



**UNIVERSIDAD DE CHILE**

**FACULTAD DE MEDICINA**

**ESCUELA DE POSGRADO**

“Identificación y caracterización de genes para la síntesis de metabolitos especializados y bacteriocinas de *Pseudomonas* sp. I1 involucrados en la capacidad de inhibir el crecimiento de patógenos microbianos”

**Diana Eugenia Guaya Iñiguez**

**AFE PARA OPTAR AL GRADO DE MAGÍSTER EN MICROBIOLOGÍA**

Prof. Dr. Juan Carlos Salazar  
Garrido

**Director de AFE**

Prof. Dr. Víctor Antonio García  
Angulo

**Director de AFE**

2021

UNIVERSIDAD DE CHILE  
FACULTAD DE MEDICINA  
ESCUELA DE POSTGRADO

## INFORME DE APROBACIÓN AFE DE MAGÍSTER

Se informa a la Comisión de Grados Académicos de la Facultad de Medicina, que la AFE de Magíster presentada por la candidata

DIANA EUGENIA GUAYA IÑIGUEZ

ha sido aprobada por la Comisión Informante de Tesis como requisito para optar al Grado de Magíster en Microbiología en el Examen de Defensa de AFE rendido el día 20 de mayo de 2021

Prof. Dr. Juan Carlos Salazar  
Garrido  
Director de Tesis/AFE  
Programa de Microbiología y Micología,  
ICBM  
Facultad de Medicina, Universidad de Chile

Prof. Dr. Víctor Antonio García  
Angulo  
Director de Tesis/AFE  
Programa de Microbiología y Micología,  
ICBM  
Facultad de Medicina, Universidad de Chile

### COMISIÓN INFORMANTE DE AFE

Prof. Dr. Felipe Del Canto F.

Prof. Dra. María Teresa Ulloa F.

Prof. Dra. Marcela Wilkens A.

Prof. Dra. Cecilia Toro U.

**Presidente Comisión**

## **DEDICATORIA**

A Dios, mis padres, mis hermanas, mis tios y a tí mi ángel en el cielo, quienes han sido mi guía para recorrer este camino y llegar a este punto de mi carrera profesional, por su amor y apoyo pese a la distancia y a todo lo que estos años han implicado estar lejos de casa.

A Pablo, por su amor y apoyo incondicional, por estar en los buenos y malos momentos durante toda esta etapa.

A Daniela, Constanza y Shuliana por ser las mejores amigas y por estar siempre presente.

Diana

## **AGRADECIMIENTOS**

Primeramente, quiero agradecer a la Facultad de Medicina de la Universidad de Chile especialmente al programa de Magíster de Microbiología, por la oportunidad, enseñanza, paciencia, apoyo e interés durante esta etapa de aprendizaje. A la Dra. Cecilia Toro quien ha estado pendiente de cada uno de los procesos requeridos durante el magíster.

Así mismo, agradezco a mis tutores de tesis: Dr. Víctor García y Dr. Juan Carlos Salazar, por compartir sus conocimientos, por la dedicación y apoyo que han brindado a este trabajo.

Agradezco a la comisión evaluadora: Dr. Felipe Del Canto Fuentes, Dra. Cecilia Toro Ugalde, Dra. María Teresa Ulloa, Dra. Marcela Wilkens Anwandter por participar activa y constructivamente en el desarrollo de esta tesis.

Finalmente agradezco a la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT) de Ecuador, por la beca otorgada para poder realizar el magíster.

## ÍNDICE GENERAL

ÍNDICE DE FIGURAS.....	7
ÍNDICE DE TABLAS.....	8
ABREVIATURAS.....	9
1. RESUMEN.....	11
2. ABSTRACT.....	13
3. INTRODUCCIÓN.....	15
3.1. Resistencia antimicrobiana.....	15
3.2. Bacteriocinas.....	16
3.2.1. Clasificación de las bacteriocinas.....	17
3.3. Bacteriocinas de clase I y II.....	18
3.4. Bacteriocinas de clase III.....	18
3.4.1. Microcinas.....	18
3.4.2. Colicinas.....	19
3.4.2.1. Organización genética, biosíntesis y liberación.....	19
3.4.2.2. Mecanismo de acción.....	21
3.5. Metabolitos secundarios.....	21
3.5.1. Sintetasas de péptidos no ribosomales (NRPS).....	22
3.6. Otros productos del metabolismo secundario con capacidad antimicrobiana.....	24
3.6.1. Policétidos.....	24
3.6.2. Betalactonas.....	24
3.6.3. Terpenos.....	25
3.7. Aislados bacterianos del Río Mapocho con capacidad inhibitoria contra patógenos.....	25
3.8. Bacteriocinas y metabolitos secundarios antimicrobianos producidos por <i>Pseudomonas</i> .....	27
3.9. Minería Genómica.....	29
4. HIPÓTESIS.....	31
5. OBJETIVOS.....	<b>¡Error! Marcador no definido.</b>
5.1. OBJETIVO GENERAL.....	<b>¡Error! Marcador no definido.</b>
5.2. OBJETIVOS ESPECÍFICOS.....	<b>¡Error! Marcador no definido.</b>
5.2.1. Objetivo específico 1:.....	<b>¡Error! Marcador no definido.</b>
5.2.2. Objetivo específico 2:.....	31
6. MATERIALES Y MÉTODOS.....	32

6.1. Identificación de <i>clusters</i> genéticos que codifican para la biosíntesis de metabolitos especializados y bacteriocinas en <i>Pseudomonas</i> sp. I1. ....	32
6.2. Análisis de los <i>clusters</i> obtenidos por antiSMASH y BAGEL4 .....	32
7. RESULTADOS.....	34
7.1. Análisis del genoma de la bacteria inhibidora <i>Pseudomonas</i> sp. I1 mediante minería genómica.....	34
7.2. Identificación <i>in silico</i> de genes para la biosíntesis de lokisina en <i>Pseudomonas</i> sp. I1 .....	35
7.3. Identificación <i>in silico</i> de genes para la biosíntesis de pioverdina en <i>Pseudomonas</i> sp. I1 .....	41
7.4. Identificación de una vía que codifica hipotéticamente para betalactona.....	43
7.5. Caracterización de las bacteriocinas identificadas en el genoma de <i>Pseudomonas</i> sp. I1 .....	45
7.6. Búsqueda de homólogos de la piocina S14 e identificación de motivos conservados en el dominio citotóxico de las bacteriocinas identificadas por BAGEL4.....	50
8. DISCUSIÓN .....	55
8.1. Identificación de metabolitos especializados y bacteriocinas mediante minería genómica.....	55
8.2. Identificación de un <i>cluster</i> de genes en <i>Pseudomonas</i> sp. I1 para la síntesis de lokisina. ....	56
8.3. Identificación de un <i>cluster</i> de genes en <i>Pseudomonas</i> sp. I1 para la síntesis de pioverdina.....	57
8.4. Identificación de una vía que codifica hipotéticamente para betalactona.....	58
8.5. Caracterización de las bacteriocinas identificadas en <i>Pseudomonas</i> sp. I1 por BAGEL4 .....	59
8.6. Identificación de homólogos de la piocina S14 y motivos conservados del dominio citotóxico C-terminal. ....	61
9. CONCLUSIONES .....	64
10. BIBLIOGRAFÍA.....	65

## ÍNDICE DE FIGURAS

Figura 1. Genes implicados en la producción de colicina E9 en <i>E. coli</i> .....	20
Figura 2. Organización de los dominios proteicos de la colicina E9. ....	20
Figura 3. <i>Cluster</i> de genes biosintéticos de metabolitos secundarios que codifican proteínas con actividades enzimáticas .....	22
Figura 4. Organización de las NRPS. ....	23
Figura 5. Ensayos de inhibición de los aislados I1 e I2 contra patógenos bacterianos humanos.....	26
Figura 6. Organización de dominios de la piocinas tipo S. ....	28
Figura 7. <i>Clusters</i> identificados en <i>Pseudomonas</i> sp. I1 por antiSMASH para la síntesis de lokisina. ....	40
Figura 8. <i>Cluster</i> de vías NRPS que codifican para pioverdina. A y B.....	43
Figura 9. <i>Cluster</i> identificado por antiSMASH para la biosíntesis de una betalactona en <i>Pseudomonas</i> sp. I1. ....	44
Figura 10. Genes de bacteriocinas identificados en el genoma de <i>Pseudomonas</i> sp. I1 por BAGEL4.....	<b>¡Error! Marcador no definido.</b>
Figura 11. Alineamiento de carocina D hipotética de <i>Pseudomonas</i> sp. I1 (parte superior) con la carocina D de <i>P. carotovorum</i> (parte inferior). ....	<b>¡Error! Marcador no definido.</b>
Figura 12. Organización de dominios de la piocina S14.....	50
Figura 13. Dominios hipotéticos de colicina E9 y colicina de <i>Pseudomonas</i> sp. I1 identificados por BAGEL4.....	50
Figura 14. Predicción de estructura secundaria del dominio C-terminal de las bacteriocinas.....	53
Figura 15. Alineamiento de las 953 proteínas que conservan el dominio C-terminal de la piocina S14.....	54

## ÍNDICE DE TABLAS

Tabla 1. <i>Clusters</i> biosintéticos de metabolitos secundarios y bacteriocinas en el genoma de <i>Pseudomonas</i> sp. I1 identificados mediante antiSMASH y BAGEL4. ....	35
Tabla 2. Bacteriocinas incluidas en la base de datos local.. <b>¡Error! Marcador no definido.</b>	
Tabla 3. Proteínas homólogas de la piocina S14 .....	50

## **ABREVIATURAS**

**OMS:** Organización Mundial de la Salud

**LAB:** Bacterias productoras de ácido láctico

**PK:** Policétido

**PKS:** Policétido sintasa

**NRPS:** Sintetasas de péptidos no ribosomales

**NRP:** Péptidos no ribosomales

**RiPPs:** Péptidos de síntesis ribosomal y modificados postraduccionalmente

**BGC:** *clusters* de genes biosintéticos

**MiBIG:** Información mínima sobre grupos de genes biosintéticos

**A:** Adenilación

**T:** Tiolación

**C:** Condensación

**E:** Epimerización

**N-terminal:** amino-terminal

**C-terminal:** carboxilo-terminal

**TE:** Tioesterasa

**KS:** Cetoácido sintasa

**AT:** Acil transferasa

**ACP:** Transportador de grupos acilo

**PCR:** Reaccion en cadena de la polimerasa

**MRSA:** *Staphylococcus aureus* resistente a la meticilina

**LB:** Medio *Lysogeny Broth*

**NCBI:** Centro Nacional de Información Biotecnológica

**L-Glu:** Ácido L-glutámico

**D-Tyr:** D-tirosina

**L-Dab:** Diamino butirato

**HHX<sub>14</sub>NX<sub>8</sub>HX<sub>8</sub>H:** H, histidina. N, asparagina, la X hace referencia a cualquier aminoácido y el subíndice al número de estos que hay entre los aminoácidos conservados.

**MX<sub>46</sub>GX<sub>2</sub>LSF:** M, metionina. G, glicina. L, leucina. S, serina. F, fenilalanina, la X hace referencia a cualquier aminoácido y el subíndice al número de aminoácidos que hay entre los aminoácidos conservados.

## 1. RESUMEN

La resistencia a los antimicrobianos está generando un problema de salud pública a nivel mundial. Se estima que para el año 2050 no se contará con antimicrobianos efectivos para combatir las enfermedades causadas por bacterias multirresistentes especialmente en pacientes inmunocomprometidos, lo que aumentaría drásticamente la tasa de morbilidad y mortalidad. La mayor parte de los antimicrobianos utilizados en medicina son inicialmente producidos por microorganismos. Por lo tanto, investigar metabolitos especializados de bacterias con funciones antimicrobianas puede dirigir a encontrar nuevos antibióticos. En nuestro grupo de investigación, se aislaron dos bacterias del Río Mapocho en Santiago de Chile con potencial para la producción de compuestos antimicrobianos, evidenciado por la inhibición del crecimiento de diversos patógenos bacterianos humanos. Mediante secuenciación genómica estos aislados se identificaron como *Pseudomonas* sp. I1 y *Desemzia incerta* I2. Sin embargo, se desconocen los mecanismos de inhibición de dichos aislados. En este trabajo se investigaron los probables determinantes de inhibición en el genoma de *Pseudomonas* sp. I1. Con base en su potencial inhibitorio, se postula que esta bacteria posee genes que codifican bacteriocinas y/o proteínas responsables de la síntesis de metabolitos especializados asociados a la inhibición del crecimiento de patógenos microbianos. Para corroborar dicha hipótesis, se planteó aplicar técnicas de minería genómica con el objetivo de identificar y caracterizar genes para la síntesis de metabolitos especializados y bacteriocinas de *Pseudomonas* sp. I1 probablemente involucrados en la capacidad de inhibir el crecimiento de patógenos microbianos.

En primer lugar, se identificaron *clusters* de genes que codifican para la síntesis de metabolitos especializados y bacteriocinas hipotéticos en el genoma de *Pseudomonas* sp. I1, empleando los programas de minería genómica antiSMASH y BAGEL4. El programa antiSMASH reveló la presencia de cinco *clusters* con vías tipo NRPS (sintetasas de péptidos no ribosomales) para la producción de lokisina, nunamicina, cichopectina y pioverdina (2 *clusters*), además de una vía hipotética para la producción de una betalactona. El análisis utilizando BAGEL4 permitió

identificar la presencia de tres *clusters* que codifican bacteriocinas, correspondientes a piocinas tipo S. Posteriormente, se determinó que tres de los *clusters* de vías NRPS identificados por antiSMASH, lokisina, nunamicina y cichopectina, no son independientes, sino que corresponden a fragmentos de un mismo *cluster*, cuyos productos están involucrados en la biosíntesis de lokisina y que fueron localizados en diferentes *contigs*. El análisis de los dos *clusters* que codifican hipotéticamente para pioverdina determinó que éstos codifican enzimas que forman parte de una misma vía y poseen los genes cuyos productos son esenciales para la biosíntesis de pioverdina. El *cluster* para betalactona hipotética posee muy baja identidad de genes con la vía más parecida en la base de datos y posiblemente produciría un nuevo tipo de betalactona. Respecto a las bacteriocinas hipotéticas, en dos de ellas no fue posible identificar todos los dominios típicos de las piocinas S, mientras que la tercera bacteriocina presenta una combinación de dominios nueva, ausente en las piocinas caracterizadas hasta el momento, y se denominó piocina S14. El análisis mediante BLASTP identificó 20 homólogos de la piocina S14, conservados en diferentes especies de *Pseudomonas*. Se determinó que la piocina S14 posee un dominio carboxilo-terminal que no conserva motivos nucleasas ya descritos, por lo tanto, pertenecería a una nueva familia de DNAsas. Una búsqueda del dominio carboxilo-terminal de la piocina S14 identificó 953 proteínas en diversas especies bacterianas de  $\gamma$ -proteobacterias que poseen este dominio. Un alineamiento múltiple de estas secuencias identificó 13 aminoácidos conservados que podrían ser esenciales para la función de esta nueva familia de DNAsas.

