



FACULTAD DE CIENCIAS AGRARIAS

FACULTAD DE CIENCIAS BIOQUÍMICAS Y FARMACÉUTICAS

UNIVERSIDAD NACIONAL DE ROSARIO

Secuenciación genómica y caracterización de una cepa nativa de *Photorhabdus* sp. P32.

Dr. Jorge A. Zimmermann

**TRABAJO FINAL PARA OPTAR AL TÍTULO DE ESPECIALISTA EN
BIOINFORMÁTICA**

DIRECTOR: Dr. Leopoldo Palma Dovis

CODIRECTOR: Dr. Lucas Daurelio

2025

Secuenciación genómica y caracterización de una cepa nativa de *Photorhabdus* sp.

P32

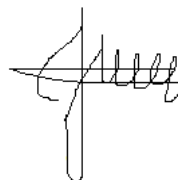
Jorge A. Zimmermann

Dr. en Cs. Veterinarias, Médico Veterinario – Universidad Nacional del Litoral

Este Trabajo Final es presentado como parte de los requisitos para optar al grado académico de Especialista en Bioinformática, de la Universidad Nacional de Rosario y no ha sido previamente presentada para la obtención de otro título en esta u otra Universidad. El mismo contiene los resultados obtenidos en investigaciones llevadas a cabo en el Instituto de Ciencias Veterinarias del Litoral (CONICET – UNL), durante el período comprendido entre julio 2022 – marzo 2024, bajo la dirección del Dr. Leopoldo Palma Dovis y la codirección del Dr. Lucas Daurelio.

Nombre y firma del autor:

Zimmermann Jorge A.



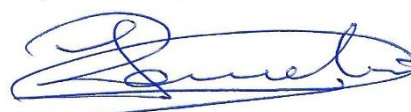
Nombre y firma del Director

Leopoldo Palma Dovis



Nombre y firma del Co - Director

Lucas Daurelio



Defendida: 19 de marzo de 2025.

Presentaciones en congresos

2022 - Genome characterization of the novel *Photorhabdus* sp. strain p32 isolated in Argentina from *Galleria mellonella* larvae.

Jorge Zimmermann, Leopoldo Palma, Federico Acosta, Laureano Frizzo, Eleodoro Eduardo del Valle.

International Congress on Invertebrate Pathology and Microbial Control and 54th Annual Meeting of the Society for Invertebrate Pathology. Nelson Mandela Bay, South Africa. 1 - 4 August 2022

2024 - Primera descripción de *Photorhabdus aballayi* en Argentina: exploración de su genoma y análisis de su potencial insecticida.

Jorge Zimmermann, Leopoldo Palma, Federico Acosta, Laureano Frizzo, Eleodoro, Sauka Diego, Eduardo del Valle.

XVI Congreso Argentino de Microbiología 2024. Ciudad Autónoma de Buenos Aires, Argentina. 21 – 23 de agosto 2024.

Abreviaturas y símbolos

A: adenina

ADN: ácido desoxirribonucleico

ARN: ácido ribonucleico

BLAST: del inglés "Basic Local Alignment Search Tool"

BPPRC: del inglés "Bacterial Pesticidal Protein Resource Center"

C: citosina

CDSs: del inglés "protein coding sequences", secuencias codificantes de proteínas

CRISPR: del inglés "Clustered Regularly Interspaced Short Palindromic Repeats", repeticiones palindrómicas cortas agrupadas y regularmente espaciadas

dDDH: hibridación digital ADN-ADN

EPN: nematodos entomopatógenos

G: guanina

g: unidad de fuerza centrífuga relativa

G + C: guanina + citosina

GBDP: del inglés "Genome BLAST Distance Phylogeny"

Gbp: del inglés "gigabase pairs", pares de gigabases

IC: intervalo de confianza

ID: identificación

INTA: Instituto nacional de tecnología agropecuaria

Jl: juveniles infectivos

LB: Luria-Bertani, medio de cultivo

Mb: megabases

N: nucleótido no asignado

NBTA: del inglés “Nutrient blue tetrasodium agar”, agar tetrasodio nutriente azul

NCBI: del inglés “National Center for Biotechnology Information”

ncRNAs: del inglés “non codificant ribonucleic acid”, ARN no codificante

Ni-siderophore: del inglés “NRPS-independent siderophore”, sideróforo independiente de NRPS

NRPS: del inglés “nonribosomal peptide synthetase”, péptido sintetasa no ribosomal

pb: pares de bases

RNAs: del inglés “ribonucleic acids”, ARN

sp: especie

subsp: subespecie

T: timina

tRNAs: del inglés “transfer ribonucleic acid”, ARN de transferencia

Glosario

N25: tamaño del contig que, junto a los contigs más grandes, contiene al 25% de la secuencia del genoma.

N50: tamaño del contig que, junto con los contigs más grandes, contiene al 50% de la secuencia del genoma.

N75: tamaño del contig que, junto a los contigs más grandes, contiene al 75% de la secuencia del genoma.

Cluster: grupo de genes o secuencias genéticas que están cercanas entre sí en un genoma y que comparten similitudes estructurales o funcionales.

Contig: conjunto de secuencias o segmentos ensamblados de ADN que representan regiones consenso de mayor tamaño.

Gap: falta de una o más bases en una secuencia de ADN. Un gap puede corresponder tanto a una delección como a una inserción.

Genoma core: conjunto de genes o secuencias genéticas presentes en todas las cepas o individuos de una especie o grupo taxonómico dado.

OrthoANI: valor, identidad promedio de nucleótidos por ortólogo. Los genes ortólogos son genes de diferentes especies que se originaron por el descenso vertical de un solo gen del último ancestro común.

Pan genoma: conjunto completo de todos los genes encontrados en las cepas o individuos de una especie o grupo taxonómico dado. Incluye no solo los genes compartidos (genoma core), sino también los genes accesorios o variables, que pueden estar presentes solo en ciertas cepas o individuos de la población.

Phred33: codificación utilizada por Illumina para las puntuaciones de calidad. En esta codificación, el nivel de calidad Q se representa como un carácter ASCII sumando 33 al valor numérico de Q. Por ejemplo, un nivel de calidad de 30 se representaría mediante el carácter ASCII '6' (33 + 30).

Pseudogen: secuencia de ADN que se asemeja estrechamente a un gen funcional pero que ha perdido su función original debido a la acumulación de mutaciones (ej. codon de stop prematuro).

Quality Score: es una medida para evaluar la confiabilidad de cada base o nucleótido en una secuencia de ADN o ARN determinada. Cada base secuenciada tiene asociada una puntuación de calidad que indica la probabilidad de que la base sea incorrecta. Estas puntuaciones se expresan comúnmente como valores numéricos o como caracteres ASCII específicos.

Scaffold: regiones consenso de unión de contigs con gaps.

Singletons: genes que están presentes en una sola cepa o individuo dentro de un conjunto de organismos estudiados. Estos genes son únicos en el sentido de que no se encuentran en ninguna otra cepa o individuo de la población analizada.

Trimming: eliminación o recorte de bases no deseadas o de baja calidad de las secuencias de ADN o ARN.

Resumen

Los nematodos del género *Heterorhabditis* son parásitos de insectos y han demostrado poseer potencial para ser utilizados como agentes de control biológico. Estos nematodos penetran a través de las aberturas naturales de los insectos o su cutícula y una vez en el interior, liberan bacterias simbióticas entomopatógenas en el hemocele. El género *Heterorhabditis* está asociado con enterobacterias del género *Photorhabdus*. Estas bacterias liberan toxinas insecticidas que producen la muerte de los insectos. Además, pueden sintetizar otros metabolitos secundarios con potencial biotecnológico y farmacéutico, destinadas a inhibir la proliferación de otros microorganismos competidores y oportunistas. El objetivo de este trabajo fue secuenciar el genoma de la cepa P32 de *Photorhabdus* sp., aislada de larvas de *Galleria mellonella* infestadas con *Heterorhabditis amazonensis* y realizar una caracterización de la misma mediante la utilización de diferentes herramientas bioinformáticas. La bacteria fue aislada a partir de muestras de suelo obtenidas en Calilegua, provincia de Jujuy, Argentina. El tamaño del borrador del genoma de *Photorhabdus* P32 fue de 5.36 Megabases con un contenido G+C de 42,5%. Se identificaron 4614 CDSs, de los cuales 141 corresponden a funciones desconocidas. El análisis de identidad promedio de nucleótidos (orthoANI) mostró un 96,94 % de similitud entre *Photorhabdus* P32 y *Photorhabdus aballayi*, confirmando que esta cepa pertenece a dicha especie. A través de este estudio, pudimos caracterizar el genoma de *Photorhabdus* P32, en el que identificamos proteínas con potencial actividad insecticida, quitinasas y regiones de clusters de genes de metabolitos secundarios con actividad citotóxica y de sideróforo. Finalmente, el análisis filogenético identificó a *Photorhabdus* P32 como una cepa perteneciente a la especie *P. aballayi*

Palabras claves: *Photorhabdus*, enterobacteria, insecticida, genoma, *Heterorhabditis*

Genomic sequencing and characterization of a native *Photorhabdus* sp. strain

Abstract

Nematodes of the *Heterorhabditis* genus are insect parasites and have demonstrated the potential to be used as biological control agents. These nematodes penetrate through the natural openings of insects or their cuticle and once inside, they release symbiotic entomopathogenic bacteria into the hemocoel. The genus *Heterorhabditis* is associated with enterobacteria of the *Photorhabdus* genus. These bacteria release insecticidal toxins that kill the insects. In addition, they can synthesize other secondary metabolites with biotechnological and pharmaceutical potential, intended to inhibit the proliferation of other competing and opportunistic microorganisms. The objective of this work was to sequence the genome of the strain P32 of *Photorhabdus* sp., isolated from *Galleria mellonella* larvae infested with *Heterorhabditis amazonensis* and to characterize it using different bioinformatics tools. The bacteria was isolated from soil samples obtained in Calilegua, province of Jujuy, Argentina. The size of the draft genome of *Photorhabdus* P32 was 5.36 Megabases with a G + C content of 42.5%. A total of 4614 CDSs were identified, 141 of which correspond to unknown functions. The average nucleotide identity (orthoANI) analysis showed a 96.94% similarity between *Photorhabdus* P32 and *Photorhabdus aballayi*, confirming that this strain belongs to that species. Through this study, we were able to characterize the genome of *Photorhabdus* P32, in which we identified proteins with potential insecticidal activity, chitinases and regions of secondary metabolite gene clusters with cytotoxic and siderophore activity. Finally, phylogenetic analysis identified *Photorhabdus* P32 as a strain belonging to the species *P. aballayi*.

