphere*, 30, 100905.

Peltier, J., Hamiot, A., Garneau, J. R., Boudry, P., Maikova, A., Hajnsdorf, E., ... & Soutourina, O. (2020). Type I toxin-antitoxin systems contribute to the maintenance of mobile genetic elements in *Clostridioides difficile*. *Communications biology*, 3(1), 718.

Pfeifer, E., & Rocha, E. P. (2024). Phage-plasmids promote recombination and emergence of phages and plasmids. *Nature Communications*, 15(1), 1545.

Plyusnin, I., Kant, R., Jääskeläinen, A. J., Sironen, T., Holm, L., Vapalahti, O., & Smura, T. (2020). Novel NGS pipeline for virus discovery from a wide spectrum of hosts and sample types. *Virus evolution*, 6(2), veaa091.

Prysak, M. H., Mozdierz, C. J., Cook, A. M., Zhu, L., Zhang, Y., Inouye, M., & Woychik, N. A. (2009). Bacterial toxin YafQ is an endoribonuclease that associates with the ribosome and blocks

translation elongation through sequence-specific and frame-dependent mRNA cleavage. *Molecular microbiology*, 71(5), 1071-1087.

Qi, Q., Rajabal, V., Ghaly, T. M., Tetu, S. G., & Gillings, M. R. (2023). Identification of integrons and gene cassette-associated recombination sites in bacteriophage genomes. *Frontiers in microbiology*, 14, 1091391.

Qiu, J., Zhai, Y., Wei, M., Zheng, C., & Jiao, X. (2022). Toxin–antitoxin systems: classification, biological roles, and applications. *Microbiological research*, 264, 127159.

R Core Team (2025). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Austria. <https://www.R-project.org/>

Ranjan Mishra, R., Ranjan Swain, M., Kanti Danga, T., & Thatoi, H. (2012). Diversity and seasonal fluctuation of predominant microbial communities in Bhitarkanika, a tropical mangrove ecosystem in India. *Revista de biología tropical*, 60(2), 909-924.

Rangel-Pineros, G., Almeida, A., Beracochea, M., Sakharova, E., Marz, M., Reyes Muñoz, A., Hölzer, M. & Finn, R. (2023) VIRify: An integrated detection, annotation and taxonomic classification pipeline using virus-specific protein profile hidden Markov models. *PLoS Comput Biol*, 19(8): e1011422.

Reid, S., De Malmanche, H., Chan, L., Popham, H., & Van Oers, M. M. (2023). Production of entomopathogenic viruses. In *Mass production of beneficial organisms* (pp. 375-406). Academic Press.

Reinhold-Hurek, B., & Hurek, T. (2011). Living inside plants: bacterial endophytes. *Current opinion in plant biology*, 14(4), 435-443.

Ren, J., Ahlgren, N. A., Lu, Y. Y., Fuhrman, J. A., & Sun, F. (2017). VirFinder: a novel k-mer based tool for identifying viral sequences from assembled metagenomic data. *Microbiome*, 5, 1-20.

Robertson, J., & Nash, J. H. (2018). MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microbial genomics*, 4(8), e000206.

Rodríguez-Beltrán, J., DelaFuente, J., León-Sampedro, R., MacLean, R. C., & San Millán, Á. (2021). Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. *Nature Reviews Microbiology*, 19(6), 347-359.

Rodriguez-R, L. M., & Konstantinidis, K. T. (2014). Estimating coverage in metagenomic data sets and why it matters. *The ISME journal*, 8(11), 2349-2351.

Rohrmann, G. F. (2019). Baculoviruses as insecticides: Four examples. *Baculovirus Molecular Biology (4th ed.)*. National Center for Biotechnology Information (US).

Rosenblueth, M., & Martínez-Romero, E. (2006). Bacterial endophytes and their interactions with hosts. *Molecular plant-microbe interactions*, 19(8), 827-837.

Roux, S., Enault, F., Hurwitz, B. L., & Sullivan, M. B. (2015). VirSorter: mining viral signal from microbial genomic data. *PeerJ*, 3, e985.

Sahu, P. K., Tilgam, J., Mishra, S., Hamid, S., Gupta, A., K, J., Verma, S., K. & Kharwar, R. N. (2022). Surface sterilization for isolation of endophytes: Ensuring what (not) to grow. *Journal of Basic Microbiology*, 62(6), 647-668.

Sánchez, P. A. G., Mesa, H. J., & Montoya, M. M. (2016). Next generation sequence analysis of the forage peanut (*Arachis pintoi*) virome. *Revista Facultad Nacional de Agronomía Medellín*, 69(2), 7881-7891.

Sbardellati, D. L., & Vannette, R. L. (2024). Targeted viromes and total metagenomes capture distinct components of bee gut phage communities. *Microbiome*, 12(1), 155.

Scharf, B. E., Hynes, M. F., & Alexandre, G. M. (2016). Chemotaxis signaling systems in model beneficial plant–bacteria associations. *Plant molecular biology*, 90, 549-559.

Schierstaedt, J., Bziuk, N., Kuzmanović, N., Blau, K., Smalla, K., & Jechalke, S. (2019). Role of plasmids in plant-bacteria interactions. *Current Issues in Molecular Biology*, 30(1), 17-38.

Smalla, K., & Sobecky, P. A. (2002). The prevalence and diversity of mobile genetic elements in bacterial communities of different environmental habitats: insights gained from different methodological approaches. *FEMS Microbiology ecology*, 42(2), 165-175.

Sessitsch, A., Howieson, J. G., Perret, X., Antoun, H., & Martínez-Romero, E. (2002). Advances in Rhizobium research. *Critical Reviews in Plant Sciences*, 21(4), 323-378.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), 2068-2069.

Siguier, P., Gourbeyre, E., & Chandler, M. (2014). Bacterial insertion sequences: their genomic impact and diversity. *FEMS microbiology reviews*, 38(5), 865-891.

Singh, G., Yadav, M., Ghosh, C., & Rathore, J. S. (2021). Bacterial toxin-antitoxin modules: classification, functions, and association with persistence. *Current Research in Microbial Sciences*, 2, 100047.

Siyanbola, K. F., Ejiohuo, O., Ade-adekunle, O. A., Adekunle, F. O., Onyeaka, H., Furr, C. L. L., ... & Oladipo, E. K. (2024). Bacteriophages: sustainable and effective solution for climate-resilient agriculture. *Sustainable Microbiology*, 1(1), qvae025.

Smits, S. L., Bodewes, R., Ruiz-González, A., Baumgärtner, W., Koopmans, M. P., Osterhaus, A. D., & Schürch, A. C. (2015). Recovering full-length viral genomes from metagenomes. *Frontiers in microbiology*, 6, 1069.

Sonika, S., Singh, S., Mishra, S., & Verma, S. (2023). Toxin-antitoxin systems in bacterial pathogenesis. *Heliyon*, 9(4).

Song, X., Sun, J., Mikalsen, T., Roberts, A. P., & Sundsfjord, A. (2013). Characterisation of the plasmidome within *Enterococcus faecalis* isolated from marginal periodontitis patients in Norway. *PLoS One*, 8(4), e62248.

Soontharapirakkul, K., Promden, W., Yamada, N., Kageyama, H., Incharoensakdi, A., Iwamoto-Kihara, A., & Takabe, T. (2011). Halotolerant cyanobacterium *Aphanothece halophytica* contains an Na⁺-dependent F1F0-ATP synthase with a potential role in salt-stress tolerance. *Journal of Biological Chemistry*, 286(12), 10169-10176.

Sorokan, A., Cherepanova, E., Burkhanova, G., Veselova, S., Rumyantsev, S., Alekseev, V., ... & Maksimov, I. (2020). Endophytic *Bacillus* spp. as a prospective biological tool for control of viral diseases and non-vector *Leptinotarsa decemlineata* Say. in *Solanum tuberosum* L. *Frontiers in Microbiology*, 11, 569457.

Squeglia, F., Maciejewska, B., Łatka, A., Ruggiero, A., Briers, Y., Drulis-Kawa, Z., & Berisio, R. (2020). Structural and functional studies of a Klebsiella phage capsule depolymerase tailspike: mechanistic insights into capsular degradation. *Structure*, 28(6), 613-624.

Srikanth, S., Lum, S. K. Y., & Chen, Z. (2016). Mangrove root: adaptations and ecological importance. *Trees*, 30, 451-465.

Stockdale, S. R., & Hill, C. (2023). Incorporating plasmid biology and metagenomics into a holistic model of the human gut microbiome. *Current Opinion in Microbiology*, 73, 102307.

Su, Y., Yu, H., Gao, C., Sun, S., Liang, Y., Liu, G., Zhang, X., Dong, Y., Liu, X., Chen, G., Shao, H., McMinn, A., & Wang, M. (2024). Effects of vegetation cover and aquaculture pollution on viral assemblages in mangroves sediments. *Journal of Hazardous Materials*, 476, 135147.

Sudhir, S., Arunprasath, A., & Vel, V. S. (2022). A critical review on adaptations, and biological activities of the mangroves. *Journal of Natural Pesticide Research*, 1, 100006.

Szczepankowska, A. K., & Łobocka, M. (2024). Exploring the role of phage plasmids in gene transfers. *Trends in Genetics*, 40(7), 555–557. <https://doi.org/10.1016/J.TIG.2024.04.011>

Taghavi, S., Barac, T., Greenberg, B., Borremans, B., Vangronsveld, J., & van der Lelie, D. (2005). Horizontal gene transfer to endogenous endophytic bacteria from poplar improves phytoremediation of toluene. *Applied and environmental microbiology*, 71(12), 8500-8505.

Tan, B., Li, Y., Xie, H., Dai, Z., Zhou, C., Qian, Z. J., ... & Li, C. (2022). Microplastics accumulation in mangroves increasing the resistance of its colonization Vibrio and Shewanella. *Chemosphere*, 295, 133861.

Taylor-Brown, A., Vaughan, L., Greub, G., Timms, P., & Polkinghorne, A. (2015). Twenty years of research into Chlamydia-like organisms: a revolution in our understanding of the biology and pathogenicity of members of the phylum Chlamydiae. *FEMS Pathogens and Disease*, 73(1), 1-15

Taylor-Brown, A., & Polkinghorne, A. (2017). New and emerging chlamydial infections of creatures great and small. *New Microbes and New Infections*, 18, 28-33.

Tokuda, M., & Shintani, M. (2024). Microbial evolution through horizontal gene transfer by mobile genetic elements. *Microbial Biotechnology*, 17(1), e14408.

Trivedi, P., Batista, B. D., Bazany, K. E., & Singh, B. K. (2022). Plant–microbiome interactions under a changing world: responses, consequences and perspectives. *New Phytologist*, 234(6), 1951-1959.

Turner, T. R., James, E. K., & Poole, P. S. (2013). The plant microbiome. *Genome biology*, 14, 1-10.

Umene, K., & Shiraishi, A. (2013). Complete nucleotide sequence of *Bacillus subtilis* (natto) bacteriophage PM1, a phage associated with disruption of food production. *Virus Genes*, 46(3), 524-534.