En conclusión, se determinó que el genoma de *Pseudomonas* sp. I1 posee genes cuyos productos estarían asociados a la síntesis de moléculas con potencial inhibitorio, como péptidos de síntesis no ribosomal, betalactona y una nueva bacteriocina.

**Palabras clave:** bacteriocinas, resistencia antimicrobiana, metabolitos especializados

## 2. ABSTRACT

Antimicrobial resistance comprises a major public health problem worldwide. It is estimated that by 2050 there will be no effective antimicrobials to treat diseases caused by multi-drug resistant bacteria, especially in immunocompromised patients, drastically increasing the morbidity and mortality rates. Most of the antimicrobials used in medicine are produced by microorganisms. Therefore, research on specialized metabolites with antimicrobial functions produced by bacteria may lead to the development of new, effective antibiotics. In our research group, two bacteria isolated from the Mapocho River, in Santiago of Chile, display a potential for the production of antimicrobial compounds, evidenced by the growth inhibition of human bacterial pathogens. Through genome sequencing and phylogenetic analysis, these isolates were identified as *Pseudomonas* sp. I1 and *Desemzia incerta* I2. Nonetheless, their mechanisms of inhibition are still unknown. In this work, the determinants of inhibition capacity in the *Pseudomonas* sp. I1 genome were investigated. On the basis of his inhibitory activity, it is hypothesized that in the genome of this species are encoded proteins responsible for the synthesis of specialized metabolites and bacteriocins associated with the inhibition of the growth of microbial pathogens. To corroborate this hypothesis, genome mining techniques were applied to identify and characterize genes for the synthesis of specialized metabolites and bacteriocins of *Pseudomonas* sp. I1 putatively involved in the ability to inhibit the growth of pathogens.

First, we identified hypothetical genetic clusters coding for the synthesis of specialized metabolites and bacteriocins in the genome of *Pseudomonas* sp. I1 using the genome mining programs antiSMASH and BAGEL4. Using The antiSMASH software the presence of 5 clusters with NRPS-like pathways for the production of lokisin, nunamycin, chichopeptin, and pyoverdine (2 clusters), in addition to a hypothetical pathway for production of a new betalactone, were detected. The analysis using BAGEL4 allowed the identification of three clusters that encode bacteriocins, all corresponding to S-type pyocins. Then, these clusters were characterized by comparisons with similar clusters in databases. These analyses

determined that three of the NRPS clusters identified by antiSMASH, namely the lokisin, nunamycin and chichopectin clusters, were not independent but comprise fragments of a single cluster whose products are involved in lokisin biosynthesis, located in different *contigs*. Furthermore, the two hypothetical pyoverdine clusters were determined to form a single pathway and to contain the genes whose products are essential for pyoverdine biosynthesis. The hypothetical betalactone cluster identified by antiSMASH may produce a new type of betalactone. While three hypothetical bacteriocins were identified, for two of them it was not possible to identify all the typical domains of S pyocins. The other bacteriocin has a new combination of domains not identified in the pyocins described so far and it was herein named pyocin S14. Using a BLASTP search, 20 homologs of the pyocin S14 were identified conserved in different species of *Pseudomonas*. Further, the pyocin S14 has a carboxyl-terminal domain that lacks the nucleases motifs described to date, therefore, it would represent a new family of DNases. In addition, 953 proteins that conserve the carboxyl-terminal domain of pyocin S14 in several bacterial species of  $\gamma$ -proteobacteria group were identified. A multiple alignment of these sequences identified 13 conserved amino acids that could be essential for the function of this new family of DNases.

In conclusion, the genome of *Pseudomonas* sp. I1 has genes whose products may be associated with the synthesis of molecules with inhibitory potential, such as non-ribosomal synthesis peptides, betalactone and a new bacteriocin.

**Keywords:** bacteriocins, antimicrobial resistance, specialized metabolites

### 3. INTRODUCCIÓN

#### 3.1. Resistencia antimicrobiana

El éxito de la medicina moderna se basa en gran parte en el uso de antibióticos para el control de infecciones bacterianas, siendo de particular importancia en pacientes inmunocomprometidos, procedimientos quirúrgicos y en pacientes que están en terapia contra el cáncer. Sin embargo, el manejo inadecuado a nivel hospitalario y comunitario y la falta de desarrollo de nuevos antimicrobianos ha favorecido la aparición de cepas con resistencia a los antimicrobianos actualmente en uso. Esto ha causado problemas de salud pública a nivel mundial, principalmente en el aumento de los costos en atención y hospitalización (Aslam et al. 2018; Morehead et al. 2018). Otra de las causas del desarrollo de resistencia a antimicrobianos es el uso descontrolado en la industria ganadera y alimentaria, donde se utilizan no solo para la prevención y el tratamiento de enfermedades, sino también para promover el crecimiento de los animales (Aslam et al. 2018).

La organización global sin fines de lucro *Research and Development Corporation* de Estados Unidos sostiene que alrededor del año 2050 la población se verá afectada al quedarse sin un agente antimicrobiano efectivo para tratar infecciones bacterianas, con un incremento dramático en las tasas de morbilidad y mortalidad. Se calcula que el costo de la resistencia a los antibióticos será aproximadamente de 100 billones de dólares en todo el mundo, causando 10 millones de muertes por año (Aslam et al. 2018; Morehead et al. 2018). En el marco de esta crisis mundial de resistencia a los antibióticos, la OMS elaboró una lista de priorización global de bacterias resistentes para guiar el desarrollo de nuevos antimicrobianos efectivos. Entre los patógenos con prioridad crítica se encuentran los Gram negativo *Acinetobacter baumannii* y *Pseudomonas aeruginosa* resistentes a carbapenemas y las *Enterobacteriaceae* (*Klebsiella pneumoniae*, *Escherichia coli*, *Enterobacter* spp., *Serratia* spp., *Proteus* spp., *Providencia* spp. y *Morganella* spp.) resistentes a carbapenemas y a cefalosporinas de 3<sup>era</sup> generación. Dentro de los patógenos con prioridad alta se encuentran *Enterococcus faecium* resistente a vancomicina, *Staphylococcus aureus* resistente a meticilina (MRSA), *Helicobacter pylori*

resistente a claritromicina, *Campylobacter* y *Salmonella* spp. resistentes a fluoroquinolonas y *Neisseria gonorrhoeae* resistente a cefalosporinas de 3<sup>era</sup> generación y a fluoroquinolonas. Entre los patógenos que conforman el nivel de prioridad media se encuentran *Streptococcus pneumoniae* resistente a la penicilina, *Haemophilus influenzae* resistente a ampicilina y *Shigella* spp. resistente a fluoroquinolonas (Tacconelli et al. 2018). Existen múltiples mecanismos de resistencia a antimicrobianos desarrollados por las bacterias. Por ejemplo, los principales mecanismos de resistencia de bacterias Gram negativo del género *Pseudomonas* comprenden: presencia de enzimas modificadoras del antibiótico (como las betalactamasas), alteraciones de membrana debido a la presencia de bombas de expulsión y mutaciones de las porinas transmembranas. En bacterias Gram positivo, como *S. aureus*, la resistencia en ocasiones se basa en la modificación de proteínas a las cuales el antibiótico usualmente inhibe, como en el caso de la resistencia a antibióticos betalactámicos (Medina et al. 2016; Subedi et al. 2018).

Los antibióticos actuales pueden ser producidos naturalmente por microorganismos como las bacterias saprófitas u hongos ambientales, pueden ser modificados sintéticamente (semisintéticos), o ser completamente sintetizados *in vitro* (sintéticos). En los microorganismos, la capacidad de producir moléculas antimicrobianas se desarrolla durante las interacciones antagónicas microorganismo-microorganismo.

Ante el problema de la resistencia antimicrobiana es necesario buscar alternativas para identificar moléculas con capacidad antimicrobiana que contribuyan al desarrollo terapéutico. Entre las opciones que existen se encuentran las bacteriocinas y metabolitos secundarios (Aslam et al. 2018; Morehead et al. 2018).

### **3.2. Bacteriocinas**

Las bacteriocinas son péptidos de síntesis ribosomal producidos por diferentes especies bacterianas para su autoconservación y ventaja competitiva. Las bacteriocinas pueden tener actividad antimicrobiana de amplio espectro o de

espectro reducido, esta última posibilita el control de patógenos específicos. El término bacteriocina es un concepto amplio que abarca desde péptidos catiónicos anfifílicos pequeños de entre 12 a 100 aminoácidos hasta proteínas completas multimodulares (Simons et al. 2020). La mayoría de las bacteriocinas conocidas son producidas por bacterias Gram positivo y en menor cantidad por especies Gram negativo (Meade et al. 2020; Zheng et al. 2018). Algunas bacteriocinas son utilizadas en la conservación y seguridad de alimentos, siendo de interés las producidas por bacterias productoras de ácido láctico (LAB, por sus siglas en inglés), que ayudan a controlar microorganismos durante la fermentación de productos lácteos. Sin embargo, las LAB producen bacteriocinas que no actúan frente a bacterias Gram negativo, debido a la composición de la pared bacteriana de estas últimas. No obstante, existen bacterias Gram negativo capaces de producir bacteriocinas que muestran actividad contra otras bacterias Gram negativo (Zheng et al. 2018; Langa et al. 2017). El mecanismo de acción de las bacteriocinas generalmente se basa en la formación de poros que causan desestabilización y permeabilización de la membrana de células sensibles o pueden tener actividad nucleasa sobre el DNA, rRNA y tRNA. Entre las potenciales aplicaciones propuestas para las bacteriocinas se encuentra su uso contra bacterias patógenas multirresistentes (Meade et al. 2020).

### **3.2.1. Clasificación de las bacteriocinas**

Actualmente las bacteriocinas son agrupadas en tres clases principales. La clase I está formada por pequeños péptidos, termolábiles y activos a nivel de membrana. Contienen aminoácidos inusuales como la dihidroalanina,  $\beta$ -metil-lantionina y lantionina. La presencia de estos aminoácidos es el resultado de modificaciones postraduccionales que consisten en la deshidratación y ciclación de residuos de aminoácidos específicos (Simons et al. 2017). Las bacteriocinas de clase II son polipéptidos pequeños que no contienen aminoácidos modificados y son termoestables, como la pediocina, bacteriocinas de dos péptidos y bacteriocinas lineales y circulares (Zheng et al. 2018). Las de clase III incluyen péptidos termolábiles relativamente grandes producidos por bacterias Gram negativo como *E. coli*, la cual produce bacteriocinas denominadas colicinas y microcinas. Entre

otras bacteriocinas de esta clase se encuentran también las producidas por *Pseudomonas* denominadas piocinas (Radaic et al. 2020; Meade et al. 2020; Simons et al. 2020).

### **3.3. Bacteriocinas de clase I y II**

Las bacteriocinas de clase I y II son producidas por bacterias Gram positivo, generalmente LAB. El sistema de biosíntesis de dichas bacteriocinas se encuentra codificado en *clusters* de genes que incluyen un gen estructural que codifica el pre-péptido, genes que codifican enzimas responsables de modificaciones postraduccionales (bacteriocinas de clase I), proteínas responsables del transporte, proteína de inmunidad que protege la célula productora y proteínas involucradas en la regulación de la expresión de los genes de biosíntesis. Estas bacteriocinas son sintetizadas como pre-péptidos inactivos que contienen un péptido señal en la región N-terminal, el cual es removido durante el transporte del pre-péptido. La acción antimicrobiana de estas bacteriocinas se debe a la desestabilización funcional de las membranas citoplasmáticas de las células sensibles. Este proceso se desarrolla en tres etapas básicas: unión de la bacteriocina a la membrana citoplasmática, inserción en la misma y formación de poros (Zheng et al. 2018; Meade et al. 2020; Chanos et al. 2016).

### **3.4. Bacteriocinas de clase III**

La mayoría de las bacteriocinas de clase III se ha aislado de *E. coli*; sin embargo, otros géneros como *Pseudomonas* o *Klebsiella* también pueden producir estas proteínas. Este grupo de bacteriocinas se ha clasificado en cuatro categorías: microcinas, colicinas, bacteriocinas similares a las colicinas (como las producidas por *Pseudomonas* llamadas piocinas o las producidas por *Klebsiella* denominadas klebicinas) y bacteriocinas de tipo cola de fago como las piocinas R y F producidas por *P. aeruginosa* (Simons et al. 2020).

#### **3.4.1. Microcinas**

Las microcinas son péptidos tóxicos secretados por enterobacterias. Se han descrito principalmente en *E. coli* y a diferencia de las colicinas poseen un tamaño

molecular inferior a 10 kDa. Las microcinas son generalmente hidrofóbicas y muestran una alta estabilidad al calor, pH extremo, proteasas y son solubles en metanol (Poey et al. 2006; Gillor et al. 2004).

Los genes implicados en la producción de las microcinas están organizados en operones que pueden encontrarse en el cromosoma o plásmidos. Dichos operones están constituidos por al menos tres tipos de genes: i) un gen que codifica el precursor de la microcina, ii) un gen que codifica el factor de inmunidad que protege a la cepa productora, y iii) genes que codifican el sistema de exportación necesario para la secreción de estos compuestos al medio. La microcina es sintetizada como un pre-péptido que durante la exportación es procesado para escindir el péptido líder. El sistema encargado de transportar a la microcina desde el citosol al medio extracelular es un transportador del tipo ABC que requiere de la hidrólisis de ATP para su función (Duquesne et al. 2007).