Índice de tablas

Tabla 1. Resumen del análisis de las secuencias forward y reverse en FASTQC de la secuenciación del genoma <i>Photorhabdus</i> P32	Pág. 20
Tabla 2. Mediciones de contigs del ensamblaje “de novo” en CLC Genomics Workbench del genoma (incluyendo regiones scaffolded) y luego del mapeo a la referencia para la optimización del ensamblaje	Pág. 22
Tabla 3. Distribución de nucleótidos con scaffolding del ensamblaje “de novo”	Pág. 22
Tabla 4. Proteínas potencialmente insecticidas identificadas con BLAST personalizado (custom) en el genoma de <i>Photorhabdus</i> sp. P32 con especies objetivo	Pág. 25
Tabla 5. Comparaciones por pares de genomas de cepa <i>Photorhabdus</i> P32 versus genomas de cepa tipo	Pág. 31
Tabla 6. Genes en Islas genómicas de <i>Photorhabdus</i> P32 contra <i>Photorhabdus luminescens</i> subsp. <i>laumondii</i> TTO1.	Pág. 33

Índice de figuras

Figura 1. Análisis de calidad de secuencias en FASTQC	Pág. 21
Figura 2. Clasificación funcional de genes en la cepa P32 de <i>Photorhabdus</i>	Pág. 24
Figura 3. Clústeres de genes potencialmente productores de metabolitos secundarios detectados con antiSMASH en el genoma de <i>Photorhabdus</i> P32 con 100% de similitud	Pág. 27
Figura 4. Mapa de calor de <i>Photorhabdus</i> sp. P32 con especies estrechamente relacionadas generado con OrthoANI, valores calculados a partir del software OAT	Pág. 29
Figura 5. Árbol filogenético inferido con FastME 2.1.6.1 a partir de distancias GBDP calculadas a partir de secuencias de genomas. Las longitudes de las ramas se escalan en términos de la fórmula de distancia GBDP d_5 . Los números sobre las ramas son valores de soporte de pseudo-bootstrap de GBDP > 60 % de 100 replicas, con un soporte de rama promedio del 93,4 %. El árbol tenía sus raíces en el punto medio	Pág. 32
Figura 6. Islas genómicas de <i>Photorhabdus</i> P32 contra <i>Photorhabdus luminescens</i> subsp. <i>laumondii</i> TTO1. El círculo exterior verde representa el genoma en megabases de la cepa P32.	Pág. 34
Figura 7. Diagrama de Venn de <i>Photorhabdus</i> P32 (4), <i>Photorhabdus aballayi</i> , <i>Photorhabdus bodei</i> , <i>Photorhabdus laumondii</i> subsp. <i>laumondii</i> TTO1.	Pág. 35

Índice

Presentaciones en congresos	Pág. 2
Abreviaturas y símbolos	Pág. 3
Glosario	Pág. 5
Resumen	Pág. 7
Abstract	Pág. 8
Índice de tablas	Pág. 9
Índice de figuras	Pág. 10
Índice	Pág. 11
Introducción	Pág. 13
Hipótesis, objetivo general y objetivos específicos	Pág. 15
Materiales y métodos	
1. Purificación de ADN y secuenciación	Pág. 16
2. Control de calidad de lecturas, ensamblado del genoma y optimización	Pág. 16
3. Anotación del genoma	Pág. 16
4. Construcción de bases de datos personalizadas o custom en BLAST (Basic Local Alignment Search Tool) y búsqueda de factores de virulencia por custom BLAST.	Pág. 17
5. Búsqueda de clústers de genes potencialmente productores de metabolitos secundarios con antiSMASH	Pág. 18
6. Análisis filogenéticos	Pág. 18
Resultados y discusión	Pág. 20
1. Secuenciación, control de calidad de lecturas, ensamblado del genoma y optimización	Pág. 20
2. Anotación del genoma	Pág. 23
3. Construcción de bases de datos personalizadas en BLAST y búsqueda de factores de virulencia	Pág. 24
4. Búsqueda de clústers de genes potencialmente productores de metabolitos secundarios con antiSMASH	Pág. 25
5. Análisis filogenéticos	Pág. 28

Análisis de genes ortólogos	Pág. 28
Análisis taxonómico	Pág. 29
Identificación de islas genómicas	Pág. 32
Comparación de contenido genético	Pág. 34
Conclusiones y próximos trabajos con la cepa P32	Pág. 36
Bibliografía	Pág. 37

Introducción

Las prácticas agrícolas modernas no logran controlar de manera segura y completa a las plagas de insectos y patógenos. Existe además, una creciente preocupación por los productos químicos que se utilizan para el control de plagas en las plantas, debido a los efectos perjudiciales en la salud humana, la contaminación, toxicidad ambiental y el desarrollo de resistencias (Coupland *et al.*, 2017). Debido a estas problemáticas y la necesidad de realizar prácticas agrícolas sostenibles, es necesario reducir el uso de pesticidas no seguros y sustituirlos por productos no tóxicos y ecológicos en el control de plagas en cultivos.

En los últimos años se ha puesto énfasis en el uso de bacterias benéficas en programas de manejo integrado de plagas y patógenos. Las bacterias del género *Photorhabdus*, pueden ser una alternativa favorable para el control biológico de muchas plagas y patógenos en plantas (Abd-Elgawad 2021). Estas bacterias Gram-negativas pertenecen a la clase Gammaproteobacterias, familia Enterobacteriaceae y son capaces de sintetizar diversos metabolitos secundarios bioactivos (como fenazinas, toxinas proteicas, peptidolipidos, espectinabilinas) (Orozco *et al.*, 2016; Tobias *et al.*, 2016; Parihar *et al.*, 2022). Estas bacterias del género *Photorhabdus*, están asociadas simbióticamente con nematodos entomopatógenos (EPN) pertenecientes al género *Heterorhabditis* (Nematoda: Heterorhabditidae), los cuales parasitan insectos por lo que son utilizados como agentes de control biológico (como larvas de insectos, gusanos, coleópteros) (Hussaini 2014) para la protección de cultivos de importancia agrícola (Grewal *et al.*, 2005; Lacey y Georgis, 2012). Los nematodos buscan e infectan las larvas de insectos que viven en el suelo. Los juveniles infectivos (JI) al localizar al huésped, penetran a través de las aberturas naturales de los insectos (boca, ano o espiráculos) o a través de la cutícula (Koppenhöfer *et al.*, 2007). Una vez en el hemocele, los JI liberan bacterias simbióticas presentes en su tracto digestivo mediante regurgitación. Estas bacterias proliferan rápidamente y liberan una serie de factores de virulencia, entre ellos, toxinas insecticidas que producen la muerte de los insectos mediante una combinación de toxemia generalizada y septicemia (Jurat-Fuentes & Jackson, 2012). Además, pueden sintetizar otros péptidos con potencial biotecnológico y farmacéutico (como enzimas y sustancias antimicrobianas) destinadas a inhibir la proliferación de otros

microorganismos competidores y oportunistas (Duchaud *et al.*, 2003). Los insectos infectados suelen morir rápidamente unas 48 h después de la infección (Forst y Clarke, 2002). Luego, los nematodos se alimentan de las bacterias y los tejidos digeridos, lo que permite a los JI recuperar a su bacteria simbiote en el tracto digestivo y completar de una a tres generaciones dentro del huésped. Finalmente, los JI abandonan el cadáver del hospedador y buscan un nuevo hospedador en el suelo (Hazir *et al.*, 2003).

De esta manera, los nematodos proporcionan a las bacterias un medio vital actuando como vectores, mientras que éstas suministran una diversidad de metabolitos tales como sustancias inmunosupresoras, bacteriocinas, sideróforos, enzimas y toxinas insecticidas necesarias para matar al insecto y producir el desarrollo óptimo y reproducción del nematodo en el hospedador (Machado *et al.*, 2021). La mayoría de los metabolitos producidos por esta bacteria presentan interesantes aplicaciones con potencial biotecnológico (Abd-Elgawad 2021).

El primer genoma de *Photorhabdus* spp. fue secuenciado en el año 2003 (Duchaud *et al.*, 2003) y actualmente hay 181 secuencias del genoma de varias cepas de *Photorhabdus* spp. disponibles en la base de datos de NCBI (<https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=29487>) (14/03/2025). Los genes que codifican proteínas insecticidas producidas por bacterias del género *Photorhabdus*, pueden convertirse en nuevos transgenes para la construcción de cultivos innovadores resistentes a insectos (Sheets y Aktories, 2016). La identificación de nuevas especies y cepas de *Photorhabdus*, no solo aumentará la diversidad de especies ya conocidas, sino que también podría contribuir a la disponibilidad de nuevas proteínas insecticidas y otros productos naturales con un interesante potencial biotecnológico (Hurst *et al.*, 2015). En este contexto, la cepa P32 de *Photorhabdus* sp. fue aislada a partir de larvas de *Galleria mellonella* infestadas con nematodos entomopatógenos del género *Heterorhabditis*, recuperados de muestras de suelo obtenidas en Calilegua, Jujuy, Argentina. El aislamiento de la bacteria simbiote se realizó a partir de la recolección de hemolinfa de larvas desinfectadas superficialmente y su cultivo en placas de agar NBTA. Se seleccionaron colonias con características morfológicas típicas del género. Este proceso permitió garantizar la obtención de un aislamiento axénico necesario para los estudios posteriores.

Hipótesis

La secuenciación y caracterización genómica de la cepa P32 de *Photorhabdus* sp. permitirá identificar genes de virulencia y clústeres de genes productores de metabolitos secundarios con potencial bioactivo, lo que contribuirá al desarrollo de nuevas estrategias de control biológico de plagas y enfermedades en los cultivos.

Objetivo general

Realizar la secuenciación y caracterización genómica de la cepa P32 de *Photorhabdus* sp. utilizando herramientas bioinformáticas.

Objetivos específicos

- Obtener la secuencia genómica de la cepa P32.
- Ensamblar y anotar el genoma de la cepa.
- Identificar factores de virulencia asociados con la actividad insecticida.
- Detectar clústeres de genes productores de metabolitos secundarios con potencial bioactivo.
- Analizar la relación filogenética de la cepa con otras especies del género *Photorhabdus*.