Underwood, G. J., Dumbrell, A. J., McGenity, T. J., McKew, B. A., & Whitby, C. (2022). The microbiome of coastal sediments. *The marine microbiome*, 479-534.

Van Kranenburg, R., Golic, N., Bongers, R., Leer, R. J., De Vos, W. M., Siezen, R. J., & Kleerebezem, M. (2005). Functional analysis of three plasmids from *Lactobacillus plantarum*. *Applied and environmental microbiology*, 71(3), 1223-1230.

Van Mastrigt, O., Di Stefano, E., Hartono, S., Abee, T., & Smid, E. J. (2018). Large plasmidome of dairy *Lactococcus lactis* subsp. *lactis* biovar *diacetylactis* FM03P encodes technological functions and appears highly unstable. *BMC genomics*, *19*(1), 1-19.

Van Oirschot, J. T. (1999). Bovine viral vaccines, diagnostics, and eradication: past, present, and future. *Advances in Veterinary Medicine*, *41*, 197-216.

Vergnes, S., Gayraud, D., Veyssière, M., Toulotte, J., Martinez, Y., Dumont, V., ... & Dumas, B. (2020). Phyllosphere colonization by a soil *Streptomyces* sp. promotes plant defense responses against fungal infection. *Molecular Plant-Microbe Interactions*, *33*(2), 223-234.

Vernikos, G. S., & Parkhill, J. (2006). Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands. *Bioinformatics*, *22*(18), 2196-2203.

Villalpando-Aguilar, J. L., Matos-Pech, G., López-Rosas, I., Castelán-Sánchez, H. G., & Alatorre-Cobos, F. (2022). Phage therapy for crops: concepts, experimental and bioinformatics approaches to direct its application. *International journal of molecular sciences*, *24*(1), 325.

Vincent, A. T., Emond-Rheault, J. G., Barbeau, X., Attéré, S. A., Frenette, M., Lagüe, P., & Charette, S. J. (2016). Antibiotic resistance due to an unusual ColE1-type replicon plasmid in *Aeromonas salmonicida*. *Microbiology*, *162*(6), 942-953.

Walker, A. (2012). Welcome to the plasmidome. *Nature Reviews Microbiology*, *10*(6), 379-379.

Wang, H., Liu, Y., Bai, C., & Leung, S. S. Y. (2024). Translating bacteriophage-derived depolymerases into antibacterial therapeutics: Challenges and prospects. *Acta Pharmaceutica Sinica B*, *14*(1), 155-169.

Wang, Y., Xiao, M., Geng, X., Liu, J., & Chen, J. (2007). Horizontal transfer of genetic determinants for degradation of phenol between the bacteria living in plant and its rhizosphere. *Applied Microbiology and Biotechnology*, 77, 733-739.

Wang, X., Xue, Y., Han, M., Bu, Y., & Liu, C. (2014). The ecological roles of *Bacillus thuringiensis* within phyllosphere environments. *Chemosphere*, 108, 258-264.

Wickham, H. (2016). Package ‘ggplot2’. *Create elegant data visualisations using the grammar of graphics. Version*, 2(1), 1-189.

Will, V., Frey, C., Normant, V., Kuhn, L., Chicher, J., Volck, F., & Schalk, I. J. (2024). The role of FoxA, FiuA, and FpvB in iron acquisition via hydroxamate-type siderophores in *Pseudomonas aeruginosa*. *Scientific Reports*, 14(1), 18795.

Williamson, K. E., Radosevich, M., & Wommack, K. E. (2005). Abundance and diversity of viruses in six Delaware soils. *Applied and environmental microbiology*, 71(6), 3119-3125.

Wishart, D. S., Han, S., Saha, S., Oler, E., Peters, H., Grant, J. R., Stothard, P., & Gautam, V. (2023). PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Research*, 51(W1), W443-W450.

Woo, S. L., & Pepe, O. (2018). Microbial consortia: promising probiotics as plant biostimulants for sustainable agriculture. *Frontiers in plant science*, 9, 1801.

Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome biology*, 20, 1-13.

Yamamoto, H., Serizawa, M., Thompson, J., & Sekiguchi, J. (2001). Regulation of the glv operon in *Bacillus subtilis*: YfiA (GlvR) is a positive regulator of the operon that is repressed through CcpA and cre. *Journal of bacteriology*, 183(17), 5110-5121.

Yao, J., Guo, Y., Wang, P., Zeng, Z., Li, B., Tang, K., ... & Wang, X. (2018). Type II toxin/antitoxin system ParESO/CopASO stabilizes prophage CP4So in *Shewanella oneidensis*. *Environmental microbiology*, 20(3), 1224-1239.

Zahran, H. H. (2017). Plasmids impact on rhizobia-legumes symbiosis in diverse environments. *Symbiosis*, 73(2), 75-91.

Zboralski, A., & Filion, M. (2020). Genetic factors involved in rhizosphere colonization by phytobeneficial *Pseudomonas* spp. *Computational and Structural Biotechnology Journal*, 18, 3539-3554.

Zhang, Y., Ding, C. T., Jiang, T., Liu, Y. H., Wu, Y., Zhou, H. W., ... & Chen, Y. (2023). Community structure and niche differentiation of endosphere bacterial microbiome in *Camellia oleifera*. *Microbiology Spectrum*, 11(6), e01335-23.

Zhang, M., Zhao, Y., & Shi, H. (2023). Metabolomics Reveals the Effects of Lighting Time on the Growth and Development of Flue-Cured Tobacco. *Russian Journal of Plant Physiology*, 70(3), 30.

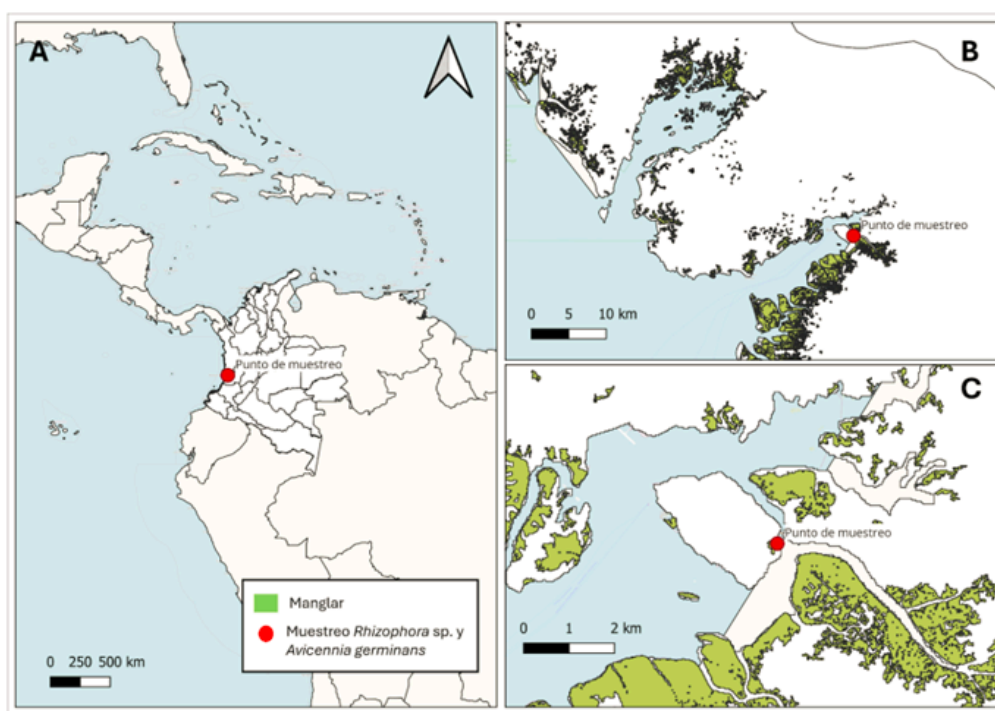
Zheng, J., Ge, Q., Yan, Y., Zhang, X., Huang, L., & Yin, Y. (2023). dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Research*, 51(W1), W115-W121.

Zheng, Y., Cao, X., Zhou, Y., Ma, S., Wang, Y., Li, Z., ... & Zhang, C. S. (2024). Purines enrich root-associated *Pseudomonas* and improve wild soybean growth under salt stress. *Nature communications*, 15(1), 3520

Zhu, Y., Li, R., Yan, S., Chen, X., Cen, S., & Xie, S. (2024). Habitat-and lifestyle-dependent structural and functional characteristics of viruses in mangrove wetlands of different functional zonings. *Environmental Research*, 252, 119070.

Anexos

Anexo 1. Ubicación de la zona de muestreo. A). Bahía de Buenaventura en el Suroccidente colombiano. B) Desembocadura del río Dagua al océano Pacífico. C) Zona de muestreo dentro del Bosque de manglar.



Anexo 2. Parámetros meteorológicos registrados en la zona de muestreo el 14 de junio de 2023 según la Dimar (Dirección General Marítima Portuaria de Colombia).

Parámetro meteorológico	Resultado	Unidad
Temperatura	30.2	°C
Índice de calor	39	°C
Índice de frío	30	°C
Punto de rocío	26.3	°C
Búlbo húmedo	27.3	°C
Humedad	78	%
Presión	14.65	PSI
Barómetro	13.24	PSI
Altura	-9.5	msnm

Anexo 3. Parámetros fisicoquímicos del suelo y del agua estuarina de la zona de muestreo.

Parámetros fisicoquímicos del suelo			Parámetros fisicoquímicos del agua		
Parámetro	Resultado	Unidad	Parámetro	Resultado	Unidad
pH	5.91	pH	pH	7.42	pH
Conductividad eléctrica	30.2	dS/m	Potencial redox ORP	119.4	mV
Carbono orgánico oxidable	1.78	%	Saturación de oxígeno	81.3	%
Materia orgánica	3.07	%	Oxígeno disuelto	5.14	ppm
Nitrógeno total	0.148	%	Conductividad Ω	41.4	mS/cm
Potasio intercambiable	642	mg/kg	Conductividad Ω	71.7	mS/cmA
Calcio intercambiable	792	mg/kg	Resistividad	0.8	Ω -cm
Magnesio intercambiable	994	mg/kg	Sólidos totales disueltos TDS	24.14	ppm
Sodio intercambiable	4480	mg/kg	Salinidad PSU	18.5	σ T
Fósforo	20.5	mg/kg	Turbidez	38.4	cm
Azufre	489	mg/kg	Temperatura	30.4	°C

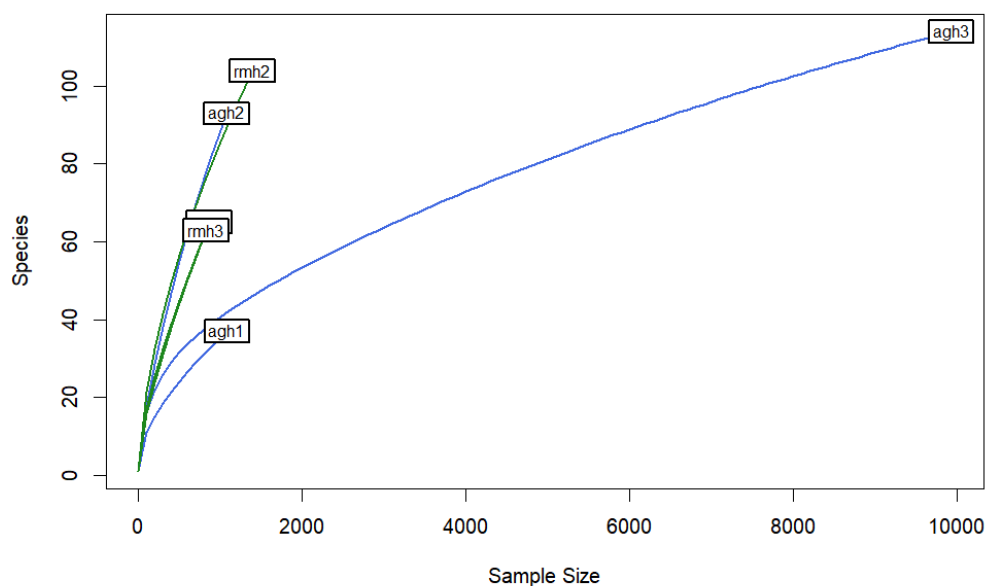
Anexo 4. Listado de accesiones del NCBI para las muestras metagenómicas limpias (Bioproject number: PRJNA1215819).