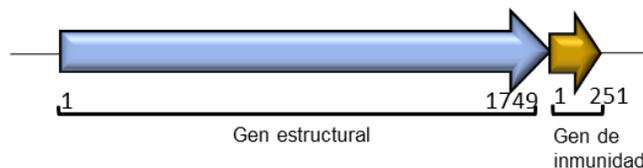
### **3.4.2. Colicinas**

#### **3.4.2.1. Organización genética, biosíntesis y liberación**

El *cluster* de genes implicado en la síntesis de colicinas está conformado típicamente por un gen estructural que codifica la bacteriocina y un gen para la proteína de inmunidad. Como ejemplo se muestra la colicina E9 producida por *E. coli* (Figura 1). Además, pueden poseer un gen de lisis que codifica una proteína que causa lisis de la célula productora para la liberación de la bacteriocina (Cascales et al. 2007). Las proteínas codificadas por los genes de lisis son pequeñas lipoproteínas que modifican la estructura de la pared celular, activan la fosfolipasa A de la membrana externa y matan a la célula productora (Ghazaryan et al. 2014). Las colicinas están organizadas en general en tres dominios proteicos. Estos se ejemplifican para la colicina E9 en la Figura 2. El dominio destructor que posee la actividad citotóxica está localizado en el extremo C-terminal (dominio III). En la parte central se localiza un dominio (II) de unión a receptor, que está involucrado en la detección de receptores específicos en la célula diana. Dichos receptores son proteínas de membrana externa que conservan una estructura de barril  $\beta$ , y generalmente corresponden a transportadores de micronutrientes como la vitamina

B12 y el hierro. Finalmente, el dominio (I) de translocación se localiza en extremo N-terminal de la proteína. Esta región promueve la internalización de la bacteriocina en la célula diana mediante la interacción con un sistema de transporte dependiente de Ton o Tol (Gillor et al. 2005; Jakes, 2016; Walker et al. 2002).

La liberación de estas bacteriocinas está acompañada por lisis de la célula productora, generalmente por acción de la proteína codificada por el gen de lisis, lo que resulta en la liberación de la bacteriocina al medio. Sin embargo, en muchos casos el gen responsable de la lisis no está en el operón de la bacteriocina y no se conoce. Las bacterias productoras de colicina están protegidas de la acción de su propia bacteriocina mediante la interacción con una proteína pequeña de inmunidad que se coproduce con la bacteriocina. La interacción de la proteína de inmunidad con la bacteriocina evita la acción del dominio citotóxico. Posteriormente, la proteína de inmunidad es liberada de la bacteriocina durante el proceso de translocación en la bacteria diana (Cascales et al. 2007; Ghazaryan et al. 2014).



**Figura 1. Genes implicados en la producción de colicina E9 en *E. coli*.** La colicina E9 está codificada en un operón bicistrónico con el gen estructural (azul) de 1749 nt y corriente abajo el gen de inmunidad (amarillo) de 251nt. En el caso de esta colicina la proteína estructural tiene un tamaño de 580 aminoácidos y la proteína de inmunidad de 86 aminoácidos.



- I Translocación
- II Unión al receptor
- III Citotóxico

**Figura 2. Organización de los dominios proteicos de la colicina E9.** El dominio de translocación (I), permite la translocación de la bacteriocina en la célula diana. Un dominio de unión a receptor (II) detecta receptores específicos de la célula diana y el dominio citotóxico (III) tiene actividad citotóxica. En el caso de esta colicina, este dominio posee actividad DNasa. Se indica la extensión de cada dominio, en aminoácidos, sobre la barra.

#### 3.4.2.2. Mecanismo de acción

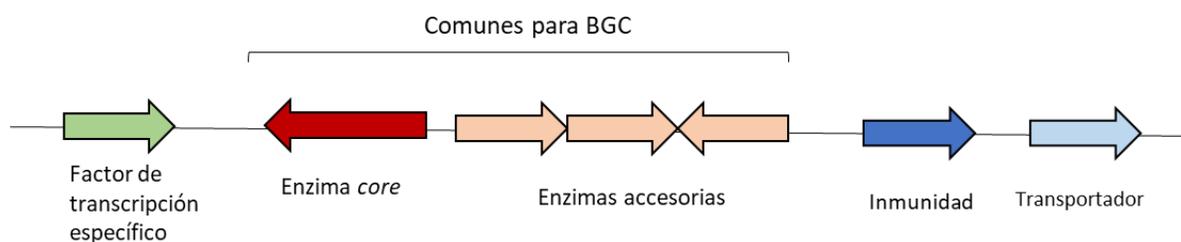
Una vez secretadas, las bacteriocinas son translocadas en la célula diana y alcanzan su objetivo letal. Este tipo de bacteriocinas posee dos mecanismos básicos de acción, la formación de poros en la membrana o una actividad endonucleasa que puede actuar sobre el DNA, rRNA o tRNA (Radaic et al. 2020; Dolejšová et al. 2019; Meade et al. 2020).

#### 3.5. Metabolitos secundarios

Los metabolitos secundarios son compuestos sintetizados por diversos microorganismos que cumplen funciones no esenciales para éstos, es decir su ausencia no es letal. Sin embargo, los metabolitos secundarios cumplen funciones importantes en los microorganismos, como las relacionadas con protección, competencia e interacciones entre especies. Además, juegan un papel importante en la adaptación fisiológica y la capacidad *fitness*, por lo que de manera más reciente se les ha denominado “metabolitos especializados” (Chavali y Rhee, 2018). Estos metabolitos varían ampliamente en estructura y actividad biológica, y algunos tienen valiosas aplicaciones farmacéuticas y agrícolas (Yan et al. 2018). Muchas de las interacciones antagónicas entre microorganismos están mediadas por metabolitos especializados. Entre las principales vías para la biosíntesis de metabolitos especializados con capacidad antimicrobiana producida por microorganismos se encuentran los grupos de genes de sintetasa de péptidos no ribosomales (NRPS), policétidos (PK), betalactonas y terpenos (Xu et al. 2019).

Los genes que codifican las actividades enzimáticas para producir cualquier metabolito secundario se organizan de manera contigua formando un *cluster* de genes biosintéticos (BGC, por sus siglas en inglés) (Figura 3). Éstos están constituidos por un gen *core* que codifica una proteína cuya actividad enzimática define la naturaleza del metabolito producido, genes accesorios de número variable que codifican enzimas que suelen modificar el esqueleto del metabolito mediante reducciones, oxidaciones, metilaciones, ciclaciones, entre otras reacciones. También pueden tener genes que codifican proteínas que se encargan del transporte, genes que codifican proteínas que protegen al microorganismo que lo

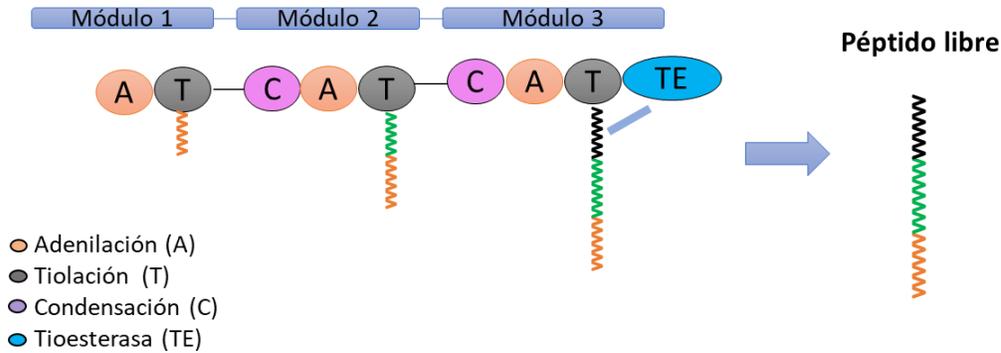
produce de su propia toxicidad y un gen que codifica un factor de transcripción específico (Keller 2019).



**Figura 3. Cluster de genes biosintéticos de metabolitos secundarios que codifican proteínas con actividades enzimáticas:** Estos *clusters* tienen al menos una enzima *core* (rojo) y enzimas accesorias (beige). Algunos BGCs contienen un factor de transcripción específico (verde), gen que codifica proteínas que protegen al microorganismo que lo produce de su toxicidad (azul) y transportador (azul celeste). Modificado de Keller 2019.

### 3.5.1. Sintetasas de péptidos no ribosomales (NRPS)

Las NRPS son enzimas multimodulares que catalizan la formación de enlaces peptídicos entre aminoácidos, para la síntesis de productos naturales muy diversos, denominados péptidos no ribosomales (NRP). Las NRPS son enzimas de gran tamaño, de aproximadamente entre 100 y 1200 kDa, y se encuentran organizadas en módulos, definidos como una parte de la enzima que incorpora de manera específica un aminoácido a la estructura del péptido. Cada módulo a su vez está conformado por tres dominios básicos (Figura 4): el dominio de adenilación (A) que reconoce y activa al aminoácido sustrato, el dominio de tiolación (T) que une covalentemente a través de un enlace tioéster al aminoácido activado y posteriormente lo transfiere al dominio de condensación (C). El dominio C cataliza la formación del enlace peptídico entre el aminoácido previo y el aminoácido del módulo siguiente. Posterior a la etapa de condensación, se cataliza la liberación de la cadena peptídica mediante la actividad del dominio tioesterasa (TE) presente en el último módulo de la enzima. Además de los dominios descritos, los módulos pueden contener dominios accesorios para realizar modificaciones al aminoácido seleccionado (Caulier et al. 2019; Le Govic et al. 2019; Agrawal et al. 2017; Miller et al. 2016).



**Figura 4. Organización de las NRPS.** Éstas se encuentran organizadas por módulos (se ejemplifica una enzima de tres módulos, 1, 2 y 3) los cuales están conformados por dominios básicos de adenilación (A), tiolación (T), condensación (C) y tioesterasa (TE). Cada espiral de color naranja, verde y negra representa un aminoácido diferente, los cuales se van incorporando secuencialmente para producir finalmente un péptido de tres aminoácidos. El número de aminoácidos en el péptido final está determinado por el número de módulos que posee la NRPS.

La síntesis de péptidos no ribosomales avanza en una dirección N-terminal a C-terminal, produciendo péptidos que tienen alrededor de 3-15 aminoácidos de longitud. Los péptidos liberados pueden ser lineales o sufrir modificaciones mediadas por dominios accesorios en los módulos de las NRPS o catalizadas por otras enzimas independientes de manera posterior a su liberación para generar productos cíclicos o cíclicos ramificados. Los módulos de adenilación de las NRPS pueden utilizar tanto aminoácidos canónicos (los 21 aminoácidos del código genético) como no canónicos (Agrawal et al. 2016; Miller et al. 2016).

Se ha reportado que algunos productos sintetizados a partir de las NRPS poseen entre otras actividades funciones antimicrobianas, antivirales, citostáticas, inmunosupresoras, antipalúdicas, antiparasitarias e insecticidas naturales (Agrawal et al. 2016). Por ejemplo, la daptomicina producida por *Streptomyces roseosporus* ha sido aprobada para el tratamiento de infecciones de la piel causadas por patógenos Gram positivo (Barry et al. 2001) y la capreomicina producida por *Saccharothrix mutabilis* posee actividad antibacteriana contra cepas de *Mycobacterium tuberculosis* resistentes a múltiples fármacos (Felnagle et al. 2008).

### **3.6. Otros productos del metabolismo secundario con capacidad antimicrobiana.**

#### **3.6.1. Policétidos**

Los policétidos (PK) son una amplia familia de productos naturales estructuralmente diversos, que incluyen compuestos con actividades antibióticas, inmunosupresoras, anticancerígenas y antiparasitarias, entre otras (Caulier et al. 2019). Estos compuestos son sintetizados por enzimas policétido sintasa (PKS), las cuales se componen de una sucesión de módulos de alargamiento, flanqueados por módulos de iniciación y terminación. Estos módulos están formados por al menos tres dominios, un dominio cetoácido sintasa (KS), una acil transferasa (AT) y un dominio ACP (dominio transportador de grupos acilos). Los PK son compuestos con potencial antimicrobiano. Por ejemplo, la macrolactina, un PK producido por *Bacillus subtilis*, muestra actividad antibacteriana contra *Burkholderia cepacia*, *S. aureus* y *Ralstonia solanacearum*, y actividad antifúngica contra *Fusarium oxysporum* (Caulier et al. 2019).

#### **3.6.2. Betalactonas**

Las betalactonas están dotadas de un sistema de anillos que les confiere a estos compuestos una reactividad única entre los productos naturales. A diferencia de los compuestos  $\beta$ -lactámicos, que se dirigen principalmente a las proteínas de unión a penicilina, las betalactonas son capaces de inhibir una variedad de enzimas como hidrolasas, transferasas, ligasas y oxidorreductasas (Kreitler et al. 2019). Hasta la fecha se han descrito más de 30 productos naturales que contienen  $\beta$ -lactonas, muchos de los cuales poseen potentes propiedades medicinales. Como por ejemplo orlistat, el agente de pérdida de peso inhibidor de lipasas, el cual es actualmente el único fármaco de  $\beta$ -lactona aprobado en el mercado, y la salinosporamida A que tiene un potencial antitumoral y anticancerígeno (Böttcher et al. 2008; Caulier et al. 2019; Robinson y Wackett. 2020).