Materiales y métodos

1. Purificación de ADN y secuenciación.

Para la obtención del ADN, la bacteria se cultivó en 5 ml de caldo Luria-Bertani (LB) a 28 °C con agitación continua (200 rpm) durante 48 h y posteriormente se centrifugó a 5000 ×g durante 5 min. El ADN genómico se obtuvo utilizando el kit de purificación de ADN genómico Wizard (Promega, Madison, WI, EE. UU.) siguiendo las instrucciones del fabricante para la purificación de ADN de bacterias Gram negativas. La secuencia genómica se realizó en plataforma Illumina, en la Unidad de Genómica de INTA Castelar.

2. Control de calidad de lecturas, ensamblado del genoma y optimización.

El control de calidad de las secuencias crudas se realizó con FASTQC (Versión 0.74) en plataforma Galaxy (Afgan *et al.*, 2022) utilizando los parámetros por defecto.

Para realizar el ensamblado en estado de draft o borrador se utilizó el software CLC Genomics Workbench (Qiagen) (versión 23.0.5). En las opciones del ensamblaje, seleccionamos el tamaño mínimo de contig en 500 pb. Elegimos el “map reads to contigs” lento, con las penalizaciones: mismatch cost 2; Insertion cost 3; Deletion cost 3. Para el mapeo de umbrales de calidad, se dejaron los valores por defecto: length fraction 0,5 y similarity fraction 0,8.

Luego de realizar el ensamblado “*de novo*” del genoma, se intentó realizar una optimización del mismo, mediante el software Geneious Prime (versión 2023.2.1) (www.geneious.com). Para esto, se utilizó la herramienta “map reads to reference” para localizar las lecturas nuevamente sobre los contigs ensamblados en el paso anterior. Luego de obtener la secuencia consenso, se realizó nuevamente el ensamblado “*de novo*” en el software CLC Genomics Workbench. Se realizaron repeticiones.

3. Anotación del genoma.

La secuencia se subió a la base de datos pública GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) y la anotación automatizada completa del genoma de la cepa P32 se realizó en el National Centre for Biotechnology Information

(NCBI) y mediante el servidor RAST (Rapid Annotation using Subsystem Technology) (Aziz *et al.*, 2008). Para ello se utilizó el ensamblaje obtenido en CLC y se exportó como un archivo multifasta.

En la plataforma RAST, se completó el envío del borrador de la siguiente manera:

- Taxonomy ID: 29487 (obtenido de NCBI taxonomy ID);
- Taxonomy string: *Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales; Morganellaceae; Photorhabdus* (obtenido automáticamente a partir del Taxonomy ID).
- Genus: *Photorhabdus*
- Species: undefined
- Strain: P32
- Genetic Code: 11 (Archea, most Bacteria, most Virii, and some Mitochondria)

Finalmente, en Annotation settings:

- Choose RAST annotation: RASTtk
- Automatically fix errors?: YES
- Fix frameshift?: YES
- Build metabolic model?: YES

4. Construcción de base de datos personalizada o custom en BLAST (Basic Local Alignment Search Tool) y búsqueda de factores de virulencia por custom BLAST.

La base de datos personalizada se construyó obteniendo secuencias específicas de proteínas insecticidas de bacterias entomopatógenas desde bases de datos públicas GenBank <https://www.ncbi.nlm.nih.gov/genbank/> (Sayers *et al.*, 2021). A partir de la anotación obtenida previamente, se realizó la búsqueda local con Blastp (Altschul *et al.*, 1997) (ncbi-blast v 2.15) localmente utilizando como query la anotación del genoma y como base de datos la colección de toxinas bacterianas. La base de datos se creó utilizando el comando *makeblastdb* sobre el archivo multifasta conteniendo las secuencias de proteínas a utilizar:

```
makeblastdb -in proteinas.fasta -out database name -dbtype prot
```

Comando: makeblastdb; Entrada: -in; Nombre base de datos:-Out; Tipo de base de datos (proteínas o nucleótidos):-dbtype opciones: nucl o prot.

Las secuencias de proteínas utilizadas como query en este estudio están disponibles como material complementario y pueden consultarse en el siguiente enlace: https://drive.google.com/drive/folders/1--MbpB_g9MEw54w8rr0ej8qeAai?usp=sharing.

Además, se realizó una búsqueda manual de proteínas relacionadas a la actividad insecticida mediante palabras claves en el software Geneious Prime, luego se realizó la traducción de las secuencias a proteína y se realizó una búsqueda en la base de datos de proteínas de NCBI (BlastP) de cada secuencia.

5. Búsqueda de clústers de genes potencialmente productores de metabolitos secundarios con antiSMASH.

Para realizar la búsqueda de clusters de genes productores de metabolitos secundarios bioactivos, se utilizó AntiSMASH (Blin *et al.*, 2023). Utilizamos el archivo fasta con el ensamblaje “*de novo*”. En la rigurosidad de detección dejamos por defecto “relaxed” y en las opciones de “Extra features” seleccionamos “All On” para un informe detallado.

6. Análisis filogenéticos.

Análisis de genes ortólogos:

Se realizó con OrthoANI (Orthologous Average Nucleotide Identity) (versión 0.93.1) (<https://www.ezbiocloud.net/tools/orthoani>). Para ello, se comparó el genoma de *Photorhabdus* P32 con 9 genomas diferentes de especies de *Photorhabdus* tipo: *Photorhabdus aballayi* (GenBank accession no. GCA_026153155.1), *Photorhabdus bodei* (GenBank accession no. GCA_003287735.1), *Photorhabdus caribbeanensis* (GenBank accession no. GCA_018448885.1), *Photorhabdus cinérea* (GenBank accession no. GCA_011189575.1), *Photorhabdus hainanensis* (GenBank accession no. GCA_018448965.1), *Photorhabdus laumondii* subsp. *laumondii* TTO1 (GenBank accession no. GCA_003343245.1), *Photorhabdus stackebrandtii* (GenBank accession no. GCA_011189495.1), *Photorhabdus tasmaniensis* (GenBank accession no. GCA_011189505.1) y *Photorhabdus thracensis* (GenBank accession no. GCA_001010285.1). Los genomas de las 9 cepas

utilizadas fueron descargadas de la base de datos de genomas del NCBI (<https://www.ncbi.nlm.nih.gov/genome/>). En los ajustes se configuró el número máximo de threads en 10 y los demás parámetros se dejaron por defecto.

Análisis taxonómico:

La taxonomía del genoma completo de la cepa se realizó en la plataforma del servidor Type Strain Genome Server (TYGS) (<https://tygs.dsmz.de/>) (Meier-Kolthoff y Goker, 2019). Este análisis permite determinar si un genoma pertenece a una especie conocida o si podría representar una nueva especie, basándose en métricas como identidad genómica promedio, distancia filogenética y relaciones con cepas tipo. Se realizó hibridación digital ADN-ADN (dDDH) con GGDC 2.1 Este comparó el genoma de *Photorhabdus* P32 con genomas de referencia de *Photorhabdus* spp. Se calculan tres valores de dDDH (d0, d4 y d6). Además se generó un árbol filogenético basado en comparación genómica con otras *Photorhabdus* utilizando el método GBDP (Genome BLAST Distance Phylogeny).

Identificación de islas genómicas:

Para la identificación de islas genómicas, se utilizó el software IslandViewer 4 (<http://www.patohgenic.sfu.ca/islandviewer/>) (Bertelli *et al.*, 2017). La anotación del genoma *Photorhabdus* P32, fue alineada con el genoma de referencia *Photorhabdus luminescens* subsp. *laumondii* TTO1 disponible en la base de datos de la plataforma.

Comparación de contenido genético:

Se utilizó el marco de base de datos eficiente para análisis comparativos del genoma utilizando la plataforma de servidor BLAST Score Ratios (EDGAR) (<https://edgar3.computational.bio.uni-giessen.de>) (Blom *et al.*, 2016) para comparar el borrador del genoma de *Photorhabdus* P32 con otras especies estrechamente relacionadas: *Photorhabdus aballayi* (GenBank accession no. GCA_026153155.1), *Photorhabdus bodei* (GenBank accession no. GCA_003287735.1) y *Photorhabdus laumondii* subsp. *laumondii* TTO1 (GenBank accession no. GCA_003343245.1).

Resultados y discusión

1. Secuenciación, control de calidad de lecturas, ensamblado del genoma y optimización.

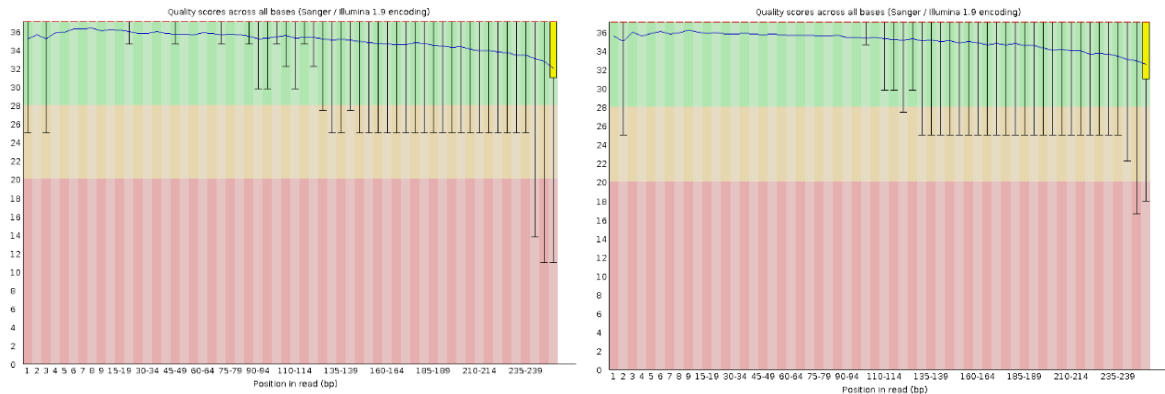
El análisis de calidad de las lecturas forward y reverse de la secuenciación del genoma *Photorhabdus* P32 en FASTQC indicó una correcta secuenciación. En la tabla 1 se presenta el resumen del análisis de las secuencias en estudio. La codificación de las secuencias es Illumina 1.9, es decir utiliza un sistema de puntuación de calidad de Phred+33. Las secuencias totales procesadas en ambos análisis fueron 7989252. No se filtraron secuencias, ya que no se establecieron condiciones de Trimming. El mismo fue realizado previamente en la Unidad de Genómica INTA. El rango de longitud de las secuencias fue de 35 – 251 pb. El contenido total de citosina y guanina en las secuencias fue de 44%.