Accession	Sample name	Tissue
SAMN46420086	RmR1d	Root

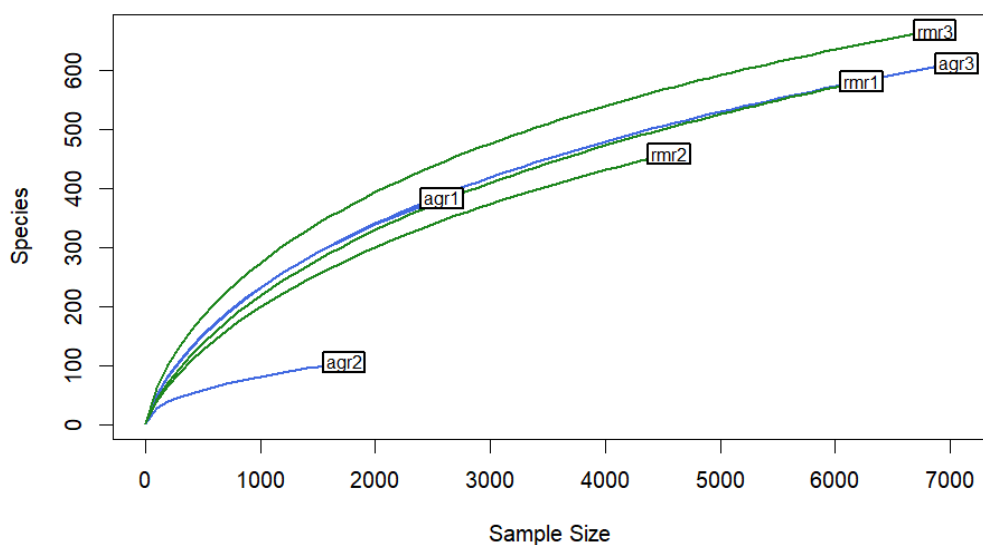
SAMN46420087	RmR2d	Root
SAMN46420088	RmR3d	Root
SAMN46420089	RmH1d	Leaf
SAMN46420090	RmH2d	Leaf
SAMN46420091	RmH3d	Leaf
SAMN46420097	AgR1d	Root
SAMN46420098	AgR2d	Root
SAMN46420099	AgR3d	Root
SAMN46420100	AgH1d	Leaf
SAMN46420101	AgH2d	Leaf
SAMN46420102	AgH3d	Leaf

Anexo 5. Curvas de rarefacción de especies virales para cada réplica. Se describen curvas de cantidad de especies virales en función de la cantidad de secuencias. A. Tejido foliar de muestras de *R. mangle* (Rm) y *A. germinans* (Ag). B. Tejido radicular de *R. mangle* (Rm) y *A. germinans* (Ag).

A.



B.



Anexo 6. Análisis multivariado permutado de varianza (PERMANOVA) de las comunidades virales endófitas de ADN, con los factores especie, tejido e interacción entre especie y tejido. Este análisis contó con 999 permutaciones y con un valor α de referencia de 0.05.

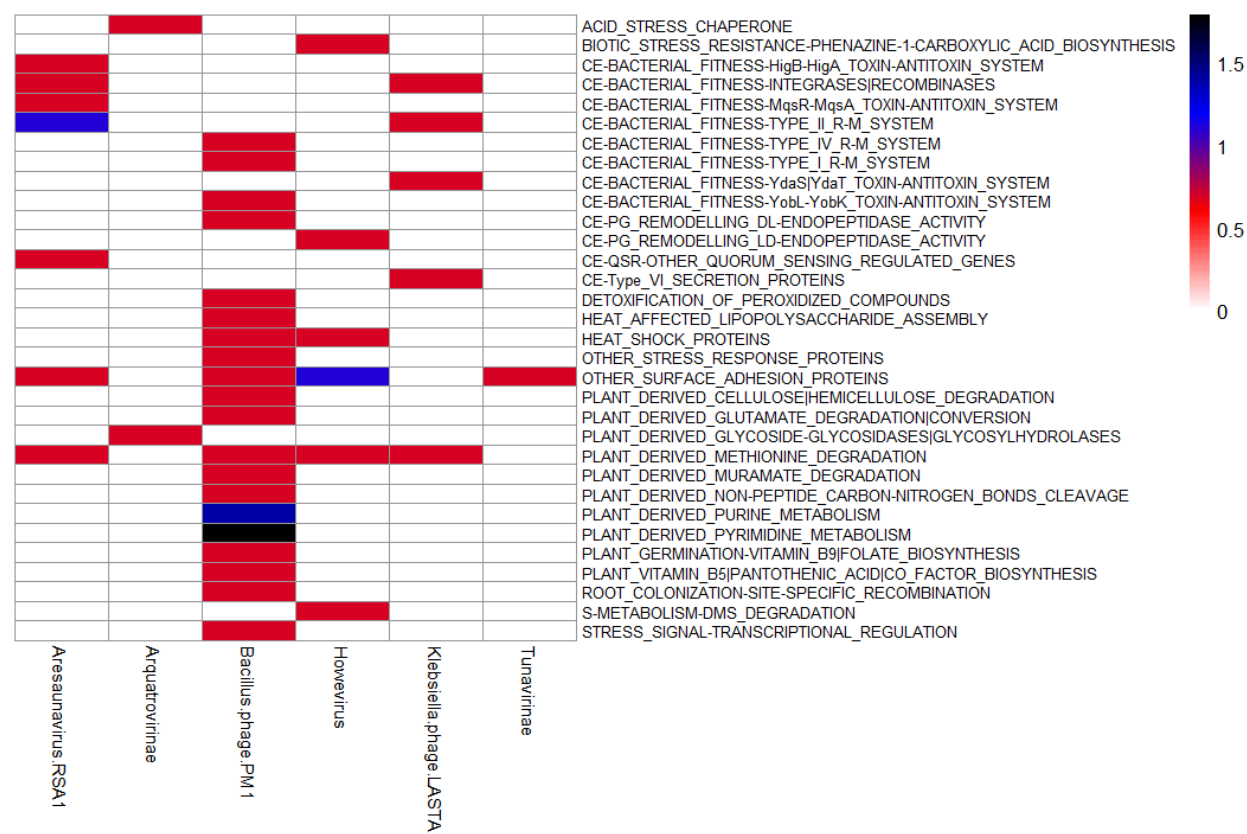
Factors	Df	SumOfSqs	R ²	F	Pr(>F)	
<i>Specie</i>	1	0,19407	0,07561	1,1132	0,303	
<i>Tissue</i>	1	0,77241	0,30094	4,4306	0,002	**
<i>Specie:Tissue</i>	1	0,20547	0,08005	1,1786	0,233	
<i>Residual</i>	8	1,39469	0,54339			
<i>Total</i>	11	2,56664	1			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

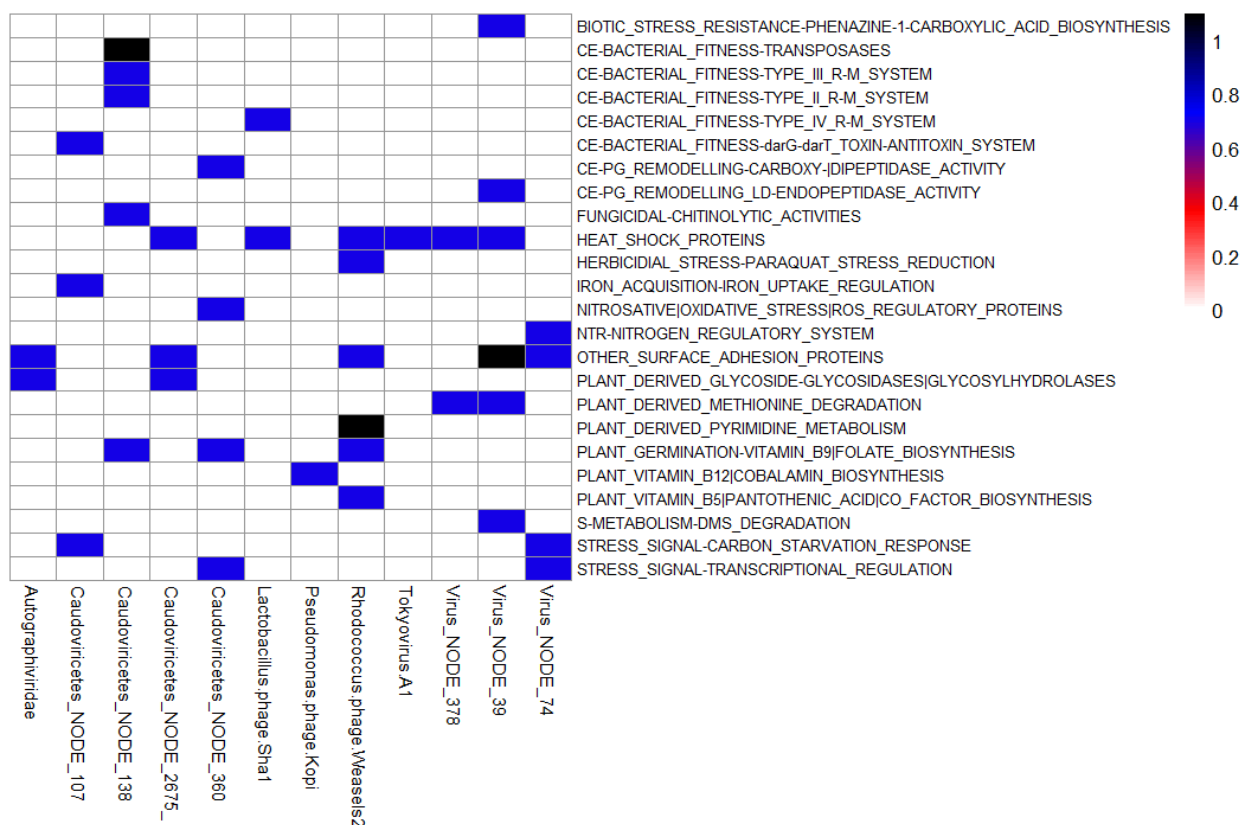
Anexo 7. Anotación de categorías génicas (Lv5) involucradas en la relación microbiota-planta, dentro de los genomas virales reconstruidos de tejidos de *A. germinans* (A.) y *R. mangle* (B.). En los genomas analizados se evidencia la abundancia de CDS, clasificados en las categorías del nivel 5 (Lv5) de la base de datos de PLaBase, los cuales están relacionados con funciones metabólicas implicadas en la interacción planta-microorganismo. En la escala ubicada a la derecha de la figura, el color negro

determina una alta abundancia de secuencias y el color blanco representa la ausencia de secuencias anotadas para esa función.