### **3.6.3. Terpenos**

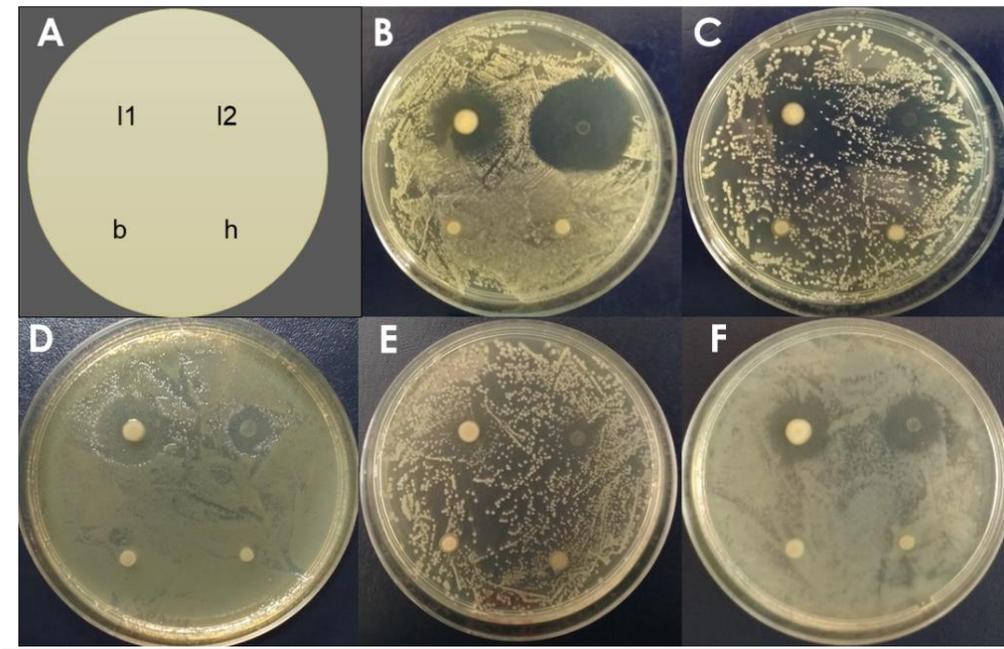
Los terpenos, también conocidos como terpenoides o isoprenoides, se encuentran entre los metabolitos secundarios. Se originan a partir de dos precursores principales: pirofosfato de isopentenilo y el pirofosfato de dimetilalilo (Rudolf et al. 2020). Se ha reportado que los terpenos y sus derivados tienen potencial antibacteriano, antiinflamatorio, antitumorigénico, anticancerígeno, transdérmico y neuroprotector (Caulier et al. 2019). Por ejemplo, la pumilina, un sesquiterpeno aislado de *Bacillus pumilus* posee actividad antibiótica frente a bacterias Gram positivo y actividad citotóxica frente a diferentes líneas celulares tumorales humanas (cerebro, glándula mamaria y pulmón) (Hamdache et al. 2011).

### **3.7. Aislados bacterianos del Río Mapocho con capacidad inhibitoria contra patógenos**

Desde el descubrimiento de la penicilina por Alexander Fleming, los productos naturales derivados de microorganismos ambientales han sido una fuente importante para el desarrollo de agentes antimicrobianos. Las bacterias ambientales son los mayores productores de productos naturales, de las cuales se han descubierto compuestos antibacterianos actualmente en uso, incluida la tetraciclina, cloranfenicol, eritromicina, vancomicina y rifamicina (Milshteyn et al. 2014; Katz et al. 2016). Diversos ecosistemas han sido explorados recientemente en búsqueda de cepas bacterianas con la capacidad de producir nuevas moléculas bioactivas, tales como áreas áridas extremas, cavernas prístinas, rizósferas, diversos ecosistemas marinos y la microbiota humana (Genilloud 2019). En particular, las bacterias marinas y derivadas de agua dulce han sido la fuente más prometedora de productos únicos. Por ejemplo, recientemente se descubrió un nuevo antibiótico, el denominado nucleósido A201A, aislado de un actinomiceto marino capaz de inhibir a MRSA (Basalla et al. 2019).

Un estudio realizado por nuestro grupo de investigación en la Facultad de Medicina de la Universidad de Chile encontró que la microbiota natural del Río Mapocho, en Santiago, Chile, es capaz de inhibir el crecimiento del patógeno *Vibrio cholerae* biovar El Tor. Se aislaron dos microorganismos con capacidad inhibitoria

denominados I1 e I2, a las que se les evaluó la capacidad inhibitoria sobre otros patógenos bacterianos humanos mediante ensayos de co-cultivo en placa. Estos análisis mostraron que I1 e I2 tienen potencial inhibitorio contra *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Vibrio cholerae*, *Escherichia coli* enterotoxigénica y *Pseudomonas aeruginosa* (Figura 5).



**Figura 5. Ensayos de inhibición de los aislados I1 e I2 contra patógenos bacterianos humanos.** A. Posición en la placa de los inhibidores y los controles. En la parte superior de la placa se encuentran el inhibidor I1 e I2 y en la parte inferior b y h, dos microorganismos aislados del Río Mapocho sin capacidad inhibitoria, que fueron utilizados como controles negativos. Se sembró un césped de cada bacteria patógena en placas de LB agar. B. *S. aureus* ATCC 43300; C. *Klebsiella pneumoniae* BAA1705; D. *V. cholerae* N16961; E. *E. coli* enterotoxigénica; F. *Pseudomonas aeruginosa* PAO1. Luego, estas placas fueron inoculadas con 1 microlitro de un cultivo de los inhibidores I1 e I2 y los controles b y h e incubadas por 24 horas. Se puede observar la formación de un halo de inhibición por parte de ambos inhibidores en todas las placas en relación con los controles.

Posteriormente, se realizó la secuenciación del genoma de ambas bacterias. El genoma de la cepa I1 se ensambló en 79 *contigs* con un tamaño total de 6,46 Mpb, mientras que el genoma de I2 se ensambló en 49 *contigs* con un tamaño total de 2,49 Mpb. Posteriormente, se realizó un análisis de identificación mediante la reconstrucción de un árbol filogenético incluyendo los genomas de cepas relacionadas con los aislados I1 e I2 en la plataforma PATRIC con el método de máxima verosimilitud. Con base en dicho análisis, se denominó a las bacterias

inhibidoras como *Pseudomonas* sp. I1 y *Desemzia incerta* I2. La capacidad de estos aislados para inhibir el crecimiento tanto de patógenos Gram positivo como Gram negativo, indica su uso potencial para la producción de compuestos antimicrobianos. Actualmente, se desconocen los genes que codifican a las moléculas inhibidoras producidas hipotéticamente por estos aislados.

A pesar que el mecanismo de inhibición del aislado *Pseudomonas* sp. I1 es desconocido, se han identificado diversas bacteriocinas y compuestos antibacterianos producidos por otros aislados de *Pseudomonas*, y está reportado que algunas cepas tienen la capacidad de suprimir enfermedades causadas por patógenos que afectan a plantas (Lozano et al. 2017).

### **3.8. Bacteriocinas y metabolitos secundarios antimicrobianos producidos por *Pseudomonas*.**

El género *Pseudomonas* está compuesto por alrededor de 200 especies Gram negativo, son microorganismos ubicuos y versátiles, que se caracterizan por su adaptabilidad a muchos ambientes diferentes como el agua, suelo, animales, plantas y humanos. Dentro de este género se encuentran especies como *P. aeruginosa* y *P. fluorescens*, esta última dividida en seis grupos, *P. syringae*, *P. lutea*, *P. putida*, *P. anguilliseptica*, *P. straminea* y *P. koreensis* (Jeong et al. 2018; Babalola et al. 2019; Lin et al. 2016).

En diferentes especies de este género, se ha determinado la presencia de bacteriocinas que se denominan piocinas y según su estructura se pueden clasificar en tres tipos: R, F y S. Las piocinas tipo R y F se asemejan a las proteínas de cola de fago, mientras que las piocinas tipo S son bacteriocinas similares a las colicinas descritas anteriormente. Se han descrito y caracterizado diversas piocinas de tipo S, incluyendo las piocinas S1-9, S11-13 y la AP41 (Ghequire et al. 2014).

Las piocinas, al igual que las colicinas, se encuentran organizadas en dominios. Sin embargo, se postula que las piocinas poseen los dominios de translocación y de unión a receptor en posición intercambiada con respecto a las colicinas. Por lo tanto, las piocinas S poseen un dominio N-terminal de unión a receptor (dominio I), un

dominio central de translocación (dominio II) y un dominio citotóxico C-terminal (dominio III) (Figura 6) (Turano et al. 2020; Sharp et al. 2017). Estas bacteriocinas poseen generalmente una secuencia corta no estructurada de entre 20 a 40 residuos de aminoácidos en el extremo N-terminal, altamente variable entre diferentes piocinas, que está involucrada en el reconocimiento del transportador, promoviendo una reorientación de la piocina, necesaria para el funcionamiento del dominio de translocación (Behrens et al. 2020).



**Figura 6. Organización de dominios de la piocina tipo S.** Dominio de unión a receptor (I); dominio de translocación (II) y dominio citotóxico (III).

Las bacterias del género *Pseudomonas* están dotadas de una amplia diversidad metabólica y fisiológica, lo que les permite colonizar diferentes ambientes. Pueden llegar a persistir de manera eficaz en el agua y el suelo viviendo con un requerimiento nutricional mínimo y tolerando variaciones en las condiciones físicas. Se caracterizan por producir un amplio espectro de metabolitos secundarios (Yan et al. 2018). Se han identificado compuestos antimicrobianos sintetizados por este género bacteriano como la mupirocina aislada de *P. fluorescens* que posee actividad antibacteriana hacia bacterias Gram positivo y ha sido desarrollada como antibiótico tópico. Clínicamente, se utiliza para infecciones estafilocócicas dérmicas, particularmente las causadas por MRSA. Un análisis de *P. koreensis* CI12 aislado de la rizosfera de soja, identificó la producción y secreción de nuevos metabolitos especializados, entre estos la corenceína, perteneciente a una nueva familia de alcaloides de tetrahidropiridina, producida por una vía de PK que inhibe selectivamente el crecimiento de *Flavobacterium johnsoniae* y otros miembros de Bacteroidetes (Lozano et al. 2019). Los metabolitos secundarios de *Pseudomonas* también exhiben otro tipo de bioactividades. Por ejemplo, la safracina aislada de *P. fluorescens* A2-2 posee actividad antitumoral (Gross et al 2007).

Los programas de minería genómica han contribuido de manera eficiente a la identificación de vías con potencial antimicrobiano, como los metabolitos secundarios y bacteriocinas en bacterias.

### **3.9. Minería Genómica**

La minería genómica es una herramienta que ha facilitado de manera significativa la identificación de vías de producción de compuestos inhibidores en aislados microbianos (Ward et al. 2018; Ziemert et al. 2016). Esta herramienta permite la búsqueda de BGC de metabolitos secundarios como PK, NRPS, péptidos de síntesis ribosomal y modificados postraduccionalmente (RiPPs), bacteriocinas, aminoglucósidos, entre varios otros, en los genomas de los microorganismos con potencial inhibitorio. Para esto, se emplean programas que realizan comparaciones de secuencias de los genomas con bases de datos de genes de producción de inhibidores ya caracterizados experimentalmente. En un principio la minería genómica se realizaba manualmente con BLAST o HMMer, pero se fueron desarrollando sistemas automatizados que agilizan el enfoque, como la plataforma DECEIPHER®. De manera más reciente, se han desarrollado herramientas adicionales que pueden ser utilizadas de forma gratuita y con interfase amigable, incluidos BAGEL4, CLUSEAN y antiSMASH, siendo esta última la plataforma más completa para realizar este análisis (Ziemert et al. 2016; Sélem-Mojica et al. 2019). Estos métodos no solo buscan genes individuales sino también secuencias parciales o *clusters* completos de BGC. Esta función está incluida en el software antiSMASH que permite comparar un BGC del genoma de interés con una colección actualizada de BGC, caracterizados en otros microorganismos en la base de datos seleccionada de MiBIG (información mínima sobre grupos de genes biosintéticos), dando información sobre vías similares presentes en el genoma de interés (Ziemert et al. 2016).

El programa antiSMASH permite identificar alrededor de 60 tipos de *clusters* diferentes que codifican para metabolitos especializados como vías con enzimas NRPS, PKS, bacteriocinas y terpenos, entre otros. Por otro lado, BAGEL4 es una herramienta que permite identificar rápidamente posibles grupos de genes que

codifican bacteriocinas (de Jong et al. 2006). Existe una gran cantidad de información respecto a los metabolitos y proteínas con características bioactivas producidas por microorganismos. Al realizar minería genómica sobre un genoma de un aislado nuevo, es probable encontrar *clusters* genéticos cuyos productos ya son conocidos y han sido caracterizados en otras cepas o especies microbianas. Por lo tanto, usualmente, una vez obtenidos los resultados de minería genómica para un organismo en particular, se aplica un análisis de desreplicación. Éste consiste en comparar las predicciones contra bases de datos para identificar BGCs con potencial para producir moléculas nuevas y separar lo ya conocido de lo nuevo (Begrem et al. 2020).

Utilizando minería genómica se han descubierto diversas vías que sintetizan metabolitos especializados con actividad biológica de interés, como por ejemplo, las NRPS producidas por *Pseudomonas fluorescens* Pf-5 que sintetizan orfamida A con efecto antibacteriano (Gross et al. 2007). Otro ejemplo es la salinosporamida K que tiene un efecto antitumoral, sintetizado por una NRPS identificada en *Salinispora pacifica* (Nikolouli et al. 2012).

El aislamiento previo de *Pseudomonas* sp. I1 como parte de la microbiota del Río Mapocho con potencial de inhibir patógenos bacterianos humanos es relevante, sin embargo, sus determinantes de inhibición son completamente desconocidos. Por lo tanto, en este estudio se aplicaron técnicas de minería genómica para identificar en el genoma de *Pseudomonas* sp. I1 *clusters* de genes involucrados en la capacidad de inhibición de patógenos microbianos. Este estudio se centró en la identificación bioinformática y caracterización *in silico* de probables determinantes genéticos cuyos productos permitan explicar el fenotipo de inhibición de crecimiento bacteriano de *Pseudomonas* sp. I1.

#### **4. HIPÓTESIS**

La bacteria *Pseudomonas* sp. I1 posee en su genoma genes que codifican proteínas responsables de la síntesis de metabolitos especializados y bacteriocinas con potencial de inhibición del crecimiento de patógenos microbianos.

#### **5. OBJETIVOS**

##### **5.1. OBJETIVO GENERAL**

Identificar y caracterizar genes que codifican proteínas para la síntesis de metabolitos especializados y bacteriocinas de *Pseudomonas* sp. I1 que se asocian a la capacidad de inhibición del crecimiento de patógenos microbianos.