Tabla 1. Resumen del análisis de las secuencias forward y reverse en FASTQC de la secuenciación del genoma *Photorhabdus* P32

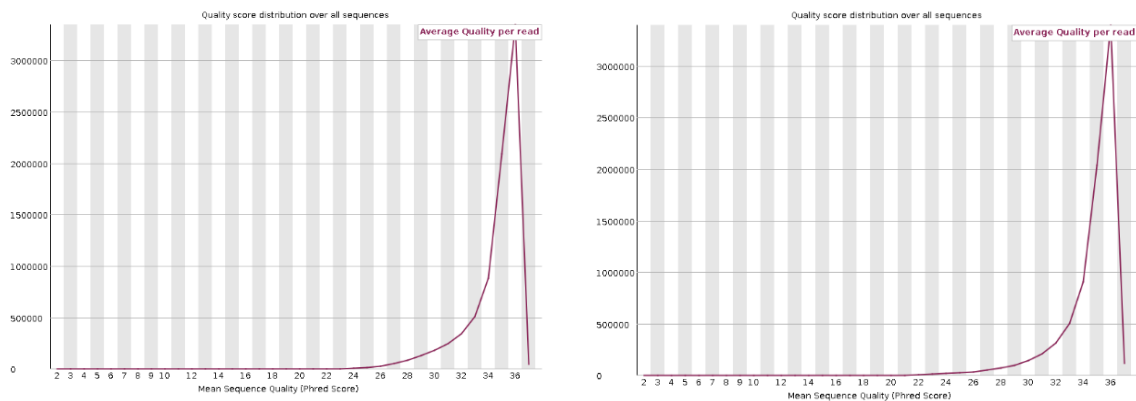
	Valores	
	Forward	Reverse
Tipo de archivos	Conventional base calls	Conventional base calls
Codificación	Sanger / Illumina 1.9	Sanger / Illumina 1.9
Secuencias totales	7989252	7989252
Total de bases	1.9 Gbp	1.9 Gbp
Secuencias filtradas	0	0
Longitud de las secuencias	35-251	35-251
%GC	44	44

Con respecto a la calidad de las bases, el análisis y el informe de puntuación de calidad por secuencia nos indican que la media (quality score) es de 33Q, con una calidad muy buena (figura 1.A y B). Podemos notar que el valor disminuye conforme nos desplazamos a lo largo de las posiciones de la secuencia, lo cual es esperado debido a las características de los sistemas de secuenciación. Estos sistemas se destacan por su eficacia inicial en las primeras bases, pero experimentan una disminución progresiva a medida que avanzan a lo largo de la secuencia. Podemos observar también que el cuartil inferior en las últimas posiciones de las lecturas de

forward y reverse se encuentran por debajo de 20. En la figura 1.C y D, podemos visualizar la calidad del conjunto de las lecturas, la mayor parte de las secuencias presenta un valor de calidad alto y muy pocas secuencias se encuentran en un bajo rango de calidad.



A. Calidad de secuencias por base. Forward (izquierda), Reverse (derecha)



B. Puntuaciones de calidad por secuencia. Forward (izquierda), Reverse (derecha)

Figura 1. Análisis de calidad de secuencias en FASTQC

Ensamblaje “de novo”:

Luego de realizar el ensamblaje en CLC Genomics Workbench, se obtuvo un genoma de 5.36 Megabases (Mb) compuesto por 255 contigs y el N50 del ensamblaje fue de 56166. El tamaño del borrador es similar al tamaño obtenido de la secuenciación de otras especies del mismo género (Palma *et al.* 2016; Somvanshi *et al.*, 2019). En la tabla 2 se presentan los resultados obtenidos del ensamblaje, incluyendo scaffolds (ensamblado inicial). El contig de menor longitud fue de 501 pb y el de mayor longitud

fue de 173982 pb. La distribución de los nucleótidos con scaffolding se presentan en la tabla 3.

Tabla 2. Mediciones de contigs del ensamblaje “de novo” en CLC Genomics Workbench del genoma (incluyendo regiones scaffolded) y luego del mapeo a la referencia para la optimización del ensamblaje

	Ensamblado Inicial	Ensamblado Optimizado
N75 (pb)	30175	34760
N50 (pb)	56166	66768
N25 (pb)	99604	103526
Mínimo (pb)	501	729
Máximo (pb)	173982	157558
Promedio (pb)	20999	35143
Conteo	255	121
Total (pb)	5354785	4252264

Tabla 3. Distribución de nucleótidos con scaffolding del ensamblaje “de novo”

Nucleótido	Cantidad	Porcentaje (%)
Adenina (A)	1540557	28,77
Citocina (C)	1135938	21,21
Guanina (G)	1137506	21,24
Timina (T)	1534554	28,66
No asignados (N)	6230	0,12

Optimización del ensamblado:

A partir del ensamblaje obtenido previamente, se intentó realizar una optimización de su calidad, utilizando la primera versión de ensamblado como referencia para el mapeo de las lecturas originales. Finalizado este paso, se tomaron las secuencias consenso y se realizó nuevamente el ensamblaje. En la tabla 2 se presentan los resultados obtenidos del nuevo ensamblaje (ensamblaje optimizado). El número de contigs obtenidos en la optimización del ensamblaje “de novo” se redujo de 255 a 121. El N50 fue de mayor tamaño y el tamaño mínimo de contig aumentó a 729. El tamaño máximo de contig se redujo a 157558 y se observa una reducción del tamaño total del genoma, de 5.35 Mb a 4.25 Mb, lo cual podría generar una pérdida importante de información teniendo en cuenta el tamaño promedio de las especies de *Photorhabdus*. Para continuar con el análisis se decidió utilizar el ensamblaje “de novo” obtenido

originalmente. El análisis preliminar del ensamblaje permitió clasificar a la cepa P32 como perteneciente a la especie *Photorhabdus aballayi*. Los detalles de esta clasificación se desarrollan más adelante.

2. Anotación del genoma.

El proyecto de genoma borrador fue depositado en la base de datos pública GenBank (Benson *et. al*, 2012) con el número de acceso JAZGJN000000000 (https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_038085095.1/). La cepa P32 tiene un tamaño de genoma de 5.354.785 pb y un contenido de G+C del 42,5%. El borrador del genoma tiene un total de 4695 genes. De estos, 4473 corresponden a genes codificantes de proteínas (CDSs), y 81 genes de ARN, que se dividen en 61 ARNt y 6 ARN no codificantes (ARNnc). Se encontraron 142 pseudogenes, lo que refleja la complejidad del genoma. Finalmente, se detectaron 4 regiones CRISPR Arrays, que podrían estar asociadas a sistemas de defensa bacteriana. El análisis funcional realizado con el servidor RAST destacó la diversidad de subsistemas genómicos presentes en la cepa, los cuales están distribuidos en 347 categorías. Las principales categorías incluyen procesos relacionados con el metabolismo de aminoácidos y derivados (334 genes), metabolismo de proteínas (202 genes) y cofactores, vitaminas y pigmentos (182 genes). Estas categorías representan funciones esenciales para la adaptabilidad de la bacteria, incluyendo su capacidad para sintetizar metabolitos secundarios y responder al estrés ambiental. Otros procesos destacados incluyen el transporte de membrana (116 genes) y la respiración (78 genes), lo que sugiere una alta capacidad metabólica y versatilidad funcional.

En la Figura 2 se presenta la distribución detallada de estas categorías, destacando el predominio de genes relacionados con el metabolismo primario y secundario. Este perfil refleja las capacidades biosintéticas y adaptativas de la cepa, las cuales son clave para su interacción simbiótica y su potencial uso biotecnológico.

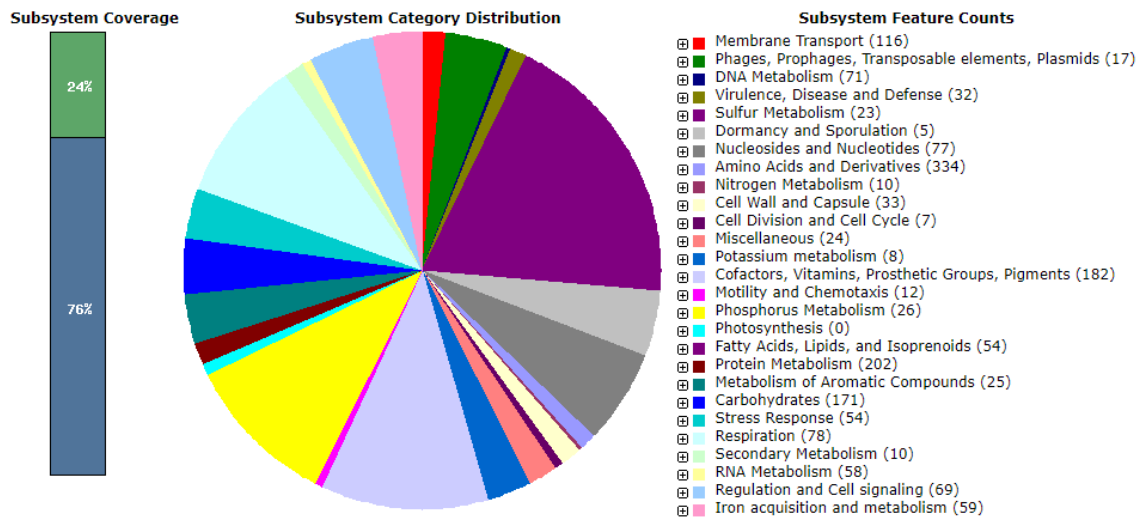


Figura 2. Clasificación funcional de genes en la cepa P32 de *Photorhabdus*.