A.



B.



Anexo 8. Tablas de genes anotados dentro de los genomas virales parcialmente reconstruidos de tejidos de *A. germinans* (A.) y *R. mangle* (B.), que están involucrados en la relación microbiota-planta. En las tablas, se muestran los genes clasificados en las categorías del nivel 2 (Lv2) de la base de datos de PLaBase.

A.

Categoría Lv2	Genes Lv6	KEGG ID
BIO-FERTILIZATION	dmsD serD	K23349
COLONIZING_PLANT_SYSTEM	celJ eglA eglS	K01179
	yubE	NA
	murU	K00992
	xlyAB	K01447
	dcm	K00558
	murI	K01776
	nrdA nrdE	K00525
	nrdB nrdF	K00526
	dut	K01520
	rdgB	K02428
	xerD	K04763
	ata sadA emaA	K21449

COMPETITIVE_EXCLUSION CE	int intR intV intQ intW xerC	K14059
	cwlO	K21471
	lytH	K21472
	toxin_mqsR	K13651
	hsdM	K03427
	dam	K06223
	EC_2_1_1_72	K00571
	mcrA	K07451
	antitoxin_YdaS	NA
	REGULATION_higA	K18831
	antitoxin_mqsA	K13655
	toxin_yobL	K21487
	vgrG	K11904
PHYTOHORMONE PLANT_SIGNAL_PRODUCTION	thyX thyI	K03465
	coaE	K00859
STRESS_CONTROL BIOCONTROL	aroL aroK	K00891
	ohrR	K23775
	hdeB	K19778
	clpP	K01358
	lapA	K08992
	yozG	K07727
	sigM rpoE	K03090

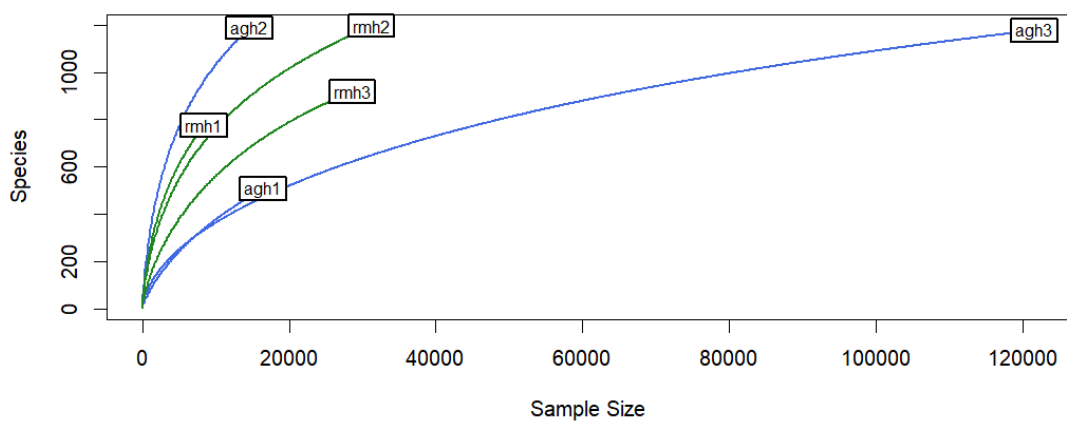
B.

Categoría Lv2	Genes Lv6	KEGG ID
BIO-FERTILIZATION	ntrX	K13599
	prhI fecI	K23514
	dmsD serD	K23349
COLONIZING_PLANT_SYSTEM	rrrD	K01185
	acm	K07273
	dcm	K00558
	nrdA nrdE	K00525
	comEB tadA	K01493
	ata sadA emaA	K21449
COMPETITIVE_EXCLUSION CE	lytH	K21472
	EC_3_4_17_14	K08640
	EC_2_1_1_72-	K00571
	mod	K07316
	mcrA	K07451
	toxin_darT	NA
	tpn	K07488
PHYTOHORMONE PLANT_SIGNAL_PRODUCTION	ribBA	K14652
	thyA	K00560
	panE apbA	K00077
	cobS	K09882

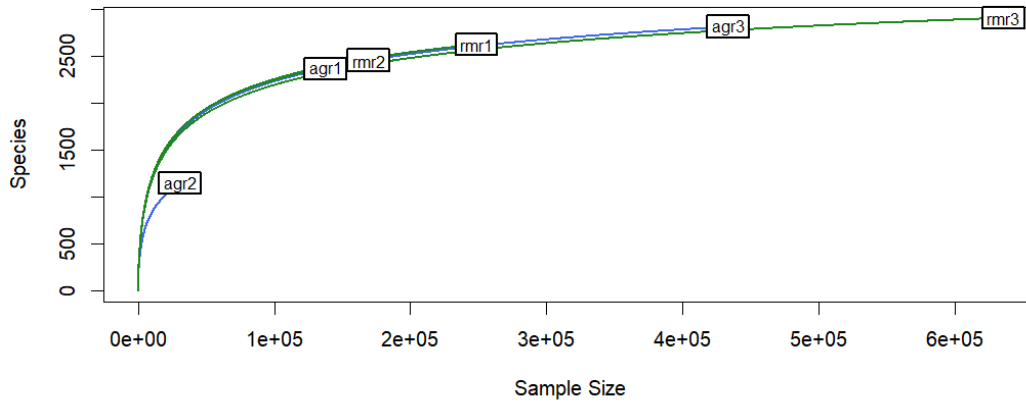
STRESS_CONTROL BIOCONTROL	putative_chitinase	K03791
	aroL aroK	K00891
	gacA uvrY varA	K07689
	yfkJ wzb	K01104
	dnaJ	K03686
	groES mopB	K04078
	clpP	K01358
	rpoE sigW	K03088
	csrA zfiA	K03563

Anexo 9. Curvas de rarefacción del conjunto de plásmidos para cada réplica. Se describen las curvas de cantidad de especies bacterianas hospederas de plásmidos, en función de la cantidad de secuencias. **A.** Tejido foliar de muestras de *R. mangle* (Rm) y *A. germinans* (Ag). **B.** Tejido radicular de *R. mangle* (Rm) y *A. germinans* (Ag).

A.



B.



Anexo 10. Análisis multivariado permutado de varianza (PERMANOVA) del conjunto de plásmidos analizados en librerías metagenómicas endófitas, con los factores especie, tejido e interacción entre especie y tejido. Este análisis contó con 999 permutaciones y con un valor α de referencia de 0.05.

Factors	Df	SumOfSqs	R2	F	Pr(>F)	
<i>Specie</i>	1	0,1787	0,05204	0,7994	0,472	
<i>Tissue</i>	1	1,269	0,36951	5,6762	0,004	**
<i>Specie:Tissue</i>	1	0,1981	0,05767	0,8859	0,398	
<i>Residual</i>	8	1,7885	0,52079			
<i>Total</i>	11	3,4343	1			

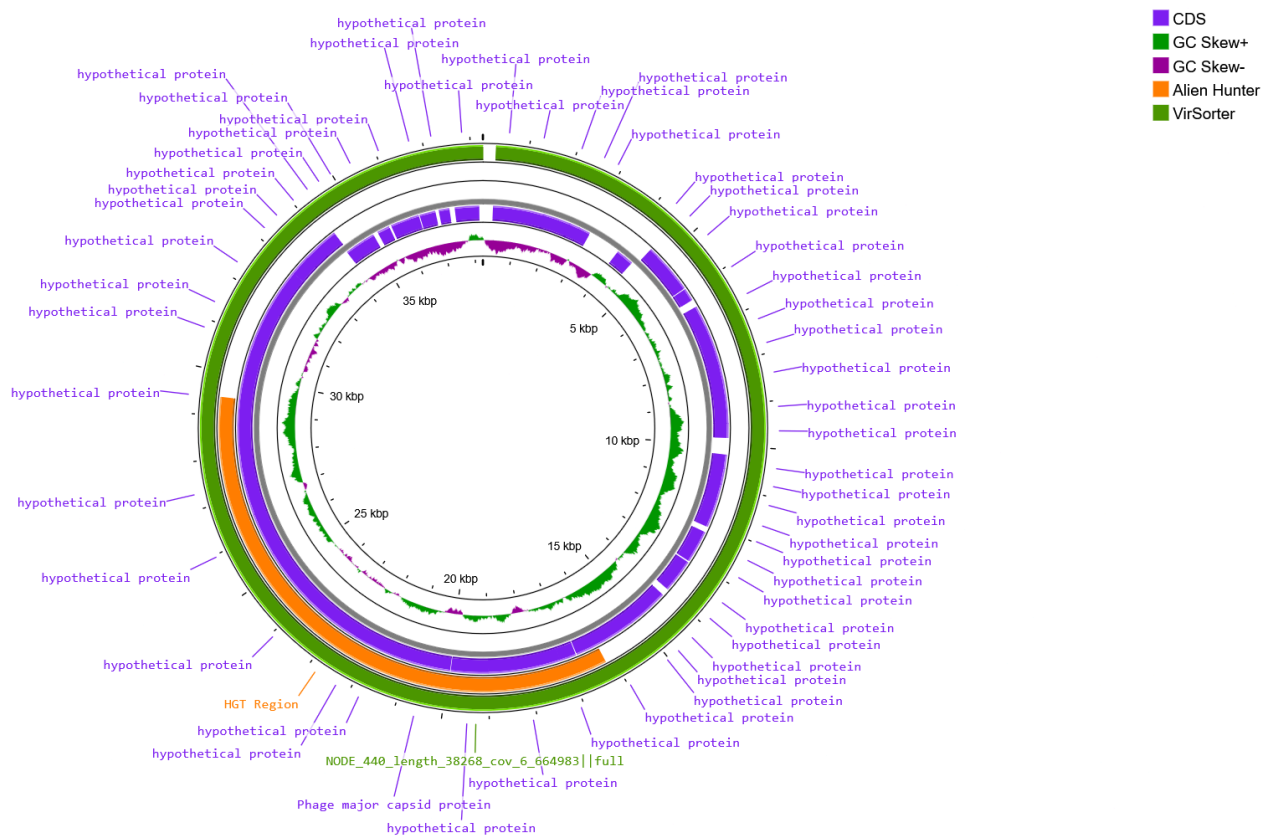
Anexo 11. Análisis estructural de los genomas virales reconstruidos, anotados y visualizados a través de Proksee, en las muestras provenientes de *A. germinans*.