##### **5.2. OBJETIVOS ESPECÍFICOS**

**5.2.1. Objetivo específico 1:** Identificar in silico los genes cuyos productos se involucran en la síntesis de metabolitos especializados y bacteriocinas en el genoma de *Pseudomonas* sp. I1

**5.2.2. Objetivo específico 2:** Caracterizar los *clusters* para los probables determinantes de inhibición presentes en *Pseudomonas* sp. I1 mediante comparaciones con genes reportados en bases de datos.

## 6. MATERIALES Y MÉTODOS

### 6.1. Identificación de *clusters* genéticos que codifican para la biosíntesis de metabolitos especializados y bacteriocinas en *Pseudomonas* sp. I1.

Para dar cumplimiento a los objetivos específicos previamente planteados, se utilizaron los siguientes materiales y métodos.

El genoma de *Pseudomonas* sp. I1, secuenciado en nuestro grupo, se depositó en la plataforma PATRIC (<https://www.patricbrc.org/>). Para la búsqueda de genes cuyos productos sintetizan metabolitos especializados, los *contigs* del genoma se obtuvieron de la plataforma PATRIC y se analizaron utilizando antiSMASH (<http://antismash.secondarymetabolites.org>), que tiene como función la búsqueda de BGCs involucrados hipotéticamente en la biosíntesis de metabolitos especializados y bacteriocinas mediante la comparación de los genes de las cepas en estudio con una extensa colección de BGC caracterizados de otros microorganismos en la base de datos MiBIG. Además, antiSMASH aplica sobre los posibles *hits* positivos algunos análisis complementarios específicos, por ejemplo, búsqueda de dominios en módulos de enzimas NRPS. También, con el objeto de identificar genes relacionados con síntesis de bacteriocinas en el genoma de *Pseudomonas* sp. I1, se utilizó el software BAGEL4 (<http://bagel4.molgenrug.nl/>), que funciona mediante búsqueda de identidad con genes de bacteriocinas en bases de datos.

Los resultados obtenidos antiSMASH y BAGEL4 se descargaron en formato SVG, para posteriormente ser editados en el programa de imágenes Inkscape versión 1.0.2. Estos análisis se realizaron en un computador básico, ya que dichos programas no requieren un computador especial o de mayor capacidad.

### 6.2. Análisis de los *clusters* obtenidos por antiSMASH y BAGEL4

Una vez obtenidos los resultados de la minería genómica, se realizaron análisis de los *clusters* obtenidos utilizando programas bioinformáticos. El programa BioEdit versión 7.2, se utilizó para crear una base de datos local de bacteriocinas, ensamblajes manuales y desde este programa se realizaron alineamientos

mediante el método *optima GLOBAL alignment*. El programa Jalview se usó para realizar alineamientos múltiples mediante la herramienta ClustalO.

Para la predicción de estructuras secundarias y la búsqueda de regiones *coiled coil*, la secuencia primaria de las piocinas indicadas se analizó utilizando los programas JPred4 (<http://www.compbio.dundee.ac.uk/jpred/>) y COILS del servidor ExPASy ([https://embnet.vital-it.ch/software/COILS\\_form.html](https://embnet.vital-it.ch/software/COILS_form.html)), respectivamente, sin realizar modificaciones de los parámetros que poseen ambas aplicaciones por defecto.

También se utilizó la base de datos del Centro Nacional de Información Biotecnológica (NCBI) (<https://www.ncbi.nlm.nih.gov/>), para búsqueda de homólogos mediante análisis de secuencia utilizando BLASTp.

La búsqueda de los motivos nucleasa conservados en las bacteriocinas DNasa HNH y no HNH según lo reportado por Sharp et al. (2017) se llevó a cabo mediante inspección manual y alineamientos con el dominio DNasa de la colicina E9 (representativa de DNasa HNH) y la secuencia del extremo carboxilo terminal de la piocina de *Pseudomonas* proporcionada por Sharp et al. (2017) representativa de DNasas no HNH (WP\_016394378.1). Además, se realizaron predicciones de estructura secundaria de los dominios del extremo carboxilo terminal de las bacteriocinas utilizando el programa Jpred4, para comparar las estructuras secundarias hipotéticas de dichos dominios DNasa.

## 7. RESULTADOS

**Objetivo específico 1:** Identificar *in silico* los genes cuyos productos se involucran en la síntesis de metabolitos especializados y bacteriocinas en el genoma de *Pseudomonas* sp. I1

### 7.1. Análisis del genoma de la bacteria inhibidora *Pseudomonas* sp. I1 mediante minería genómica.

La minería genómica permite predecir *clusters* de genes hipotéticamente involucrados en la producción de compuestos bioactivos. Para estos análisis se han desarrollado diversos programas bioinformáticos. El programa antiSMASH es la herramienta de minería genómica más utilizada y permite identificar BGCs en el genoma de interés mediante una comparación de los genes de las cepas en estudio con una base de datos de genes para vías de producción de metabolitos y bacteriocinas conocidas. Por otro lado, BAGEL es un programa que permite identificar específicamente *clusters* de genes para la síntesis de bacteriocinas mediante búsqueda de identidad con genes de bacteriocinas en bases de datos. Para el análisis del genoma de *Pseudomonas* sp. I1 se decidió utilizar estos dos programas.

Mediante el análisis de los *contigs* del ensamblado del genoma de *Pseudomonas* sp. I1 empleando antiSMASH para la búsqueda de *clusters* de metabolitos especializados y bacteriocinas, se localizaron cinco *clusters* de genes cuyos productos corresponderían a vías tipo NRPS de diferentes tamaños. Además, otro resultado del análisis mediante antiSMASH señaló que las vías más parecidas en la base de datos MiBIG para los *clusters* con vías de tipo NRPS de *Pseudomonas* sp. I1 corresponden a las vías de biosíntesis de lokisina, nunamicina, cichopectina y pioverdina (2 *clusters*). Estos *clusters* se encuentran formando parte de los *contigs* 0034, 0039, 0014, 0002 y 0006, respectivamente (Tabla 1). En el análisis se observa que para el caso de lokisina y nunamicina, todo el *cluster* corresponde al tamaño del *contig*. Adicionalmente, se identificó una vía hipotética para la producción de una betalactona. Los resultados de antiSMASH indican similitud de esta vía con la vía

de producción de fengicina. Por otro lado, el análisis utilizando BAGEL4 permitió identificar la presencia de tres *clusters* que codifican bacteriocinas hipotéticas en el genoma (Tabla 1). Las vías encontradas en el genoma de *Pseudomonas* sp. I1 son de tres tipos, las cuales se caracterizan por el potencial antimicrobiano que poseen.

**Tabla 1. Clusters biosintéticos de metabolitos secundarios y bacteriocinas en el genoma de *Pseudomonas* sp. I1 identificados mediante antiSMASH y BAGEL4.**

Plataforma de análisis	Tipo	Producto hipotético*	Longitud de BGC (pb)	Longitud de los clusters (pb)	Logitud (pb) y (número de los contigs)
antiSMASH	NRPS	Lokisina	50122	23909	23909 (0034)
		Nunamicina	115071	5484	5484 (0039)
		Cichopectina	84760	31755	150087 (0014)
		Pioverdina	175054	32500	386174 (002)
		Pioverdina	175054	63562	628500 (006)
	Betalactona	Fengicina	49542	23263	231209 (009)
BAGEL4	Bacteriocinas	Carocin_D	N/A	2105	823227 (001)
		Colicina_E9	N/A	1654	131845 (0016)
		Colicina	N/A	1406	30722 (0029)

\*Los productos hipotéticos corresponden a los productos de las BGC más parecidas caracterizadas y reportadas en bases de datos, según los resultados indicados por los programas. N/A: No disponible ¿Qué significa N/A)?

**Objetivo específico 2:** Caracterizar los *clusters* para los probables determinantes de inhibición presentes en *Pseudomonas* sp. I1 mediante comparaciones con genes reportados en bases de datos.

## 7.2. Identificación *in silico* de genes para la biosíntesis de lokisina en *Pseudomonas* sp. I1

Para cada una de las vías de biosíntesis de metabolitos secundarios de tipo NRPS, antiSMASH señala una vía similar en la base de datos MiBIG. En la Figura 7A se muestran los resultados para el primero de estos *clusters*. En la parte superior de ésta, se esquematiza el primer *cluster* tipo NRPS encontrado en el genoma de *Pseudomonas* sp. I1. Según los resultados de antiSMASH, este *cluster* contendría una vía metabólica conformada por dos enzimas *core* tipo NRPS, cuyos genes están indicados con una barra verde y coloreados de color rojo vino de acuerdo al código

de colores relativo a la función hipotética de los genes del *cluster* indicado en la figura. Además, este *cluster* también incluye un gen probablemente asociado al transporte del metabolito producido, indicado en color azul y localizado en la región corriente arriba de los dos genes para NRPS *core*. En color gris, antiSMASH indica la presencia de otros genes en esta región genética, pero que probablemente no estarían involucrados en la biosíntesis del metabolito secundario. En la parte de abajo de la Figura 7A se muestra el resultado de una comparación realizada automáticamente por antiSMASH del *cluster* encontrado en *Pseudomonas* sp. I1 con el *cluster* más parecido en la base de datos, en este caso correspondiente al BCG0001980, el cual comprende un *cluster* de producción de lokisina en *Pseudomonas* sp. COR10. En este análisis la secuencia problema corresponde al *cluster* de *Pseudomonas* sp. I1, y los genes conservados entre ella y el *cluster* BCG0001980 se encuentran coloreados del mismo color. Los resultados indican que el *cluster* de *Pseudomonas* sp. I1 conserva el 50% de los genes requeridos para la producción de lokisina de la vía modelo en el BCG0001980. Este *cluster* modelo de biosíntesis de lokisina contiene tres enzimas NRPS, de las cuales dos están conservadas en el *cluster* de la secuencia problema. Además, los cuatro primeros genes codificados corriente arriba y de manera divergente a los genes para NRPS, también tienen homología con genes de la vía de biosíntesis de lokisina. Entre estos genes está incluido el gen originalmente identificado por antiSMASH con funciones transportadoras. Para los otros tres genes, antiSMASH no había asignado originalmente una función hipotética dentro de la vía detectada en *Pseudomonas* sp. I1, pero su conservación en la vía modelo de lokisina sugiere que podrían estar involucrados en la biosíntesis del metabolito secundario producido por este *cluster*. De manera importante, esta comparación indica que el *cluster* de *Pseudomonas* sp. I1 carece del tercer gen de NRPS conservado en el *cluster* de lokisina modelo (Figura 7A, parte baja). En el genoma de *Pseudomonas* sp. I1, este *cluster* se encuentra localizado en el *contig* 0034. Al observar los resultados de la predicción original de antiSMASH, se detectó que el *cluster* está localizado en el extremo terminal 5' del *contig*. Por lo tanto, es posible que este *cluster* no esté completamente representado en el *contig*. El gen de NRPS faltante

se localiza en el extremo 5' del *cluster*, el cual correspondería al término truncado del *contig*. Por lo tanto, este gen podría estar presente en el genoma original pero no haberse ensamblado en el mismo *contig*.

De manera análoga al resultado anterior, en la parte superior de la Figura 7B se muestra el segundo *cluster* de tipo NRPS detectado por antiSMASH en el genoma de *Pseudomonas* sp. I1, mientras que en la parte inferior de la figura se muestra una comparación entre este *cluster* y el BCG más parecido, en este caso el BCG001416, con la vía de biosíntesis de nunamicina de *Pseudomonas* sp. In5. De acuerdo a los resultados, este *cluster* estaría conformado por un gen *core* para una enzima NRPS y un gen para una enzima accesoria. Sin embargo, en la comparación se muestra que éstos representan solo un 35% de los genes requeridos para la biosíntesis de nunamicina. Además, este *cluster* está localizado en el *contig* 0039, el cual es un *contig* relativamente pequeño de 5484 pb (Tabla 1). Por lo tanto, también en este caso es posible que el *cluster* identificado por antiSMASH represente un fragmento truncado de una vía metabólica.

En la Figura 7C, se muestra el tercer *cluster* de NRPS identificado por antiSMASH. Según los resultados de la comparación con el BGC más parecido en las bases de datos mostrado en la parte inferior de la figura, este *cluster* posee un 46% de genes homólogos a los de la vía metabólica modelo, el BGC0001389, que contiene los genes de la vía de biosíntesis de cichopectina de *Pseudomonas cichorii*. De manera similar al *cluster* para la lokisina, este *cluster* está localizado en el límite de un *contig* (0014) y parece estar truncado.

Colectivamente, los resultados de los tres *clusters* de NRPS identificados en *Pseudomonas* sp. I1 con similitud con los BGC de lokisina, nunamicina y cichopectina sugieren que dichos *clusters* no serían independientes, sino que podrían corresponder a fragmentos de una misma vía, localizados en diferentes *contigs*. Con base en estas observaciones se realizó un ensamblaje manual de la secuencia de nucleótidos de los tres *contigs*, ordenando éstos en todas las combinaciones posibles, y un posterior alineamiento de los ensamblajes con las vías

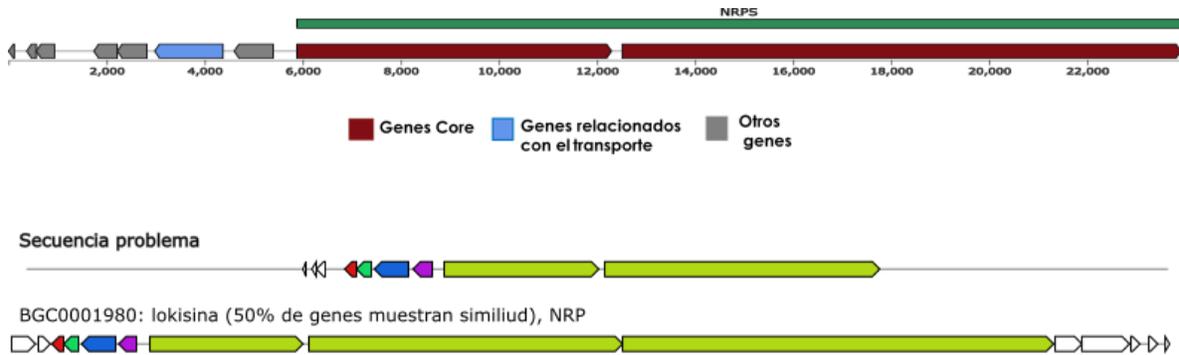
caracterizadas más parecidas indicadas por antiSMASH (lokisina, nunamicina y cichopectina), utilizando el programa BioEdit.