3. Construcción de bases de datos personalizadas en BLAST y búsqueda de factores de virulencia

Para realizar la búsqueda local en BLAST, se utilizó una base de datos con 1077 secuencias de proteínas insecticidas bacterianas. En la tabla 4 se presentan los resultados de las proteínas insecticidas identificadas con un porcentaje de identidad \geq a 80%. Las especies objetivo describen datos experimentales depositados en la base de datos de especificidad del BPPRC (Bacterial Pesticidal Protein Resource Center). Las proteínas identificadas Pra1Aa1 y Prb1Aa1 (porcentaje de identidad 98.51 y 98.33 % respectivamente), poseen actividad insecticida sobre la especie *G. mellonella* (Blackburn *et al.*, 2006) y *P. xilostella* (Daborn *et al.*, 2002). Sin embargo, las proteínas por separado, poseen baja toxicidad, siendo necesaria la presencia de las 2 proteínas para ejercer un efecto insecticida total (proteínas insecticidas binarias) (Waterfield *et al.*, 2005; Li *et al.*, 2014). La proteína App1Aa1 (identidad 85.40 %) fue identificada como alpha-xenorhabdolysin en la especie *Photorhabdus khanii*. Estudios realizados sobre las proteínas Mcf1Ac1 (identidad 81.22 %) y Mcf1Aa1 (identidad 80.57 %), sugieren actividad como toxinas proapoptótica (Dowling *et al.*, 2004; Daborn *et al.*, 2002). La proteína Mcf1Ba1 (identidad 80.50 %) es otra toxina insecticida que demostró actividad provocando lisis nuclear en estudios de expresión recombinante

en *E. coli* (Waterfield *et al.*, 2003). Se identificaron también tres quitinasas (identidad 93.22, 99.05 y 99.29 % respectivamente).

Otras proteínas identificadas corresponden a las quitinasas. El rol de estas, es el de degradar la quitina, un biopolímero constitutivo del exoesqueleto, membrana y matriz del intestino de insectos. Al degradar la quitina de la matriz peritrófica del aparato digestivo de los insectos, permitirían el acceso de las toxinas al intestino desde el lumen (ffrench-Constant, 2023).

Tabla 4. Proteínas potencialmente insecticidas identificadas con BLAST personalizado (custom) en el genoma de *Photorhabdus* sp. P32 con especies objetivo.

Proteínas insecticidas	Porcentaje de identidad	Orden objetivo	Especie objetivo
Pra1Aa1	98,51	Lepidoptera	<i>G. mellonella</i> , <i>Plutella xylostella</i>
Prb1Aa1	98,33	Lepidoptera	<i>G. mellonella</i> , <i>P. xylostella</i>
App1Aa1	85,40	-	-
Mcf1Ac1	81,22	Lepidoptera	<i>Manduca sexta</i>
Mcf1Aa1	80,57	Lepidoptera	<i>M. sexta</i>
Mcf1Ba1	80,50	Lepidoptera	<i>M. sexta</i>
Chitinase (EC 3.2.1.14) CDS	99,05	-	-
Chitinase (EC 3.2.1.14) CDS	93,22	-	-
Chitinase (EC 3.2.1.14) CDS	99,29	-	-

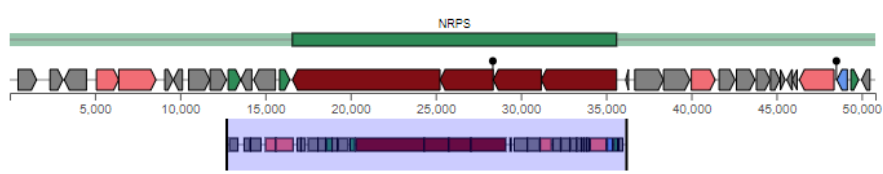
4. Búsqueda de clústers de genes potencialmente productores de metabolitos secundarios con antiSMASH.

El análisis realizado con la herramienta antiSMASH identificó cuatro regiones genómicas que contienen clústeres de genes con 100% de similitud a clústeres previamente caracterizados (figura 3). Entre estos, tres corresponden a clústeres codificantes de péptido sintetasa no ribosomal (NRPS, por sus siglas en inglés), mientras que el cuarto está relacionado con la producción de sideróforos.

- Región genómica 4.1 (contig 4, localización 1–50,805):
Se identificó un clúster de genes con 100% de similitud al grupo responsable de la biosíntesis del desipéptido “taxllaid A” en *Xenorhabdus bovienii* SS-2004 (MIBiG accession BGC0001133). Este desipéptido tiene una secuencia característica en código de una letra: T-A-X-L-L-L-A (donde X puede ser L, F o Y) y ha sido reportado con actividad citotóxica (Kronenwerth *et al.*, 2014).
- Región genómica 31.1 (contig 31, localización 1–85,290):
Este clúster corresponde al NRPS más largo y continuo descrito en bacterias, denominado Kolossin A, encontrado originalmente en *Photorhabdus luminescens* (MIBiG accession BGC0001641). Este compuesto tiene importantes implicancias en actividades antimicrobianas y de interacción ecológica (Bode *et al.*, 2015).
- Región genómica 16.1 (contig 16, localización 37,424–61,888):
Se encontró un clúster de genes involucrado en la producción de los sideróforos putrebactina y avaroferrina, con 100% de similitud al clúster de *Xenorhabdus budapestensis* (MIBiG accession BGC0001870). Este clúster utiliza una vía de biosíntesis independiente de NRPS (NIS, por sus siglas en inglés). Los sideróforos son esenciales para captar hierro en el ambiente del hospedador, transportándolo como un complejo sideróforo-hierro hacia la célula bacteriana, donde el hierro se reduce y se libera para su uso metabólico (Crosa, 1989).
- Región genómica 53.1 (contig 55, localización 1–25,043):
Corresponde a un clúster productor de compuestos bioactivos ririwpeptide A, B y C, caracterizados en *Photorhabdus laumondii* subsp. *laumondii* TTO1 (MIBiG accession BGC0002286). Estos metabolitos secundarios son de interés por su potencial aplicación biotecnológica (Wang *et al.*, 2019).

Región 4.1 NRPS

Location: 1 - 50,805 nt. (total: 50,805 nt) [Show pHMM detection rules used](#) [Download region SVG](#) [Download region GenBank file](#)

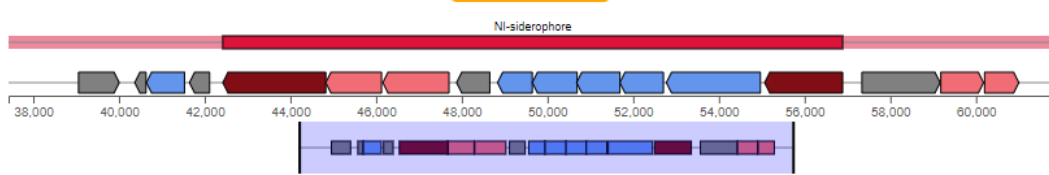


Legend:



Región 16.1 NI-siderophore

Location: 37,424 - 61,888 nt. (total: 24,465 nt) [Show pHMM detection rules used](#) [Download region SVG](#) [Download region GenBank file](#)

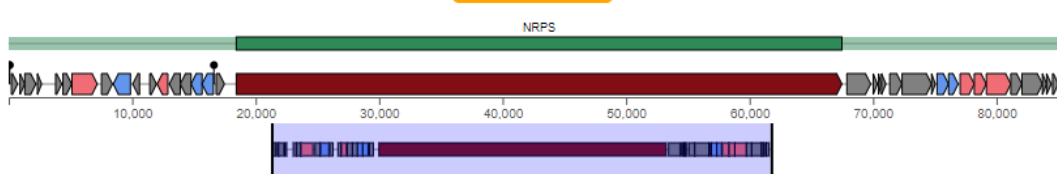


Legend:



Región 31.1 NRPS

Location: 1 - 85,290 nt. (total: 85,290 nt) [Show pHMM detection rules used](#) [Download region SVG](#) [Download region GenBank file](#)

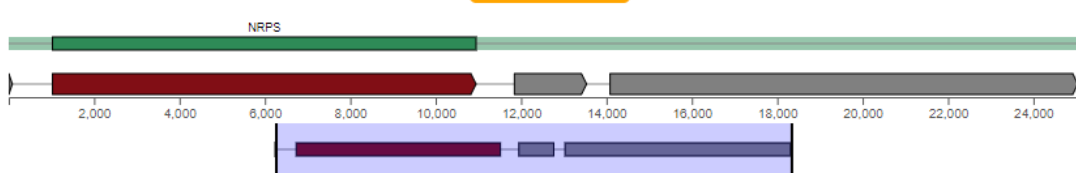


Legend:



Región 53.1 NRPS

Location: 1 - 25,043 nt. (total: 25,043 nt) [Show pHMM detection rules used](#) [Download region SVG](#) [Download region GenBank file](#)



Legend:

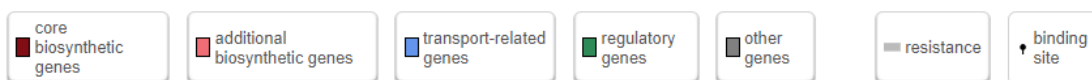


Figura 3. clústeres de genes potencialmente productores de metabolitos secundarios detectados con antiSMASH en el genoma de *Photorhabdus* P32 con 100% de similitud

5. Análisis filogenéticos.

Análisis de genes ortólogos:

El análisis de identidad promedio de nucleótidos (*OrthoANI*) mostró un valor de 96,94% entre *Photorhabdus* sp. P32 y la cepa tipo *Photorhabdus aballayi* (DSM 15197), lo que sugiere que ambas cepas pertenecen a la misma especie, según el umbral establecido para circunscribir especies procariotas (95-96%) (Richter y Rosselló-Móra, 2009).

En el análisis comparativo se incluyeron otras cepas tipo, como *Photorhabdus luminescens* subsp. *laumondii* TT01 (DSM 15139) y *Photorhabdus temperata* (DSM 19984), las cuales presentaron valores de identidad significativamente menores, iguales o inferiores al 92,28%. Estos resultados confirman que la cepa P32 no comparte una identidad cercana con otras especies analizadas, reforzando su identificación como parte de *Photorhabdus aballayi*.



Heatmap generated with OrthoANI values calculated from the OAT software. Please cite Lee *et al.* 2015.