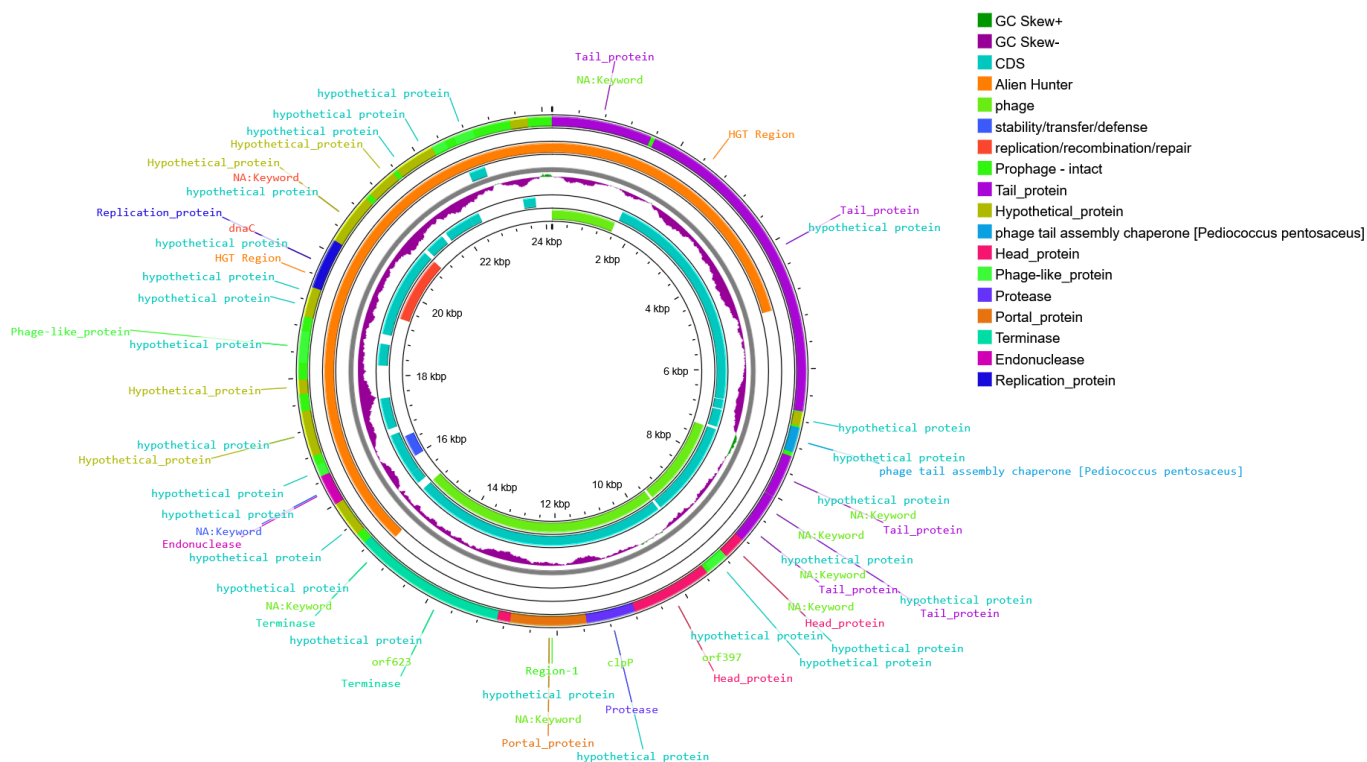




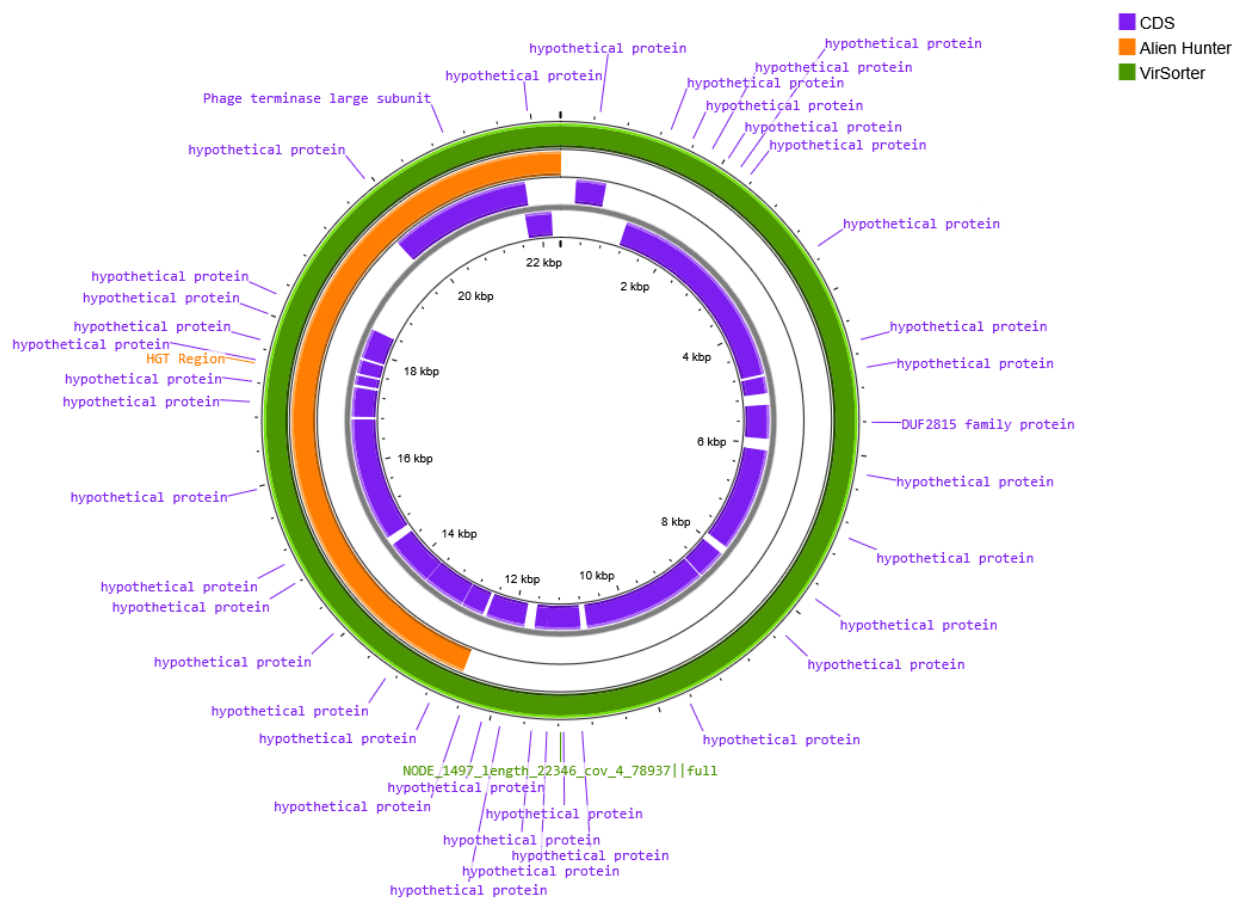
NODE 378 length 40626 cov 8.627936



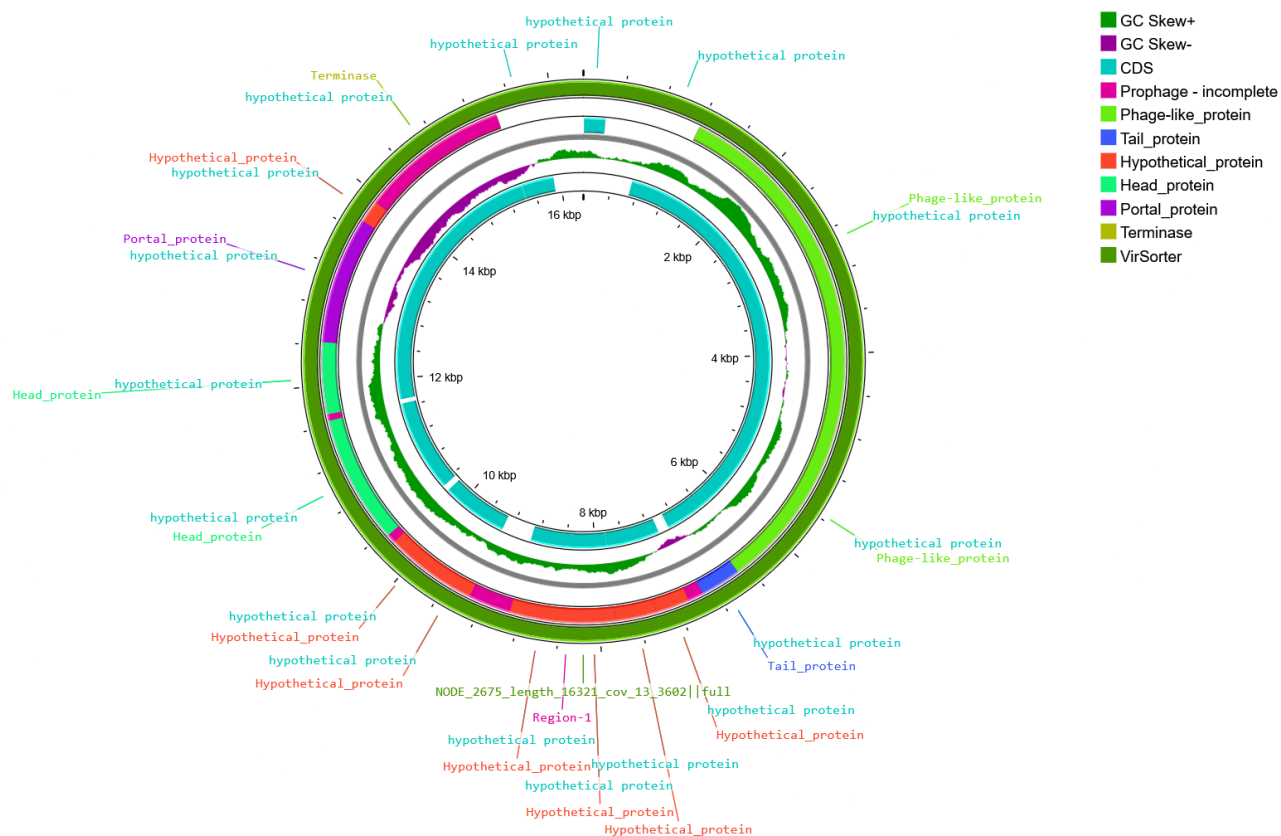
NODE_440_length_38268_cov_6.664983



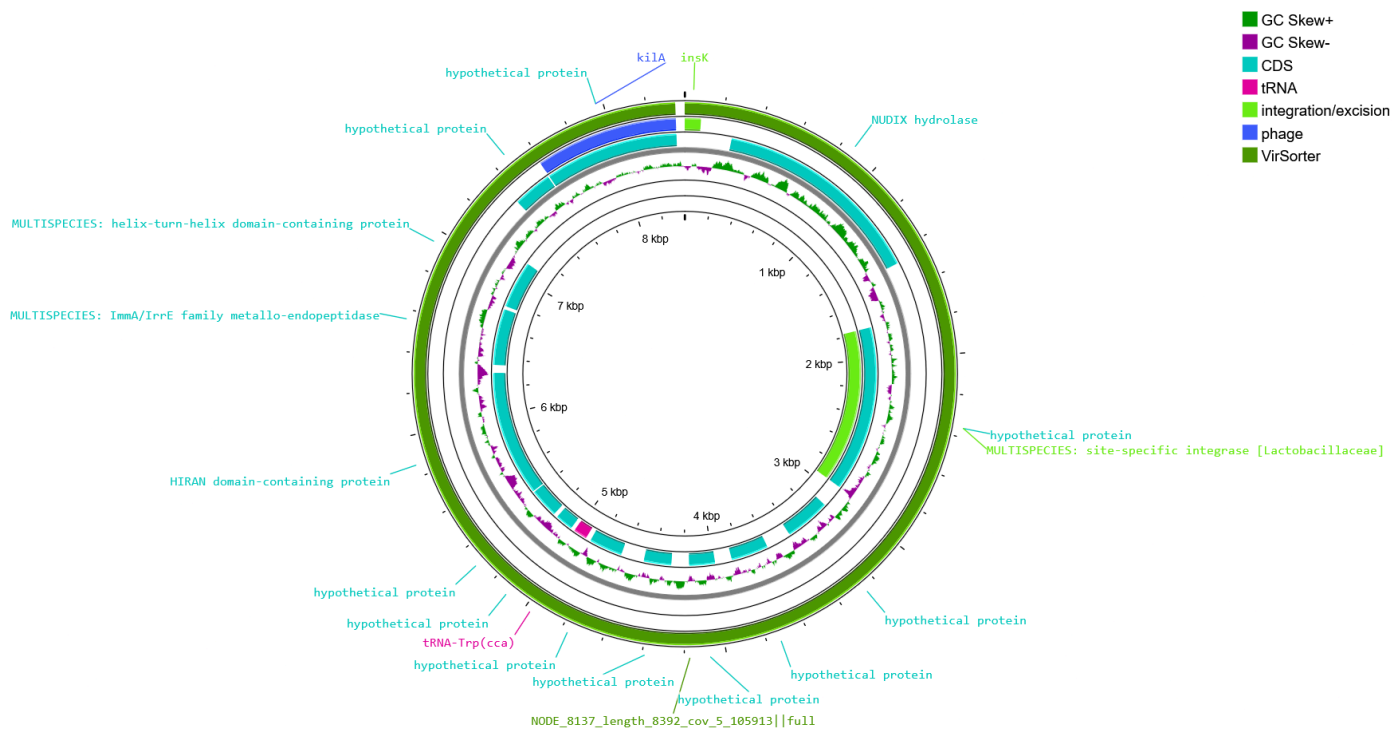
NODE_1281_length_24149_cov_5.527932



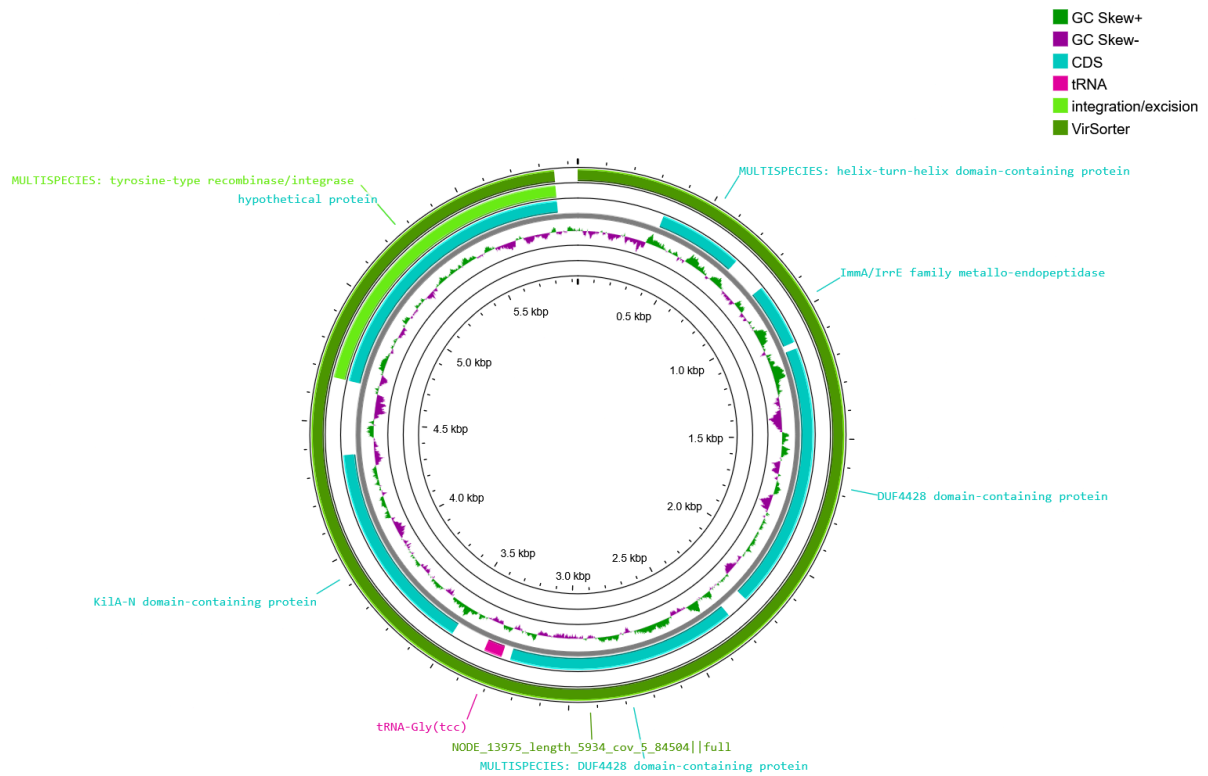