El ensamblado de los tres *clusters* en el orden 0034-0039-0014 conserva un 84,5% de identidad con el *cluster* completo de genes biosintéticos de la vía modelo de lokisina de *Pseudomonas* sp. COR10 (Figura 7D), contiene el orden genético y la mayoría de la secuencia de la vía modelo de biosíntesis de lokisina, excepto dos segmentos pequeños de 1101 pb, correspondientes a partes de los genes *lokB* y *lokC*. En el ensamblado, estas regiones estarían situadas exactamente en ambos extremos del *contig* interno (0039) y se denominaron secuencia 1 y secuencia 2 (Figura 7D). Un alineamiento de estas dos secuencias nucleotídicas determinó que comparten un 91% de identidad entre sí, por lo tanto, son secuencias repetidas.

Para determinar si estas secuencias están representadas en otros *contigs* de *Pseudomonas* sp. I1 no identificados por antiSMASH, se realizó un análisis de estas secuencias de 1101 pb contra los *contigs* del genoma mediante BLASTN. De este modo, se determinó que una copia de estas secuencias se encuentra representada en los *contigs* pequeños 0054 (612nt) y 0068 (384nt) del genoma de *Pseudomonas* sp. I1 (Figura 7E). Por lo tanto, los resultados de este análisis sugieren que los tres *clusters* identificados por antiSMASH como de biosíntesis de lokisina, nunamicina y cichopectina conforman un solo *cluster* para la biosíntesis de lokisina en el genoma de *Pseudomonas* sp. I1.

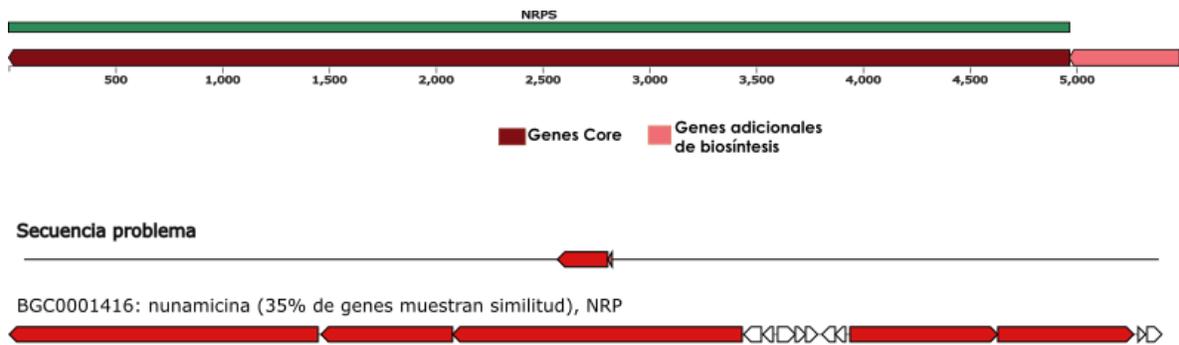
A

### Cluster del contig 0034



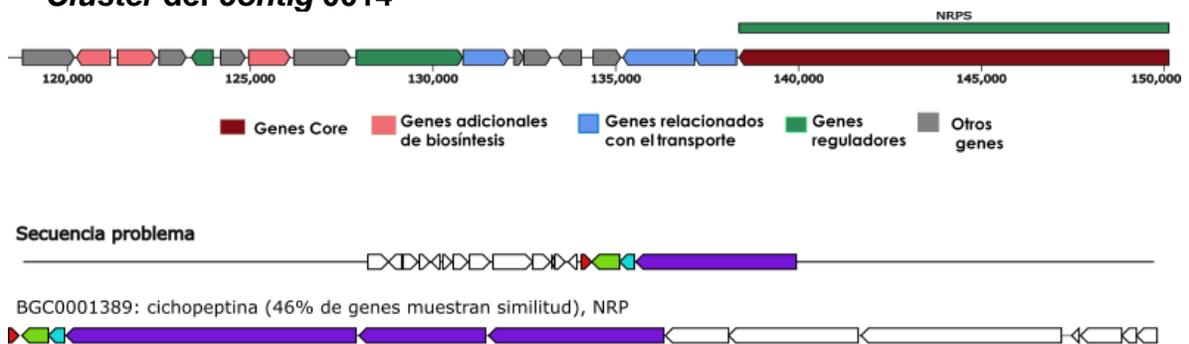
B

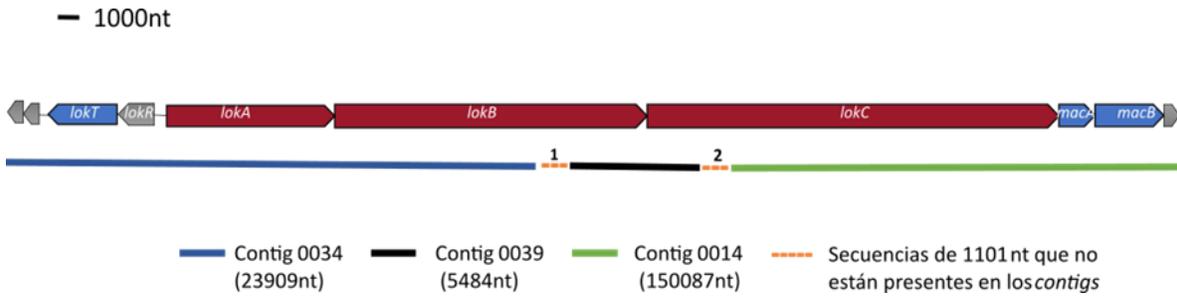
### Cluster del contig 0039



C

### Cluster del contig 0014



**D****E**

**Figura 7. Clusters identificados en *Pseudomonas* sp. I1 por antiSMASH para la síntesis de lokisina. A.** En la parte superior se encuentra el *cluster* del *contig* 0034, que está conformado por genes *core* (rojo vino), genes relacionados con el transporte (azul) y otros genes (gris). En la parte inferior se muestra la secuencia problema y la vía modelo lokisina con su respectivo porcentaje de similitud de genes. **B.** Se puede observar en la parte superior el *cluster* del *contig* 0039, el cual contiene gen *core* (rojo vino) y genes accesorios (rosa). En la parte inferior se observa la vía de numamicina que tiene genes similares a la secuencia problema. **C.** El *cluster* del *contig* 0014 está conformado por genes *core* (rojo vino), genes accesorios (rosa), genes relacionados con el transporte (azul), genes reguladores (verde y otros genes (gris), en la parte inferior se observa la vía de cichopectina con un porcentaje de similitud con la secuencia problema. El código de colores entre la secuencia problema y la vía de lokisina hace referencia a los genes similares entre ambas. **D.** *Cluster* de genes para la síntesis de lokisina de *Pseudomonas* sp. COR10 conformado por los siguientes genes: *lokA*, *lokB*, *lokC* (vino tinto); *macA*, *macB* y *lokT* (azul) y *lokR* (gris). En la parte inferior, ensamblaje de los *contigs* 0034 (línea azul), 0039 (línea negra) y 0014 (línea verde) de *Pseudomonas* sp. I1. Dos secuencias de 1101 nt de la vía modelo (líneas entrecortadas color naranja 1 y 2) no están presentes en los *contigs*. **E.** Representación del alineamiento de las dos secuencias de 1101 nt (líneas naranjas), con los *contigs* del genoma de *Pseudomonas* sp. I1. Un fragmento de la secuencia de 1101 forma parte del *contig* 0054 (línea roja) y otro fragmento del *contig* 0068 (línea morada).

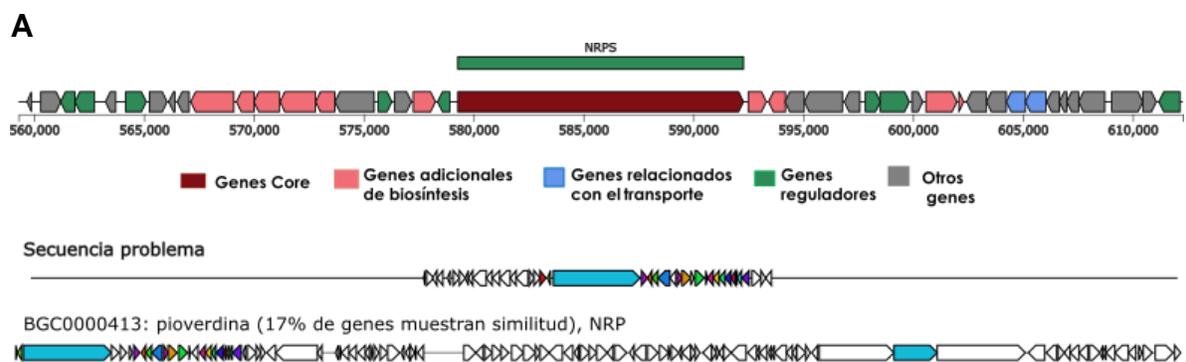
### 7.3. Identificación *in silico* de genes para la biosíntesis de pioverdina en *Pseudomonas* sp. I1

Las pioverdinas son sideróforos fluorescentes producidas por la mayoría de las especies de *Pseudomonas*, sintetizadas a partir de NRPS. La cadena de aminoácidos inicial sintetizada por las enzimas NRPS involucradas en esta vía se denomina ferribactina y es el sustrato precursor de enzimas modificadoras que proveen la estructura final de la pioverdina. La estructura básica de una pioverdina está sintetizada inicialmente por PvdL, que es una NRPS esencial. PvdL incorpora siempre los mismos tres aminoácidos a la estructura: ácido L-glutámico (L-Glu), D-tirosina (D-Tyr) y diamino butirato (L-Dab), para la biosíntesis de cromóforo, que corresponde a la primera parte de la ferribactina. Además, otras NRPS como PvdI y PvdD, son responsables de la síntesis del resto del péptido. Sin embargo, a diferencia de PvdL, estas NRPS adicionales son variables entre diferentes cepas de *Pseudomonas* y aportan secuencias de aminoácidos distintos (Ringel y Brüssel. 2018; Moon et al. 2008).

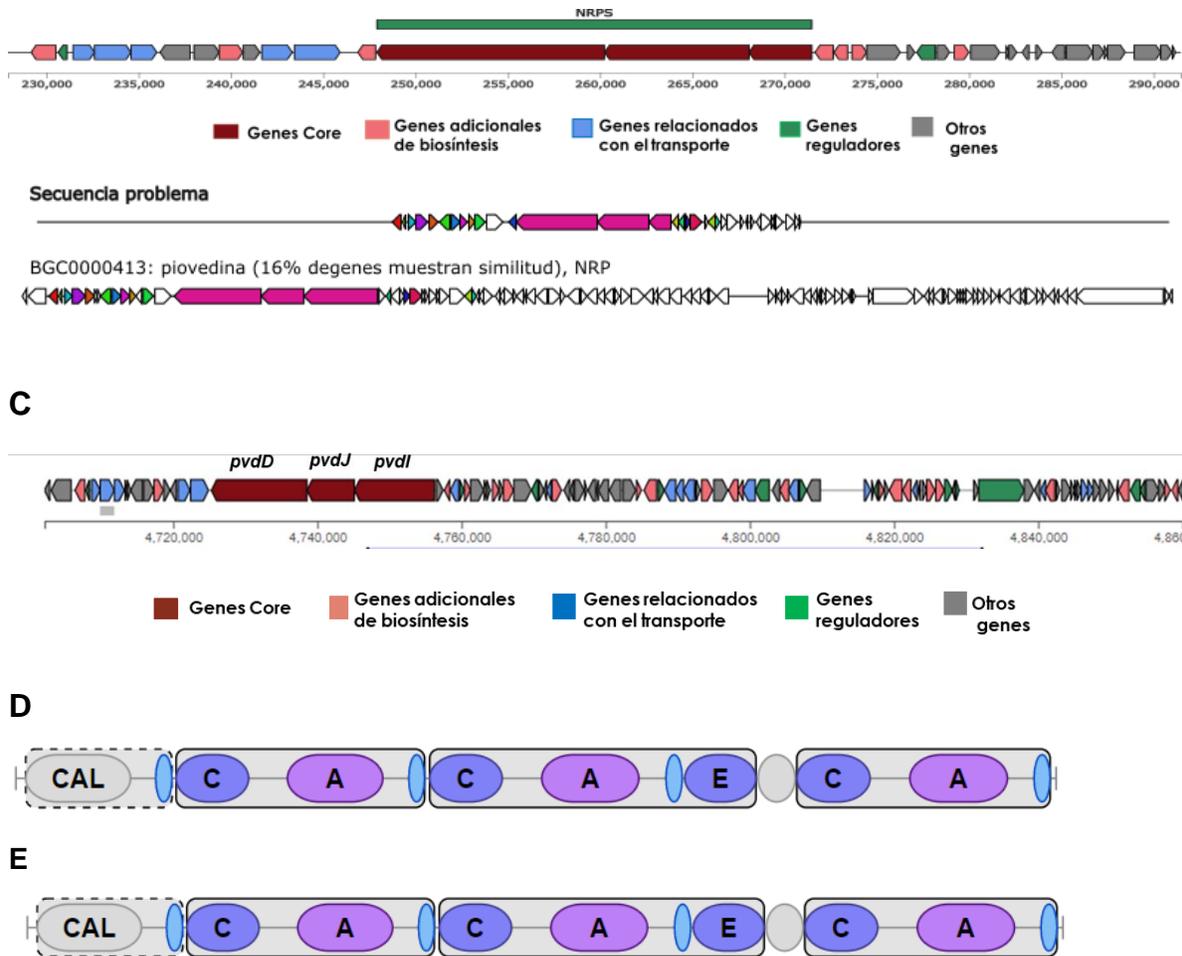
Los dos *clusters* cuyos productos sintetizan hipotéticamente pioverdina y que forman parte de las vías tipo NRPS identificadas en el genoma de *Pseudomonas* sp. I1 por antiSMASH (Figura 8A y B), también fueron analizados. Para ambos *clusters*, antiSMASH indica como la vía biosintética más parecida a la contenida en el BGC0000413, que corresponde a la vía de biosíntesis de pioverdina de *P. protegens* Pf-5. En dicha vía modelo, los *clusters* de NRPS se encuentran cercanos en el genoma y están conformados por genes para las NRPS PvdL, PvdI, PvdJ y PvdD, además de otros genes secundarios (Figura 8C). Uno de los *clusters* de las vías para pioverdina de *Pseudomonas* sp. I1 posee una NRPS (Figura 8A) con un tamaño similar al de la PvdL de la vía modelo de pioverdina de *P. protegens* Pf-5, se realizó un alineamiento entre éstas utilizando el programa BioEdit y se determinó que poseen una identidad de secuencia de aminoácidos de 84% y una similitud de 90%. Entre los resultados entregados por antiSMASH para este *cluster*, se incluyó una predicción de módulos y dominios de las enzimas NRPS. Esta información indica que la NRPS que correspondería hipotéticamente a PvdL en el *cluster* de

*Pseudomonas* sp. I1 está conformada por tres módulos con sus respectivos dominios básicos precedidos por un dominio accesorio coenzima A ligasa (CAL) (Figura 8D). Este arreglo de dominios es idéntico al presentado por la enzima PvdL *bona fide* de la vía modelo de biosíntesis de pioverdina (Figura 8E). Estos datos sugieren que este *cluster* contiene un gen para la enzima *core* de biosíntesis de pioverdina PvdL.