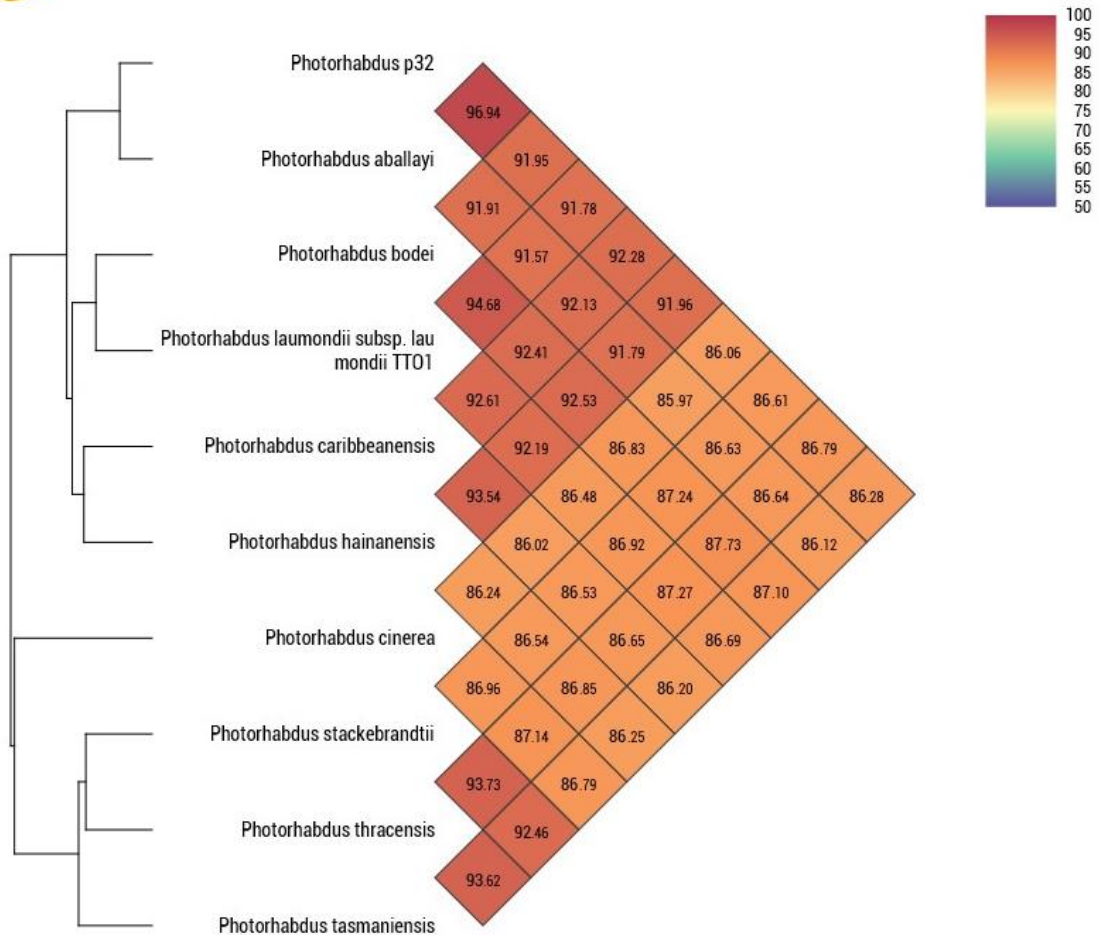


Figura 4. Mapa de calor de *Photorhabdus* sp. P32 con especies estrechamente relacionadas generado con OrthoANI, valores calculados a partir del software OAT.

Análisis taxonómico:

El análisis taxonómico de la cepa *Photorhabdus* sp. P32 mediante la estimación de dDDH utiliza distintas fórmulas aplicables a genomas borradores (draft genomes), como lo proponen Meier-Kolthoff *et al.* (2013). Los valores obtenidos están detallados en la Tabla 5. La fórmula d4 es la más adecuada para genomas incompletos, ya que es menos susceptible a variaciones en la longitud del genoma y evita sesgos asociados a regiones no secuenciadas.

La cepa P32 mostró un valor de dDDH de 71,5% con *Photorhabdus aballayi* APURE al aplicar la fórmula d4, lo que confirma su identificación dentro de esta especie, ya

que supera el umbral del 70% propuesto para delimitar especies bacterianas (Meier-Kolthoff *et al.*, 2013). Además, los valores obtenidos mediante otras fórmulas, como d0 y d6, también respaldan esta asignación al mostrar consistencia en los resultados, todos por encima del 70%.

Cabe destacar que la diferencia en el contenido de G+C entre la cepa P32 y *P. aballayi* APURE es de apenas 0,09%, lo que refuerza aún más esta clasificación. En la Figura 5 se presenta el árbol filogenético inferido con distancias GBDP, el cual confirma la relación filogenética de clado entre la cepa P32 y *Photorhabdus aballayi*. Estos resultados en conjunto brindan una base sólida para identificar inequívocamente a *Photorhabdus* sp. P32 como parte de la especie *Photorhabdus aballayi*.

Tabla 5. Comparaciones por pares de genomas de cepa *Photorhabdus* P32 versus genomas de cepa tipo

Cepa en estudio	Cepas comparadas	dDDH* (d ₀ , en %)	IC** (d ₀ , en %)	dDDH (d ₄ , en %)	IC (d ₄ , en %)	dDDH (d ₆ , en %)	IC (d ₆ , en %)	Diferencia de G+C (%)
<i>Photorhabdus</i> P32	<i>Photorhabdus aballayi</i> APURE	81,3	[77,4 - 84,7]	71,5	[68,5 - 74,4]	82,4	[79,1 - 85,3]	0,09
	<i>Photorhabdus luminescens</i> subsp. <i>venezuelensis</i> JAR	79,5	[75,5 - 82,9]	63,2	[60,3 - 66,1]	79,0	[75,6 - 82,1]	0,05
	<i>Photorhabdus luminescens</i> ATCC 29999	79,3	[75,4 - 82,8]	61,8	[58,9 - 64,6]	78,5	[75,1 - 81,6]	0,09
	<i>Photorhabdus luminescens</i> subsp. <i>mexicana</i> MEX47-22	76,8	[72,8 - 80,3]	61,3	[58,5 - 64,1]	76,3	[72,8 - 79,4]	0,01
	<i>Photorhabdus kleinii</i> DSM 23513	66,8	[62,9 - 70,4]	47,3	[44,7 - 49,9]	63,8	[60,4 - 67,0]	0,12
	<i>Photorhabdus antumapuensis</i> UCH- 936	68,8	[64,9 - 72,5]	46,4	[43,8 - 48,9]	65,2	[61,8 - 68,4]	0,02
	<i>Photorhabdus noenieputensis</i> DSM 25462T	70,6	[66,7 - 74,3]	46,0	[43,5 - 48,6]	66,5	[63,1 - 69,7]	0,14
	<i>Photorhabdus aegyptia</i> BA1T	63,5	[59,8 - 67,1]	45,9	[43,3 - 48,5]	60,7	[57,4 - 63,9]	0,04
	<i>Photorhabdus caribbeanensis</i> HG29	75,0	[71,0 - 78,6]	45,6	[43,0 - 48,1]	69,8	[66,4 - 73,1]	0,1
	<i>Photorhabdus bodei</i> LJ24-63	65,2	[61,4 - 68,8]	45,3	[42,8 - 47,9]	61,8	[58,6 - 65,0]	0,16
	<i>Photorhabdus akhurstii</i> DSM 15138	71,9	[67,9 - 75,5]	44,8	[42,3 - 47,4]	67,1	[63,7 - 70,3]	0,26
	<i>Photorhabdus hainanensis</i> DSM 22397	72,7	[68,7 - 76,3]	44,6	[42,0 - 47,1]	67,6	[64,2 - 70,9]	0,31
	<i>Photorhabdus laumondii</i> DSM 15139	66,1	[62,3 - 69,7]	44,4	[41,8 - 46,9]	62,3	[58,9 - 65,4]	0,32
	<i>Photorhabdus hindustanensis</i> H1T	71,5	[67,5 - 75,1]	44,3	[41,8 - 46,9]	66,6	[63,2 - 69,8]	0,31
	<i>Photorhabdus namnaonensis</i> PB45.5	70,9	[67,0 - 74,6]	43,9	[41,4 - 46,4]	65,9	[62,6 - 69,2]	0,2

*dDDH: Hibridación digital ADN-ADN; **IC: Intervalo de confianza

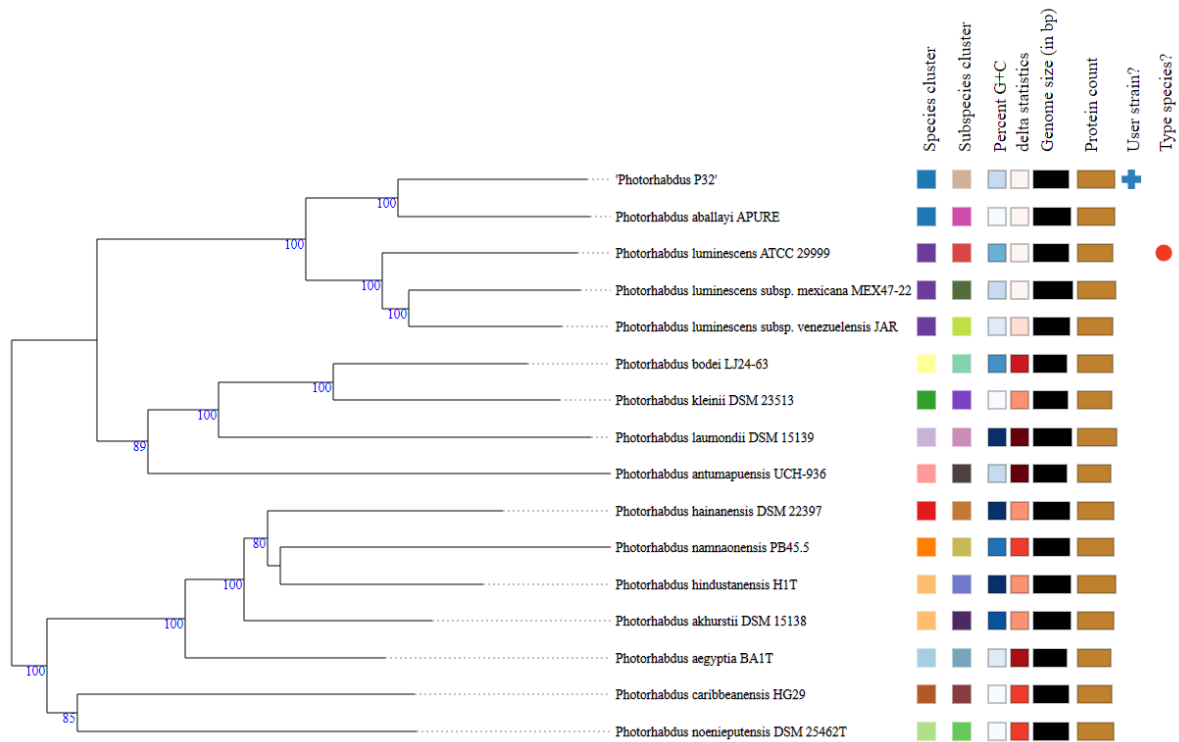


Figura 5. Árbol filogenético inferido con FastME 2.1.6.1 a partir de distancias GBDP calculadas a partir de secuencias de genomas. Las longitudes de las ramas se escalan en términos de la fórmula de distancia GBDP d5. Los números sobre las ramas son valores de soporte de pseudo-bootstrap de GBDP > 60 % de 100 replicas, con un soporte de rama promedio del 93,4 %. El árbol tenía sus raíces en el punto medio

Identificación de islas genómicas:

El genoma de *Photorhabdus* P32 contiene genes obtenidos a través de procesos de transferencia horizontal. Se reconocieron 14 islas genómicas en la alineación con el genoma de referencia de *Photorhabdus luminescens* subsp. *laumondii* TTO1 (figura 6). Se encontraron 12 genes relacionados a la biosíntesis de toxinas-antitoxinas (tabla 6). Además, se encontraron genes relacionados a la interacción con el hospedador, 4 genes relacionados a la producción de proteínas tipo fimbria y 1 gen relacionado a proteína de choque térmico.