NODE_1497_length_22346_cov_4.789377



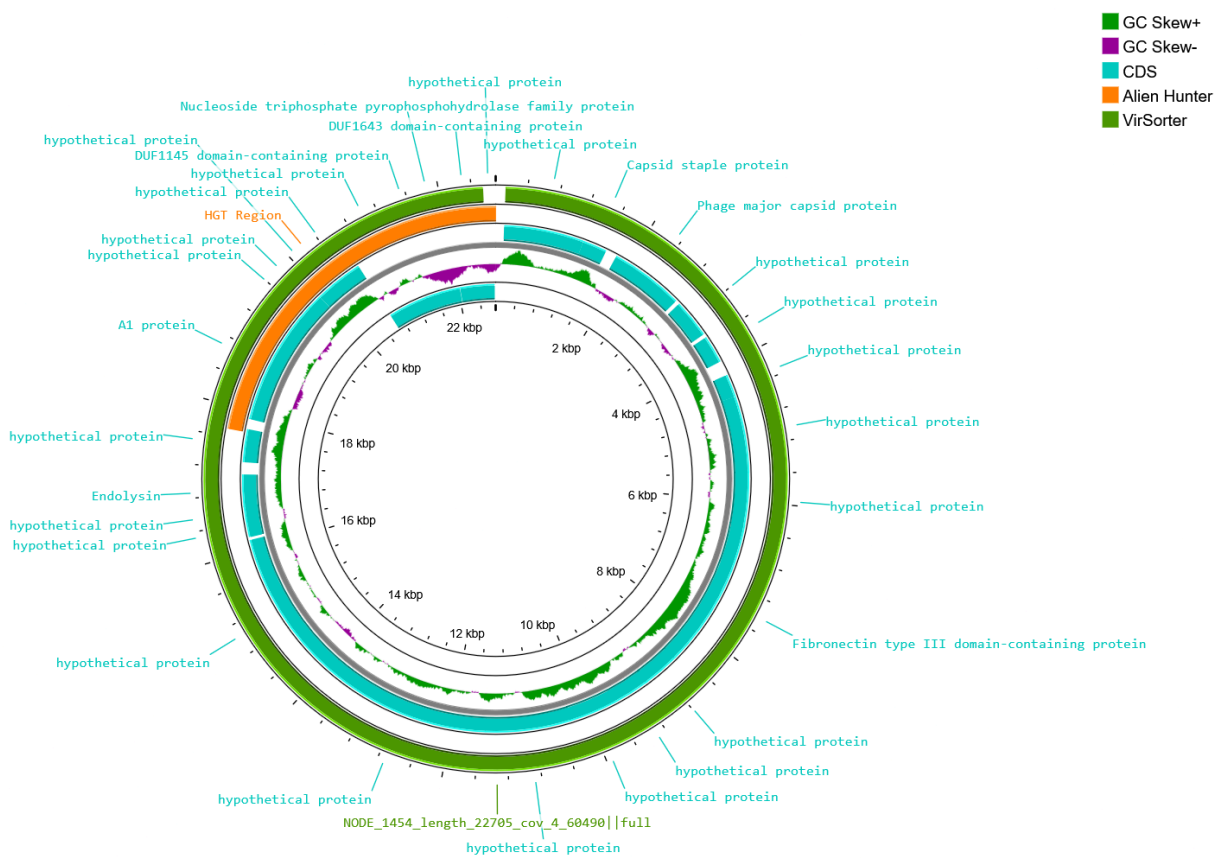
NODE_2675_length_16321_cov_13.360261



NODE_8137_length_8392_cov_5.105913

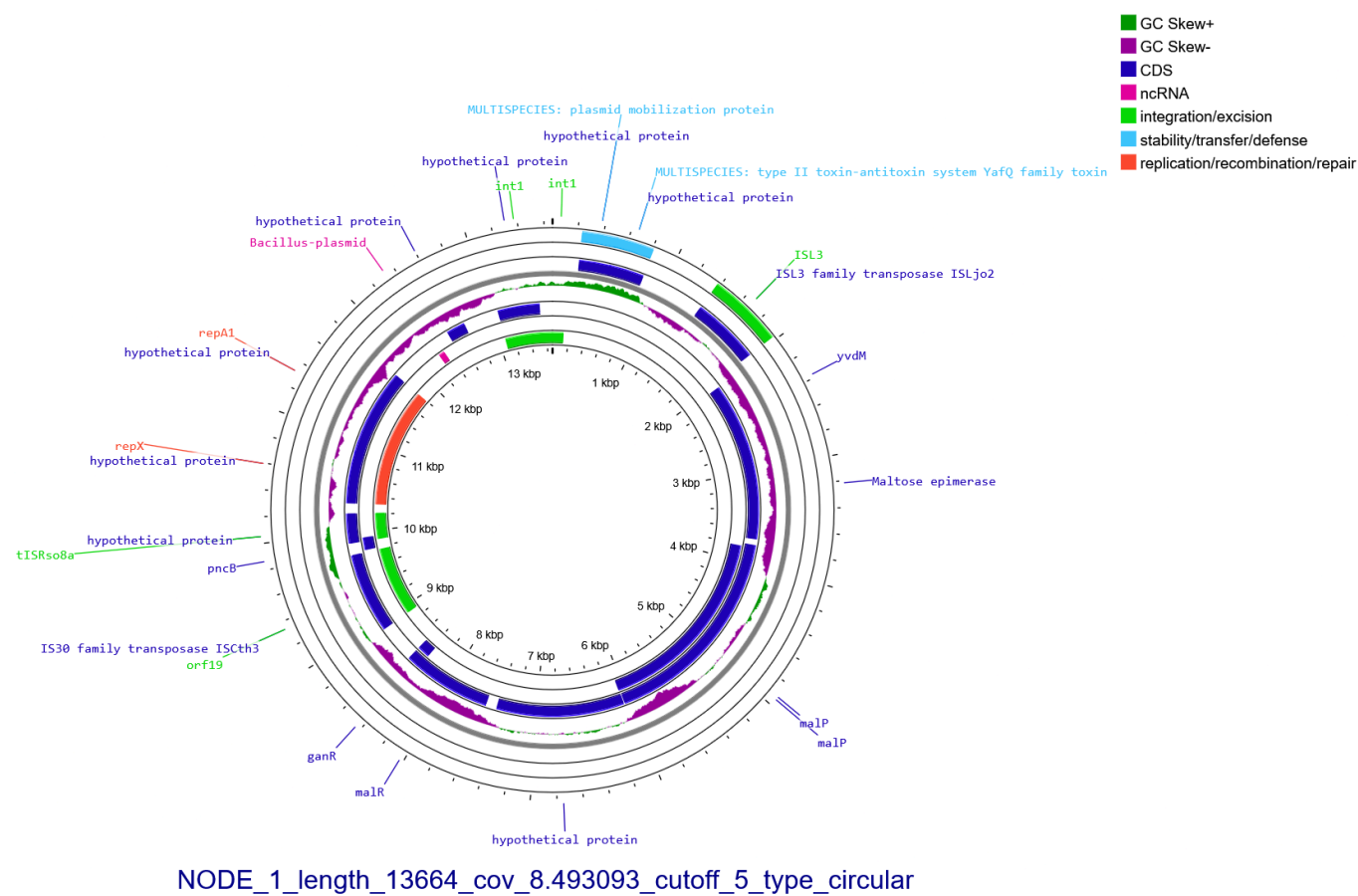


NODE_13975_length_5934_cov_5.845042

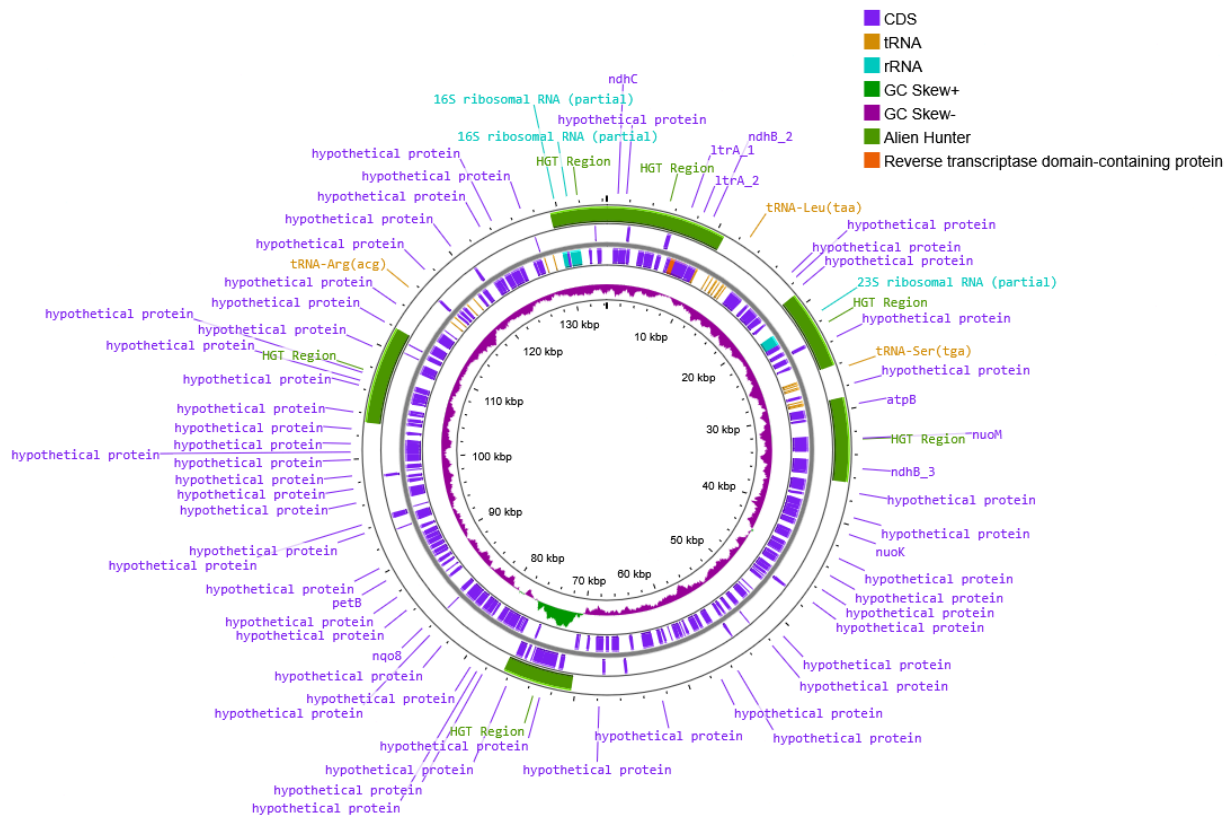