El segundo *cluster* para pioverdina en *Pseudomonas* sp. I1 (Figura 8B, parte superior) conserva tres genes para NRPS y posee identidad genética con el resto de la secuencia modelo no cubierto por el primer *cluster* (Figura 8B, parte inferior). Al comparar las tres NRPS que se encuentran en el *cluster* de *Pseudomonas* sp. I1 (Figura 8B) con las tres NRPS que conforman la vía modelo, se observa diferencias en cuanto a tamaño, lo cual es esperable ya que estas NRPS son generalmente variables entre cepas. A diferencia de los *cluster* de la vía modelo, los dos *clusters* identificados en *Pseudomonas* sp. I1 se encuentran lejanos uno del otro en el genoma. Con base en las longitudes totales de los *contigs* en los que se encuentran ambos *clusters*, se determinó que los *clusters* de biosíntesis de pioverdina en el genoma de *Pseudomonas* sp. I1 existe una distancia mínima de 134,254pb. Esta característica también se observa en otras especies de *Pseudomonas* (Moon et al. 2008). En conjunto, este análisis sugiere que los dos *clusters* para pioverdina de *Pseudomonas* sp. I1 identificados por antiSMASH forman una sola vía para producción de pioverdina, y poseen los genes NRPS esenciales para su biosíntesis.



**B**

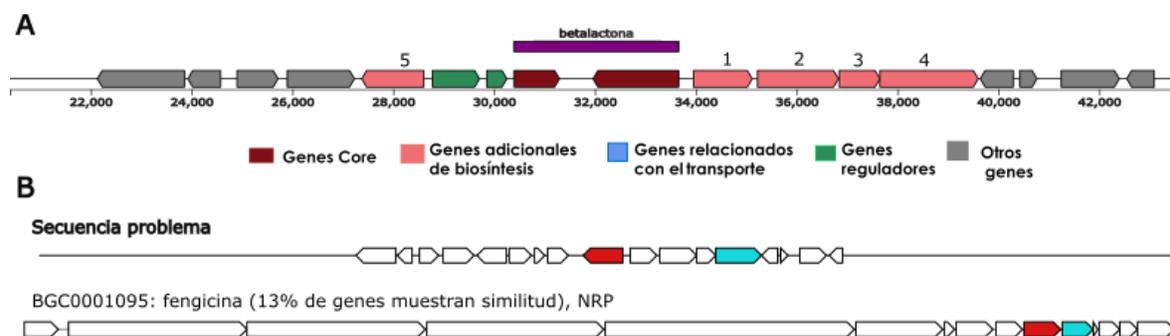


**Figura 8. Cluster de vías NRPS que codifican para pioverdina. A y B.** Clusters de vías NRPS que codifican para pioverdina en *Pseudomonas* sp. I1. Se puede observar los genes *core* (rojo vino), genes accesorios (rosa), genes relacionados con el transporte (azul), genes reguladores (verde) y otros genes (gris). **C** Cluster de genes que codifican proteínas responsables de la biosíntesis de pioverdina de *P. protegens* Pf-5, el BGC caracterizado más parecido a los clusters para pioverdina en *Pseudomonas* sp. I1, con sus respectivos genes principales (color rojo vino) *pvdD*, *pvdJ*, *pvdI* y *pvdL*. **D.** Esquematización de la predicción de módulos y dominios por antiSMASH presentes en la NRPS hipotética correspondiente a PvdL en el primer cluster para pioverdina en *Pseudomonas* sp. I1. Dominios: Adenilación (A), condensación (C), epimerización (E) y Coenzima A ligasa (CAL). **E.** Predicción de módulos y dominios presentes en la enzima PvdL del BGC modelo para biosíntesis de pioverdina en *P. protegens* pf-5.

#### 7.4. Identificación de una vía que codifica hipotéticamente para betalactona

Para el análisis del cluster identificado por antiSMASH como una vía hipotética para la biosíntesis de betalactona, se realizó una búsqueda bibliográfica de la conformación general de los clusters para vías de producción de betalactonas ya caracterizadas y se identificó el tipo de enzimas codificadas en dichos clusters. El

*cluster* identificado por antiSMASH está conformado por dos genes biosintéticos *core*, cuatro genes biosintéticos accesorios y dos genes con función de regulación (Figura 9A). Los dos genes *core* codifican productos que poseen actividad 2-isopropilmalato sintasa y sintetasa dependiente de AMP y ligasa, mientras que los genes accesorios codifican para acil-CoA deshidrogenasa (gen rosa 1 y 5), enoil-CoA hidratasa (gen rosa 3), acetil-CoA carboxilasa (gen rosa 2 y 4) (Figura 9A). Este tipo de actividades enzimáticas, relacionadas con metabolismo de ácidos grasos, están presentes en algunos *clusters* para vías de producción de betalactonas caracterizadas (Koumoutsis et al. 2004). El BGC más parecido a este *cluster* en MiBIG es el BGC0001095, el cual codifica la vía de biosíntesis de fengicina de *Bacillus velezensis* FZB42 (Figura 9B). Sin embargo, el *cluster* de *Pseudomonas* sp. I1 solo conserva 13% de genes similares a los de la vía modelo. Incluso, la vía modelo incluye cinco genes para NRPS para la biosíntesis de fengicina, los cuales están ausentes en el *cluster* de *Pseudomonas* sp. I1. El bajo porcentaje de identidad genética entre los *clusters* y la ausencia de los genes para estas enzimas clave sugieren que la vía del *cluster* de la cepa inhibidora estaría sintetizando probablemente un producto diferente a fengicina.

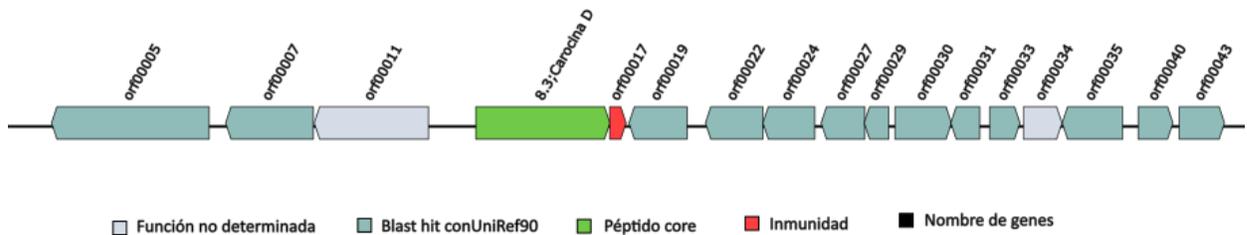


**Figura 9. Cluster identificado por antiSMASH para la biosíntesis de una betalactona en *Pseudomonas* sp. I1.** A El *cluster* contiene genes *core* (vino tinto), genes accesorios representados por los números 1, 2, 3, 4 y 5 (rosa), genes reguladores (verde) y otros genes (gris). secuencia problema y la secuencia de la vía modelo correspondiente a biosíntesis de fengicina indicada por antiSMASH. Los genes coloreados (rojo y celeste), son genes que mantienen similitud entre las dos secuencias.

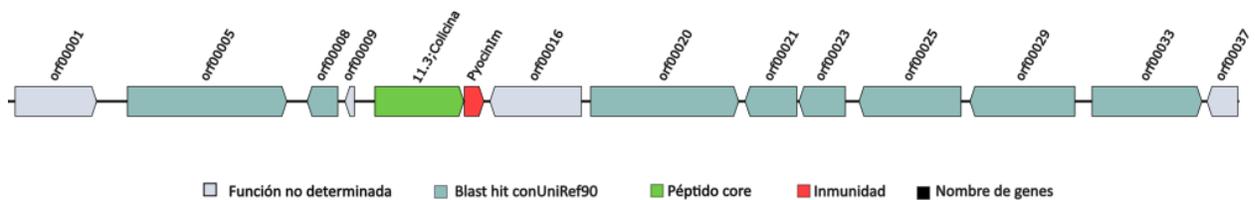
## 7.5. Caracterización de las bacteriocinas identificadas en el genoma de *Pseudomonas* sp. I1

Mediante la inspección con BAGEL4, se identificaron tres *clusters* que codifican para pocias tipo S hipotéticas en *Pseudomonas* sp. I1. Las pocias S, en general, son proteínas que poseen similitud con las colicinas, bacteriocinas multimodulares de *E. coli*. Inicialmente, las tres bacteriocinas identificadas en *Pseudomonas* sp. I1 fueron denominadas por BAGEL4 como carocina D, colicina y colicina E9 (Figura 10).

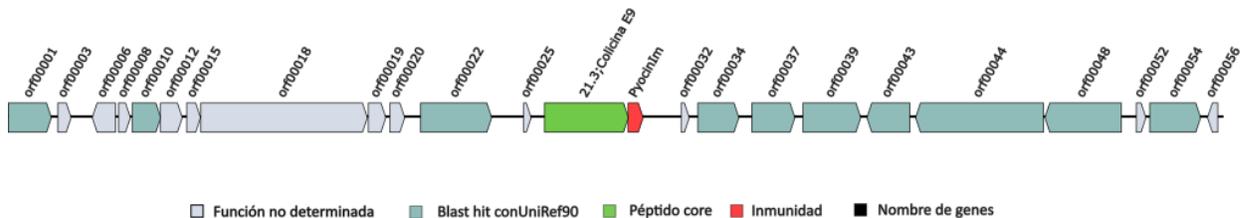
### Carocina D



### Colicina



### Colicina E9



**Figura 10. Genes de bacteriocinas identificados en el genoma de *Pseudomonas* sp. I1 por BAGEL4.** Se muestran los loci con los genes para las tres bacteriocinas. En cada caso se indica el péptido *core* que codifica para la bacteriocina (color verde) y su respectivo gen de inmunidad (color rojo).

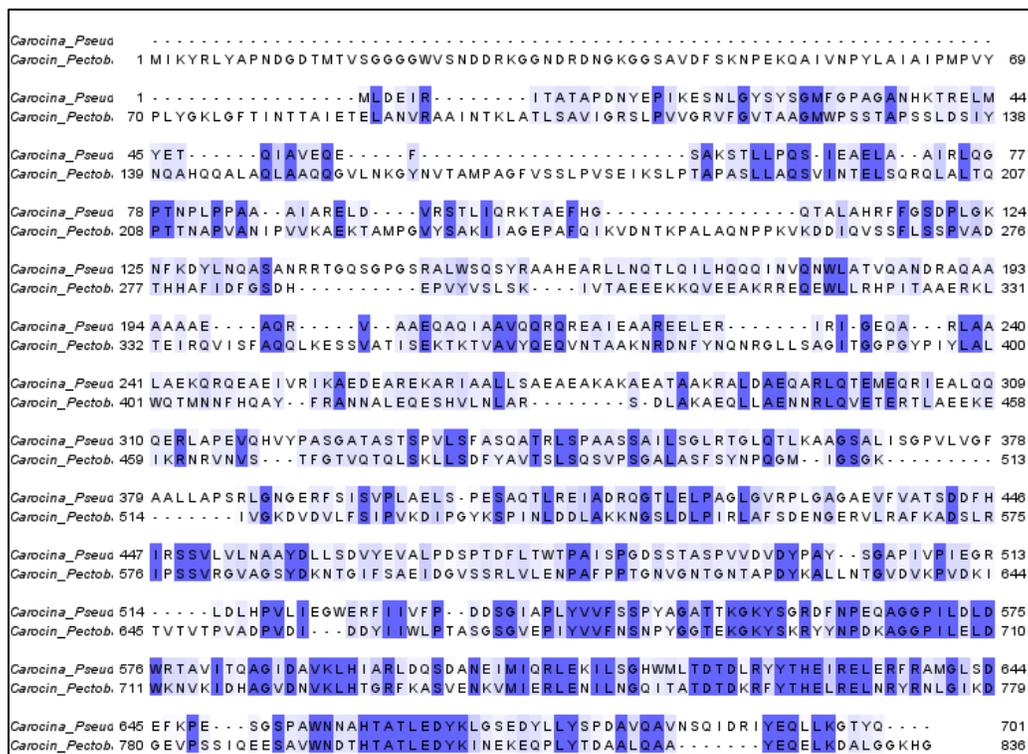
Para caracterizar estas bacteriocinas hipotéticas, se creó una base de datos con las secuencias de las colicinas y piocinas caracterizadas y reportadas en bibliografía. Las secuencias aminoacídicas de dichas bacteriocinas se obtuvieron desde la base de datos del NCBI y se utilizaron para crear una base de datos local en BioEdit. Las bacteriocinas incluidas en esta base de datos se encuentran incluidas en la Tabla 2.