Tabla 6. Genes en Islas genómicas de *Photorhabdus* P32 contra *Photorhabdus luminescens* subsp. *laumondii* TTO1.

Inicio Isla	Fin de Isla	Longitud	Inicio del gen	Fin del gen	Producto
276634	328155	51521	291232	292663	Glucosilhidrolasa (Proteína relacionada con toxinas TNT)
378363	406584	28221	399086	399463	VapC (Familia de toxinas tipo PIN, ribonucleasa)
378363	406584	28221	399463	399666	Sistema toxina-antitoxina tipo II, familia VapB antitoxina
2859823	2865701	5878	2862028	2862273	Sistema toxina-antitoxina tipo II, familia HicA toxina
3094468	3127139	32671	3095757	3095954	Sistema toxina-antitoxina tipo II, familia HicB antitoxina
3178987	3213784	34797	3199826	3199968	Sistema toxina-antitoxina tipo II, familia HicA toxina
3754847	3778311	23464	3760927	3761378	Sistema toxina-antitoxina tipo II, familia YdaT
3754847	3778311	23464	3769022	3769333	Sistema toxina-antitoxina tipo II, familia RelE/ParE toxina
4077282	4128400	51118	4128053	4128400	Sistema toxina-antitoxina tipo II, familia HicB antitoxina
5398430	5463530	65100	5399642	5399824	Módulo de adición antitoxina familia HigA
5506097	5570033	63936	5528920	5529377	Endonucleasa de restricción, toxina dominio fold 5, proteína que contiene dominio
5506097	5570033	63936	5547124	5547656	Endonucleasa de restricción, toxina dominio fold 5, proteína que contiene dominio
757424	779986	22562	774779	775318	Proteína fimbrial
757424	779986	22562	775402	775950	Proteína fimbrial
757424	779986	22562	778619	779368	Chaperona pilus/periplásmica fimbrial
757424	779986	22562	779393	779986	Proteína fimbrial
757424	779986	22562	779979	780554	Proteína fimbrial
4077282	4128400	51118	4116195	4116410	Proteína de choque en frío



Figura 6. Islas genómicas de *Photorhabdus* P32 contra *Photorhabdus luminescens* subsp. *laumondii* TTO1. El círculo exterior verde representa el genoma en megabases de la cepa P32. Los colores rojo y azul, indican la detección integrada y IslandPath-DIMOB respectivamente de las islas genómicas.

Comparación de contenido genético:

La cepa *Photorhabdus* P32 comparte un genoma core de 2941 CDS con las especies *P. laumondii* subsp. *laumondii* TTO1, *P. aballayi* y *P. bodei*, como se muestra en la figura 7. Esto es similar al encontrado por otros autores (Lulamba *et al.*, 2021; Tobias *et al.*, 2016). Esto indica una importante conservación de genes comunes en las difentes especies. Además, 370 CDS se compartieron exclusivamente con la especie *P. aballayi*. El pan genoma de las 4 especies consistió en 6916 CDS. El análisis

identificó un total de 453 genes únicos (singletons) en *Photorhabdus* sp. P32. De estos, 193 genes tienen anotaciones funcionales claras y alta confianza, mientras que el resto corresponde a genes con funciones desconocidas o predicciones con menor soporte.

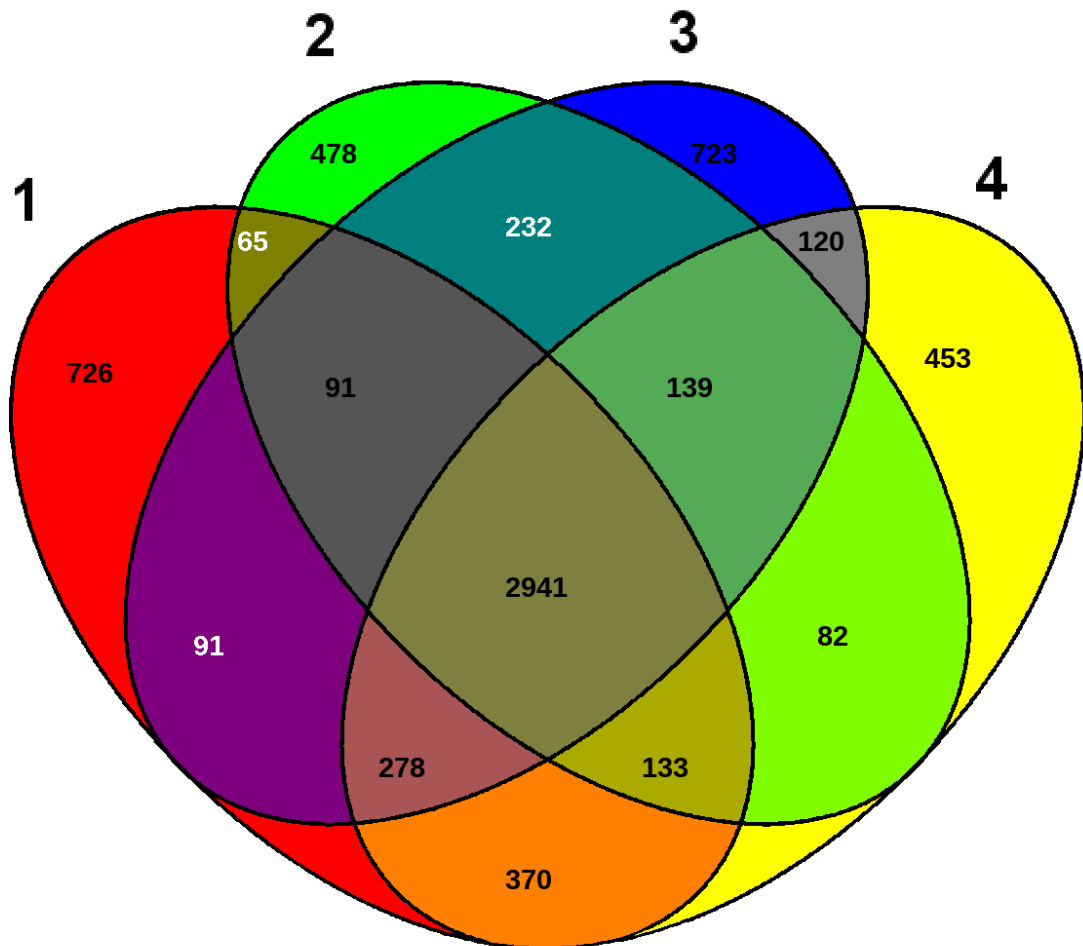


Figura 7. Diagrama de Venn de *Photorhabdus* P32 (4), *Photorhabdus* aballayi (1), *Photorhabdus* bodei (2), *Photorhabdus* laumondii subsp. laumondii TTO1 (3). Genes únicos: áreas roja, verde, azul y amarilla. Genoma core: área de intersección de todos los colores. Subconjuntos: intersecciones de 2 a 3 colores.

Conclusiones

En este trabajo, se logró secuenciar y caracterizar genómicamente la cepa *Photorhabdus* P32, aislada de larvas de *Galleria mellonella* infestadas con nematodos del género *Heterorhabditis*. Los principales hallazgos incluyen:

A través de análisis comparativos de ANI y dDDH utilizando genomas de referencia, se confirmó que la cepa P32 pertenece a la especie *Photorhabdus aballayi*, con valores superiores al 95% y 70%, respectivamente, cumpliendo con los estándares internacionales para la delimitación de especies bacterianas.

El análisis del genoma reveló la presencia de clústeres genéticos involucrados en la producción de metabolitos secundarios bioactivos, como taxllaid A (efecto citotóxico), kolossin A (péptido antimicrobiano), y sideróforos (putrebactina y avaroferrina). Estas moléculas poseen propiedades citotóxicas y capacidad de captura de hierro, destacando el potencial biotecnológico de la cepa.

Se identificaron proteínas con actividad insecticida, como toxinas del sistema tipo II y quitinasas. Estos factores refuerzan la posible aplicación de la cepa P32 como agente de control biológico.

Se detectaron 14 islas genómicas, que incluyen genes relacionados con sistemas toxina-antitoxina y adhesinas tipo fimbria, lo que podría explicar la interacción simbiótica entre *Photorhabdus* y sus hospedadores.

Este estudio respalda la capacidad de *Photorhabdus* P32 para ser utilizada en programas de manejo integrado de plagas, especialmente en cultivos agrícolas, como una alternativa ecológica a los pesticidas químicos.

Próximos estudios a realizar con la cepa *Photorhabdus* P32

1. Estudio de proteínas insecticidas

- Clonación y expresión de App1Aa1 y Mcf1Ac1 en *E. coli*.
- Bioensayos con insectos.

2. Optimización del ensamblaje

- Secuenciación con tecnología Nanopore (lecturas largas) para cerrar el genoma.

3. Producción de metabolitos secundarios

- Cultivo de *Photorhabdus* P32 para producción de metabolitos.
- Pruebas de actividad insecticida.