NODE_1454_length_22705_cov_4.604901

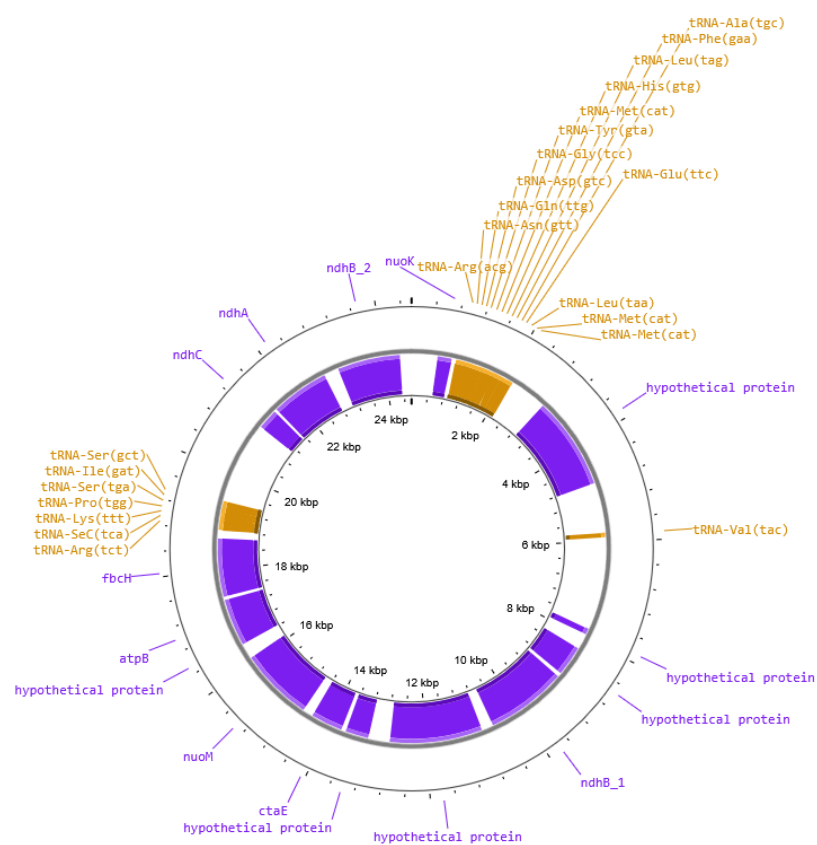
Anexo 13. Análisis estructural del plásmido reconstruido pZW5A de *Lactiplantibacillus plantarum*.



Anexo 14. Análisis estructural de los plásmidos parcialmente reconstruidos a partir del metagenoma radicular de *A. germinans*.