Bibliografía

- Abd-Elgawad, M. M. (2021). *Photorhabdus* spp.: An overview of the beneficial aspects of mutualistic bacteria of insecticidal nematodes. *Plants*, 10(8), 1660.
- Afgan, E., Nekrutenko, A., Grünig, B. A., Blankenberg, D., Goecks, J., Schatz, M. C., Briggs, P. J. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic acids research*, 50(W1).
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), 3389-3402.
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC genomics*, 9, 1-15.
- Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2012). GenBank. *Nucleic acids research*, 41(D1), D36-D42.
- Bertelli, C., Laird, M. R., Williams, K. P., Simon Fraser University Research Computing Group, Lau, B. Y., Hoad, G., Winsor, G.L. & Brinkman, F. S. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic acids research*, 45(W1), W30-W35.
- Blackburn, M. B., Farrar, R. R., Novak, N. G., & Lawrence, S. D. (2006). Remarkable susceptibility of the diamondback moth (*Plutella xylostella*) to ingestion of Pir toxins from *Photorhabdus luminescens*. *Entomologia experimentalis et applicata*, 121(1), 31-37.
- Blin, K., Shaw, S., Augustijn, H. E., Reitz, Z. L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B.R., Metcalf, W.W., Helfrich, E.N.J., van Wezel, G.P., Weber, T. (2023). antiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic acids research*, gkad344.
- Blom, J., Kreis, J., Spänig, S., Juhre, T., Bertelli, C., Ernst, C., & Goesmann, A. (2016). EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic acids research*, 44(W1), W22-W28.
- Bode, H. B., Brachmann, A. O., Jadhav, K. B., Seyfarth, L., Dauth, C., Fuchs, S. W., Kaiser, M., Waterfield, N.R., Sack, H., Heinemann, S.H., & Arndt, H. D. (2015). Structure Elucidation and Activity of Kolossin A, the D-/L-Pentadecapeptide Product of a Giant Nonribosomal Peptide Synthetase. *Angewandte Chemie International Edition*, 54(35), 10352-10355.
- Coupland, J. A. M. E. S., Abd-Elgawad, M. M., & Askary, T. H. (2017). Beneficial nematodes and the changing scope of crop protection. In *Biocontrol agents: entomopathogenic and slug parasitic nematodes* (pp. 26-42). Wallingford UK: CABI.
- Crosa, J. H. (1989). Genetics and molecular biology of siderophore-mediated iron transport in bacteria. *Microbiological reviews*, 53(4), 517-530.
- Daborn, P. J., Waterfield, N., Silva, C. P., Au, C. P. Y., Sharma, S., & Ffrench-Constant, R. H. (2002). A single *Photorhabdus* gene, makes caterpillars floppy (mcf), allows *Escherichia coli* to persist within and kill insects. *Proceedings of the National Academy of Sciences*, 99(16), 10742-10747.
- Dowling, A. J., Daborn, P. J., Waterfield, N. R., Wang, P., Streuli, C. H., & Ffrench-Constant, R. H. (2004). The insecticidal toxin Makes caterpillars floppy (Mcf) promotes apoptosis in mammalian cells. *Cellular microbiology*, 6(4), 345-353.

- Duchaud, E., Rusniok, C., Frangeul, L., Buchrieser, C., Givaudan, A., Taourit, S., Bocs, S., Boursaux-Eude, C., Chandler, M., Charles, J.F., Dassa, E., Derose, R., Derzelle, S., Freyssinet, G., Gaudriault, S., Médigue, C., Lanois, A., Powell, K., Siguier, P., Vincent, R., Wingate, V., Zouine, M., Glaser, P., Boemare, N., Danchin, A., Kunst, F. (2003). The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. *Nature biotechnology*, 21(11), 1307-1313.
- French-Constant, R.H. (2023) *Photorhabdus* toxins as novel delivery systems for agriculture and medicine. *Advances in Insect Physiology*, 65, 261-302.
- Forst, S., & Clarke, D. (2002). Bacteria-nematode symbiosis. In *Entomopathogenic nematology* (pp. 57-77). Wallingford UK: CABI Publishing.
- Grewal, P. S., Ehlers, R. U., & Shapiro-Ilan, D. I. (Eds.). (2005). *Nematodes as biocontrol agents*. Cabi publishing.
- Hazir, S., Kaya, H. K., Stock, S. P., & Keskin, N. (2003). Entomopathogenic nematodes (Steinernematidae and Heterorhabditidae) for biological control of soil pests. *Turkish journal of Biology*, 27(4), 181-202.
- Hurst, S.t., Rowedder, H., Michaels, B., Bullock, H., Jackobeck, R., Abebe-Akele, F., Durakovic, U., Gately, J., Janicki, E., Tisa, L.S. (2015). Elucidation of the *Photorhabdus temperata* genome and generation of a transposon mutant library to identify motility mutants altered in pathogenesis. *Journal of Bacteriology*, 197(13), 2201-2216.
- Hussaini, S. S. (2014). Potential of entomopathogenic nematodes in integrated pest management. In *Integrated pest management* (pp. 193-223). Academic Press.
- Jurat-Fuentes, J. L., & Jackson, T. A. (2012). Bacterial entomopathogens. *Insect pathology*, 265-349.
- Koppenhöfer, A. M., Grewal, P. S., & Fuzy, E. M. (2007). Differences in penetration routes and establishment rates of four entomopathogenic nematode species into four white grub species. *Journal of invertebrate pathology*, 94(3), 184-195.
- Kronenwerth, M., Bozhüyük, K. A., Kahnt, A. S., Steinhilber, D., Gaudriault, S., Kaiser, M., & Bode, H. B. (2014). Characterisation of Taxlllids A–G; natural products from *Xenorhabdus indica*. *Chemistry—A European Journal*, 20(52), 17478-17487.
- Lacey, L. A., & Georgis, R. (2012). Entomopathogenic nematodes for control of insect pests above and below ground with comments on commercial production. *Journal of nematology*, 44(2), 218.
- Li, Y., Hu, X., Zhang, X., Liu, Z., Ding, X., Xia, L., & Hu, S. (2014). *Photorhabdus luminescens* PirAB-fusion protein exhibits both cytotoxicity and insecticidal activity. *FEMS microbiology letters*, 356(1), 23-31.
- Lulamba, T. E., Green, E., & Serepa-Dlamini, M. H. (2021). Genome assembly and annotation of *Photorhabdus heterorhabditis* strain ETL reveals genetic features involved in pathogenicity with its associated entomopathogenic nematode and anti-host effectors with biocontrol potential applications. *Gene*, 795, 145780.
- Machado, R. A., Muller, A., Ghazal, S. M., Thanwisai, A., Pagès, S., Bode, H. B., Hussein, M.A., Khalil, K.M., Tisa, L. S. (2021). *Photorhabdus heterorhabditis* subsp. *aluminescens* subsp. nov., *Photorhabdus heterorhabditis* subsp. *heterorhabditis* subsp. nov., *Photorhabdus australis* subsp. *thailandensis* subsp. nov., *Photorhabdus australis* subsp. *australis* subsp. nov., and *Photorhabdus aegyptia* sp. nov. isolated from *Heterorhabditis* entomopathogenic nematodes. *International Journal of Systematic and Evolutionary Microbiology*, 71(1), 004610.
- Meier-Kolthoff, J. P., & Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature communications*, 10(1), 2182.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC bioinformatics*, 14, 1-14.

- Orozco, R. A., Lee, M. M., & Stock, S. P. (2014). Soil sampling and isolation of entomopathogenic nematodes (Steinernematidae, Heterorhabditidae). *JoVE (Journal of Visualized Experiments)*, (89), e52083.
- Palma, L., Del Valle, E. E., Frizzo, L., Berry, C., & Caballero, P. (2016). Draft genome sequence of *Photorhabdus luminescens* strain DSPV002N isolated from Santa Fe, Argentina. *Genome Announcements*, 4(4), 10-1128.
- Parihar, R. D., Dhiman, U., Bhushan, A., Gupta, P. K., & Gupta, P. (2022). Heterorhabditis and *Photorhabdus* symbiosis: a natural mine of bioactive compounds. *Frontiers in microbiology*, 13, 790339.
- Richter, M., & Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, 106(45), 19126-19131.
- Sayers, E. W., Beck, J., Bolton, E. E., Bourexis, D., Brister, J. R., Canese, K., Comeau, C.D., Funk, K., Kim, S., Klimke, W., Marchler-Bauer, A., Landrum, M., Lathrop, S., Lu, Z., Madden, T.L., O'Leary, N., Phan, L., Rangwala, S.H., Schneider, V.A., Skripchenko, Y., Wang, J., Ye, J., Trawick, B.W., Pruitt, K.D. & Sherry, S. T. (2021). Database resources of the national center for biotechnology information. *Nucleic acids research*, 49(D1), D10.
- Sheets, J., & Aktories, K. (2017). Insecticidal toxin complexes from *Photorhabdus luminescens*. *The molecular biology of Photorhabdus bacteria*, 3-23.
- Somvanshi, V. S., Dubay, B., Kushwah, J., Ramamoorthy, S., Vishnu, U. S., Sankarasubramanian, J., Rajendhran, J. & Rao, U. (2019). Draft genome sequences for five *Photorhabdus* bacterial symbionts of entomopathogenic Heterorhabditis nematodes isolated from India. *Microbiology Resource Announcements*, 8(4), 10-1128.
- Tobias, N. J., Mishra, B., Gupta, D. K., Sharma, R., Thines, M., Stinear, T. P., & Bode, H. B. (2016). Genome comparisons provide insights into the role of secondary metabolites in the pathogenic phase of the *Photorhabdus* life cycle. *BMC genomics*, 17, 1-11.
- Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y. M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., Wang, B., Evans, R., Miyamoto, Y., Cheng, K., Kosina, S., De Raad, M., Silva, L., Luhrs, A., Lubbe, A., Hoyt, D.W., Francavilla, C., Otani, H., Deutsch, S., Washton, N.M., Rubin, E.M., Mouncey, N.J., Visel, A., Northen, T., Cheng, J., Bode, H.B., & Yoshikuni Y. (2019). CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. *Nature microbiology*, 4(12), 2498-2510.
- Waterfield, N. R., Daborn, P. J., Dowling, A. J., Yang, G., Hares, M., & Ffrench-Constant, R. H. (2003). The insecticidal toxin makes caterpillars floppy 2 (Mcf2) shows similarity to HrmA, an avirulence protein from a plant pathogen. *FEMS Microbiology Letters*, 229(2), 265-270.
- Waterfield, N., George Kamita, S., Hammock, B. D., & Ffrench-Constant, R. (2005). The *Photorhabdus* Pir toxins are similar to a developmentally regulated insect protein but show no juvenile hormone esterase activity. *FEMS microbiology letters*, 245(1), 47-52.