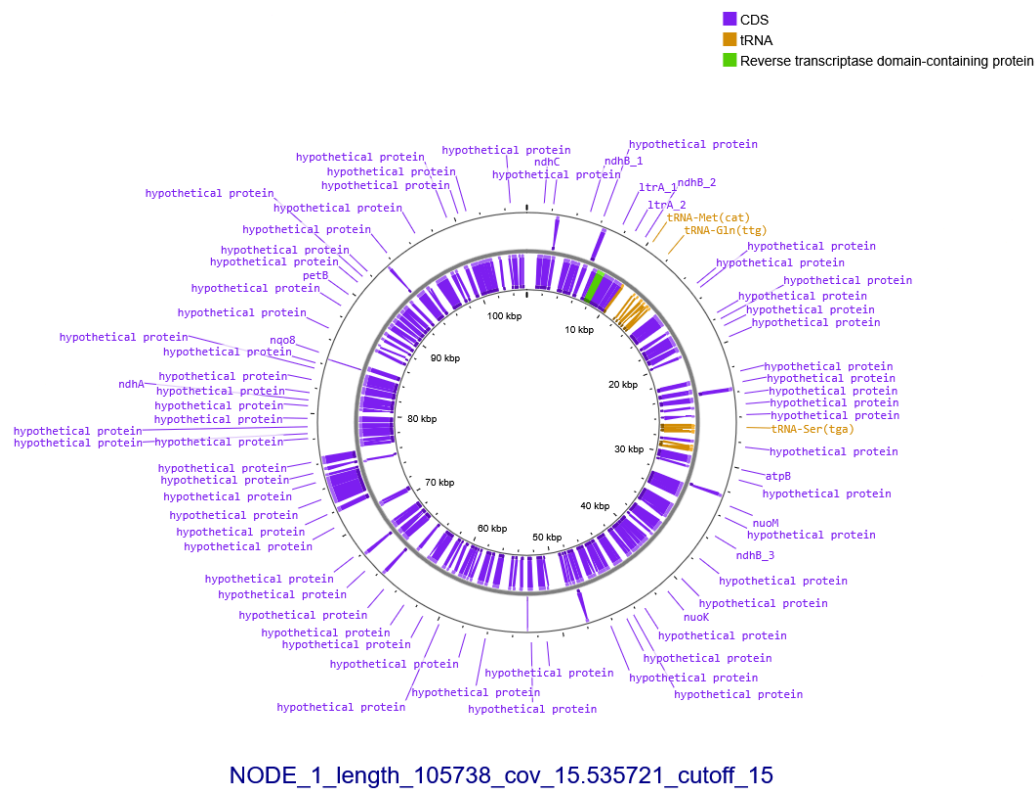


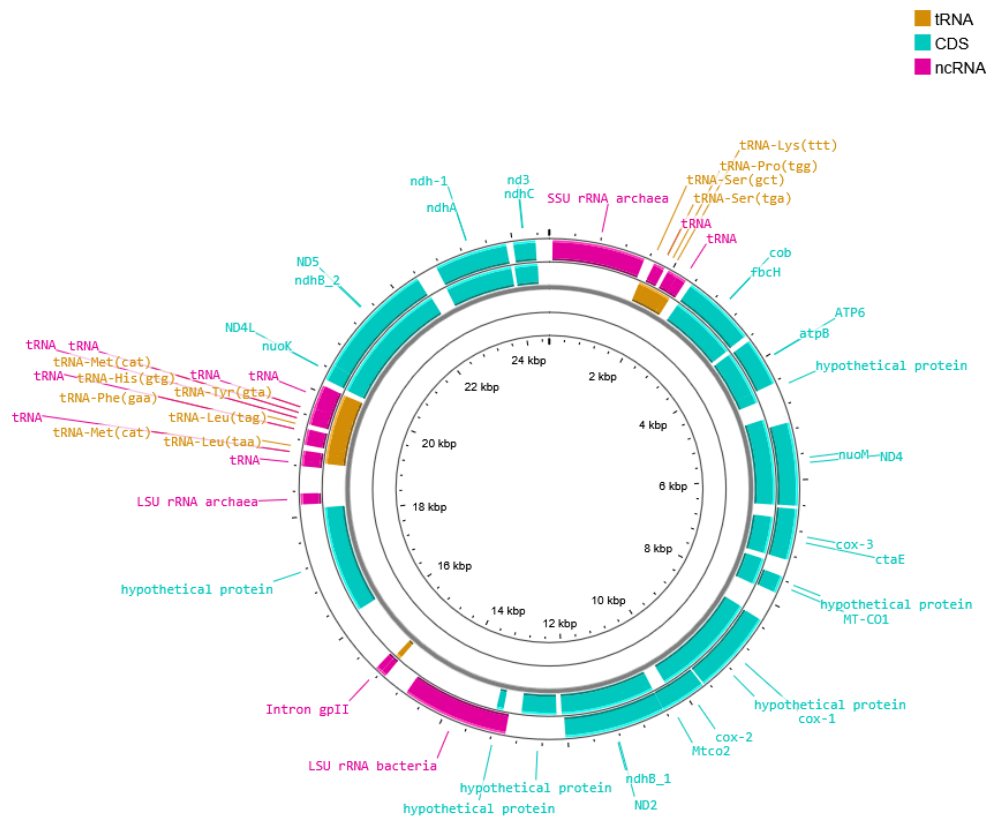
NODE_1_length_134371_cov_13.866080_cutoff_10



NODE_2_length_24580_cov_16.684333_cutoff_10

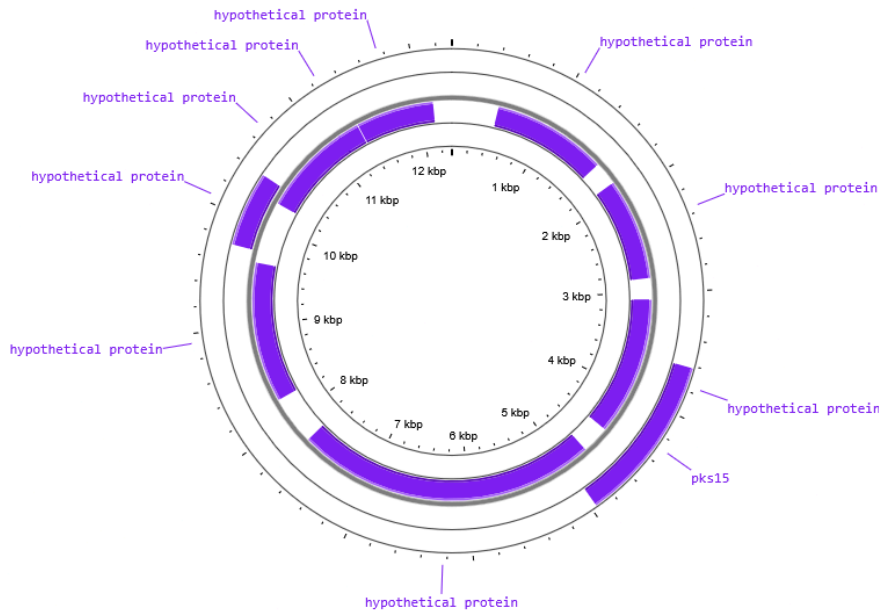
Anexo 15. Análisis estructural de los plásmidos parcialmente reconstruidos y ensamblados a partir del metagenoma radicular de *R. mangle*.





NODE_1_length_24581_cov_28.658870_cutoff_10

■ CDS



NODE_3_length_12325_cov_1.434252_cutoff_0

Anexo 16. Análisis de calidad con QUAST, de los ensamblajes empleados para la caracterización del viroma de *A. germinans* (A.) y *R. mangle* (B.).

A.

Assembly	AgH1	AgH2	AgH3	AgR1	AgR2	AgR3
# contigs (≥ 0 bp)	35744	40593	42420	858665	157173	1891896
# contigs (≥ 1000 bp)	484	483	632	19004	3084	75590
# contigs (≥ 5000 bp)	4	3	102	341	180	5674
# contigs (≥ 10000 bp)	0	0	80	49	109	1478
# contigs (≥ 25000 bp)	0	0	49	7	47	238
# contigs (≥ 50000 bp)	0	0	25	1	15	53
Total length (≥ 0 bp)	6420059	7333334	11255688	259599172	48218750	685680229
Total length (≥ 1000 bp)	809334	793563	4850948	32927894	8620460	181650786
Total length (≥ 5000 bp)	22082	17249	3997962	2727082	3731793	55427961

Total length (≥ 10000 bp)	0	0	3842956	845174	3262112	27140285
Total length (≥ 25000 bp)	0	0	3359687	246089	2261364	9890308
Total length (≥ 50000 bp)	0	0	2476826	54805	1127407	3601578
# contigs	1867	1728	2282	82042	13602	232825
Largest contig	6077	6209	262083	54805	122159	110360
Total length	1749141	1639441	5948704	75054644	15405245	287981140
GC (%)	35.58	35.63	40.71	55.93	51.89	56.17
N50	941	974	33687	902	1209	1417
N90	561	562	679	552	551	588
auN	1324.9	1320.2	57739.1	1579.7	10950.4	4486.1
L50	552	510	37	23858	2246	43604
L90	1537	1419	1256	67730	10660	179571
# N's per 100 kbp	0	0	0	0	0	0

B.

Assembly	RmH1	RmH2	RmH3	RmR1	RmR2
# contigs (≥ 0 bp)	17286	12214	31360	1928094	1182878
# contigs (≥ 1000 bp)	202	322	732	105636	35258
# contigs (≥ 5000 bp)	3	93	12	3464	408
# contigs (≥ 10000 bp)	0	80	0	436	54
# contigs (≥ 25000 bp)	0	46	0	48	4
# contigs (≥ 50000 bp)	0	21	0	18	0
Total length (≥ 0 bp)	3220153	5947353	6348541	695553807	383124252
Total length (≥ 1000 bp)	322568	3621562	1288195	201311348	55806948
Total length (≥ 5000 bp)	21406	3268541	77528	27157796	3131305
Total length (≥ 10000 bp)	0	3176532	0	7333194	769031
Total length (≥ 25000 bp)	0	2549074	0	2281611	119227
Total length (≥ 50000 bp)	0	1580163	0	1257730	0
# contigs	722	944	2164	307318	145133
Largest contig	8590	178123	9464	95065	42803
Total length	673457	4044786	2274342	340262456	130174016

GC (%)	31.35	41.89	35.08	40.43	42.77
N50	973	42083	1112	1189	899
N90	558	968	584	590	554
auN	1382.6	46671.5	1619	2319.8	1328.2
L50	217	31	588	76960	45056
L90	595	342	1744	244567	120377
# N's per 100 kbp	0	0	0	0	0

Anexo 17. Análisis de calidad con QUAST, de los ensamblajes de metaplasmidSPADES para la caracterización del plasmidoma de *A. germinans* y *R. mangle* por cada tipo de tejido.

Assembly	AgH-plasmidome	AgR-plasmidome	RmH-plasmidome	RmR-plasmidome
# contigs (≥ 0 bp)	5	68	3	155
# contigs (≥ 1000 bp)	5	68	3	155
# contigs (≥ 5000 bp)	0	16	1	27
# contigs (≥ 10000 bp)	0	6	1	9
# contigs (≥ 25000 bp)	0	2	0	5
# contigs (≥ 50000 bp)	0	1	0	3
Total length (≥ 0 bp)	9202	425474	16367	856064
Total length (≥ 1000 bp)	9202	425474	16367	856064
Total length (≥ 5000 bp)	0	296921	13664	498956
Total length (≥ 10000 bp)	0	235385	13664	384319
Total length (≥ 25000 bp)	0	176173	0	326323
Total length (≥ 50000 bp)	0	134371	0	243186
# contigs	5	68	3	155
Largest contig	2853	134371	13664	105738
Total length	9202	425474	16367	856064
GC (%)	31.84	38.39	39.54	43.45
N50	1827	13395	13664	7203
N90	1155	2188	1549	2529
auN	2008.3	50733.7	11635.4	32195.4
L50	2	4	1	15

L90	5	41	2	107
# N's per 100 kbp	0	0	0	0

Anexo 18. Análisis de homogeneidad de varianzas (PERMDISP), de la comunidad viral analizada en librerías metagenómicas endófitas, con los factores especie (A.), tejido (B.) e interacción entre especie y tejido (C.). Este análisis contó con un valor α de referencia de 0.05.

A.

Factors	Df	SumOfSqs	MeanSq	F	Pr(>F)
<i>Especie</i>	1	0,01774	0,017740	0,6022	0,4557
<i>Residual</i>	10	0,29458	0,029458		

B.

Factors	Df	SumOfSqs	MeanSq	F	Pr(>F)
<i>Tissue</i>	1	0,00002	0,000024	6e-04	0,9812
<i>Residual</i>	10	0,40494	0,040494		

C.

Factors	Df	SumOfSqs	MeanSq	F	Pr(>F)
<i>Especie:Tissue</i>	3	0,00505	0,001684	0,0256	0,994
<i>Residual</i>	8	0,52652	0,065816		

Anexo 19. Análisis de homogeneidad de varianzas (PERMDISP), del conjunto de plásmidos analizados en librerías metagenómicas endófitas, con los factores especie (A.), tejido (B.) e interacción entre especie y tejido (C.). Este análisis contó con un valor α de referencia de 0.05.

A.

Factors	Df	SumOfSqs	MeanSq	F	Pr(>F)
<i>Especie</i>	1	0,000004	0,000004	2e-04	0,9879
<i>Residual</i>	10	0,166976	0,016698		

B.

Factors	Df	SumOfSqs	MeanSq	F	Pr(>F)
<i>Tissue</i>	1	0,04735	0,047348	1,3623	0,2702
<i>Residual</i>	10	0,34756	0,034756		

C.

Factors	Df	SumOfSqs	MeanSq	F	Pr(>F)
<i>Specie:Tissue</i>	3	0,083884	0,027961	0,8071	0,5245
<i>Residual</i>	8	0,277167	0,034646		