**Tabla 2. Bacteriocinas incluidas en la base de datos local.**

Número de acceso en GenBank	Especie	Bacteriocina
ADH95192.1	<i>Pectobacterium carotovorum</i> subsp. <i>carotovorum</i> PCC21	Carocin D
Q06583.2	<i>Pseudomonas aeruginosa</i>	Piocina S1
PUA09355.1	<i>Pseudomonas aeruginosa</i>	Piocina S2
WP_012075804.1	<i>Pseudomonas aeruginosa</i>	Piocina S3
WP_003148085.1	<i>Pseudomonas aeruginosa</i>	Piocina S4
WP_003115311.1	<i>Pseudomonas aeruginosa</i>	Piocina S5
ERW54179.1	<i>Pseudomonas aeruginosa</i>	Piocina S7
APW77308.1	<i>Pseudomonas aeruginosa</i>	Piocina S8
ERV77365.1	<i>Pseudomonas aeruginosa</i>	Piocina S9
CAW28924.1	<i>Pseudomonas aeruginosa</i>	Piocina S11
ERY67097.1	<i>Pseudomonas aeruginosa</i>	Piocina S12
WP_023116694.1	<i>Pseudomonas aeruginosa</i>	Piocina S13
BAA02196.1	<i>Pseudomonas aeruginosa</i>	Piocina APA41
P02978.1	<i>Escherichia coli</i>	Colicina E1
WP_172690082.1	<i>Escherichia coli</i>	Colicina E2
WP_000012964.1	<i>Escherichia coli</i>	Colicina E3
WP_052980126.1	<i>Escherichia coli</i>	Colicina E6
WP_021530049.1	<i>Escherichia coli</i>	Colicina E7
WP_012644886.1	<i>Escherichia coli</i>	Colicina E9
WP_016245160.1	<i>Escherichia coli</i>	Colicina D
WP_058645033.1	<i>Salmonella enterica</i>	Colicina
WP_061371543.1	<i>Salmonella enterica</i>	Colicina E3
WP_058216742.1	<i>Salmonella enterica</i>	Salmocina E7
EAT2521868.1	<i>Salmonella enterica</i>	Salmocina E1a
OIN32443.1	<i>Salmonella enterica</i>	Salmocina E1b

Posteriormente, las secuencias predichas de aminoácidos para las piocinas identificadas inicialmente como carocina D, colicina E3 y colicina por BAGEL4, se obtuvieron a partir de los *contigs* del genoma de *Pseudomonas* sp. I1. Se procedió a buscar los dominios proteicos característicos de las piocinas en estos candidatos y evaluar si corresponden a homólogos de piocinas conocidas.

La primera bacteriocina analizada, la denominada inicialmente como carocina D, fue identificada por BAGEL4 con base en su parecido con la carocina D, la cual es una bacteriocina de *Pectobacterium carotovorum* (Roh et al. 2010). Por lo tanto, se realizó un alineamiento entre la bacteriocina de *Pseudomonas* sp. I1 y la carocina D de *P. carotovorum* utilizando el programa BioEdit por el método *optima GLOBAL alignment* y para crear el gráfico se utilizó el programa Jalview. Se observó que la proteína de *Pseudomonas* sp. I1 posee identidad con el dominio citotóxico localizado en el extremo C-terminal de la carocina D, pero no con el resto de la secuencia (Figura 11).



**Figura 11. Alineamiento de carocina D hipotética de *Pseudomonas* sp. I1 (parte superior) con la carocina D de *P. carotovorum* (parte inferior).** El color azul indica aminoácidos idénticos y color azul celeste aminoácidos similares entre ambas bacteriocinas.

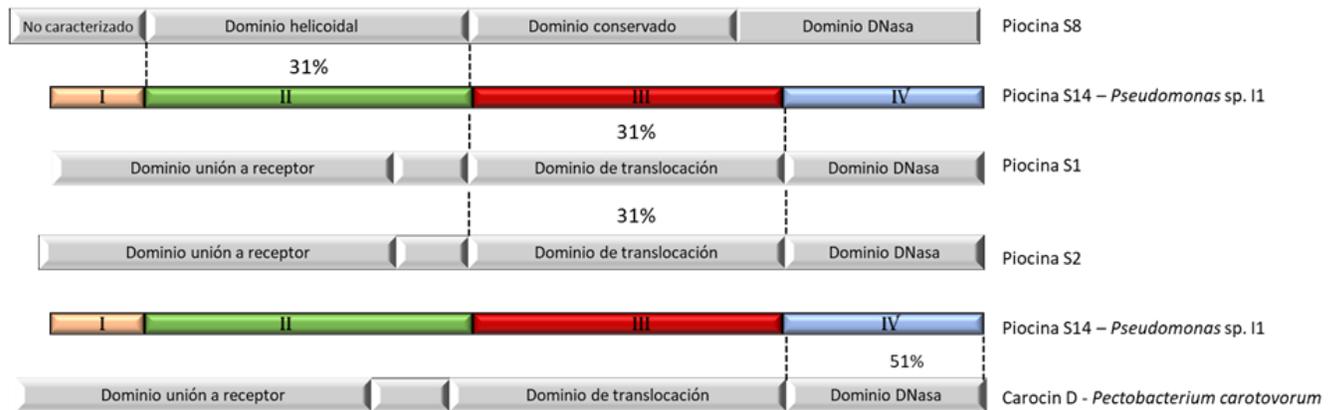
Es decir, la bacteriocina identificada *in silico* como carocina D en *Pseudomonas* sp. I1 sería una nueva bacteriocina que comparte el dominio C-terminal con carocina D, pero difiere en la secuencia N-terminal de la proteína, por ello se denominó en esta tesis como piocina S14.

El dominio C-terminal de la carocina D corresponde al dominio citotóxico y se ha documentado experimentalmente que posee actividad DNasa (Roh et al. 2010), por lo tanto, la piocina S14 correspondería hipotéticamente a una DNasa.

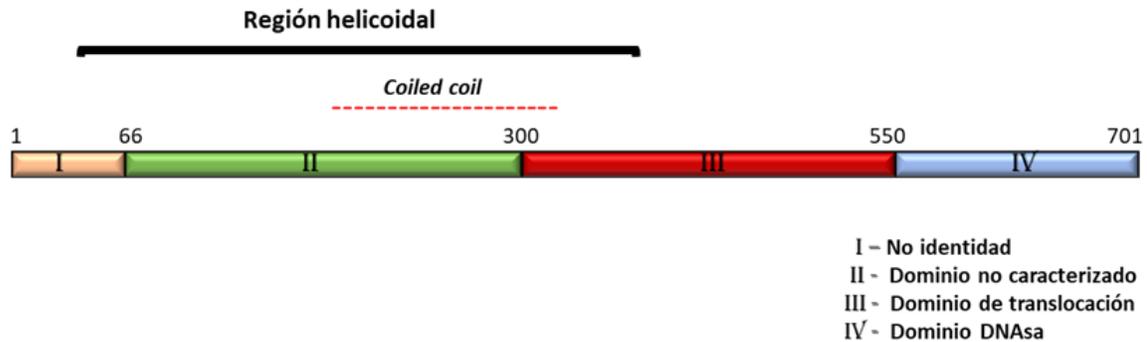
Luego, para identificar otros dominios hipotéticos presentes en la piocina S14, utilizando BLAST se realizó un análisis de comparación de la secuencia de aminoácidos de la piocina S14 contra una base de datos local de bacteriocinas ya caracterizadas. Una vez identificadas las bacteriocinas que comparten identidad con la secuencia de la piocina S14 se realizaron alineamientos específicos con éstas. Este análisis confirmó que la piocina S14 posee una región de 151 aminoácidos en el extremo C-terminal que tiene una identidad de 51% con el dominio citotóxico DNasa de carocina D de *P. carotovorum*. Posteriormente, se identificó una región que tiene identidad del 31% con los dominios de translocación de las piocinas S1 y S2, seguida de una región con un 31% de identidad con un dominio helicoidal identificado en la piocina S8. Por último, se determinó que la piocina S14 tiene una región de 66 aminoácidos en el extremo N-terminal que no posee identidad con secuencias caracterizadas (Figura 12A).

Para finalizar este análisis, se llevó a cabo una predicción de estructura secundaria y búsqueda de motivos *coiled coil* hipotéticos mediante los programas Jpred4 y COILS. Dicha búsqueda identificó una región helicoidal en la posición 37-315, la cual corresponde a la región de identidad con la piocina S8, y un motivo *coiled coil* en la posición 220-320 (Figura 12B), predicciones que son importantes debido a que los dominios de unión a receptor de las colicinas generalmente tienen estas características (Zakharov et al. 2002). Este análisis reveló una nueva bacteriocina, denominada piocina S14, ya que posee una combinación de dominios diferente a la de las bacteriocinas conocidas.

A



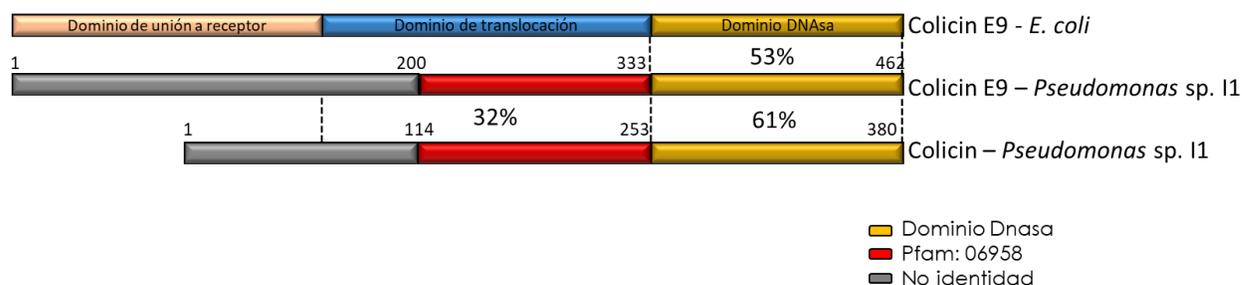
B



**Figura 12. Organización de dominios de la piocina S14.** A. Se muestra una región similar en el C-terminal (azul claro) al dominio DNasa de *P. carotovorum*, una región (rojo) que tiene identidad con los dominios de translocación de las piocinas S1 y S2, seguido de una región (verde) similar al dominio helicoidal no caracterizado de la piocina S8 y, por último, una región que no tiene identidad con nada caracterizado (beige). B. Predicción de la organización de dominios funcionales en la piocina S14, predicción de una región helicoidal (línea negra) y una región *coiled coil* (línea entrecortada rojo) en la estructura.

Mediante un análisis similar de comparación de las bacteriocinas identificadas por BAGEL4 como colicina E9 y colicina con las bacteriocinas en la base de datos local de piocinas caracterizadas, se identificó el dominio citotóxico C-terminal, ya que estas regiones poseen una identidad de 53% y 61%, respectivamente, con el dominio DNasa de la colicina E9 de *E. coli*. Estas dos colicinas conservan una secuencia perteneciente al Pfam 06958, el cual está descrito como dominio conservado en las piocinas tipo S. Además, en la región central comparten una identidad del 32% entre ellas. Sin embargo, no poseen identidad con otras

bacteriocinas, ni con ninguna otra proteína en las bases de datos, en el extremo amino terminal (Figura 13). Estos resultados revelan que estas dos bacteriocinas no poseen identidad con dominios presentes en el extremo N-terminal de bacteriocinas conocidas, por lo cual no se puede postular una posible función para estos genes.



**Figura 13. Dominios hipotéticos de colicina E9 y colicina de *Pseudomonas sp. I1* identificados por BAGEL4.** Una región en el C-terminal (amarillo) de las dos piocinas es similar al dominio DNasa de la colicina E9 de *E. coli*, poseen una región (rojo) correspondiente al Pfam 06958, dominio que conservan algunas piocinas tipo S, y un fragmento en el N-terminal (gris).

Búsqueda de homólogos de la piocina S14 e identificación de motivos conservados en el dominio citotóxico de las bacteriocinas identificadas por BAGEL4. De acuerdo con los resultados descritos anteriormente, la piocina S14 presenta una combinación de dominios nueva, ausente en las piocinas caracterizadas hasta la fecha. Dada la relevancia de este hallazgo, se planteó identificar la presencia de esta piocina en otras bacterias de las bases de datos, para lo cual se realizó una búsqueda de homólogos de la piocina S14 de *Pseudomonas sp. I1* mediante BLASTP en el servidor NCBI. Se seleccionaron como *hits* positivos aquellas proteínas que conservan identidad a lo largo de la secuencia para incluir solo homólogos que posean los tres dominios hipotéticos (unión a receptor, translocación y actividad DNasa) similares a los de la piocina S14 y además que conserven el dominio específico de la piocina S14 que se encuentra en el extremo N-terminal. En total se encontraron 20 homólogos conservados exclusivamente en especies de *Pseudomonas* (Tabla 3).

**Tabla 3. Proteínas homólogas de la piocina S14**

Número de acceso en GenBank	Especie	Identidad %	Similitud %	Cobertura %
-----------------------------	---------	-------------	-------------	-------------

WP_116254650.1	<i>Pseudomonas</i> sp. OV221	77	84	99
WP_095112657.1	<i>Pseudomonas</i> sp. Irchel 3H7	74	86	99
WP_100846211.1	<i>Pseudomonas baetica</i>	67	77	99
WP_150617165.1	<i>Pseudomonas fluorescens</i>	65	74	100
WP_007914223.1	unclassified <i>Pseudomonas</i>	70	81	99
WP_073473732.1	<i>Pseudomonas moraviensis</i>	67	77	99
WP_133334773.1	<i>Pseudomonas moraviensis</i>	67	78	99
WP_130928305.1	<i>Pseudomonas</i> sp. Sample_20	67	77	99
WP_122750882.1	unclassified <i>Pseudomonas</i>	67	76	99
WP_130901479.1	<i>Pseudomonas</i> sp. Sample_23	67	77	99
WP_122747695.1	<i>Pseudomonas viridiflava</i>	67	77	99
WP_197890289.1	<i>Pseudomonas moraviensis</i>	67	77	99
MBJ7372397.1	<i>Pseudomonas</i> sp.	66	77	99
WP_130887180.1	<i>Pseudomonas</i> sp. Sample_21	66	77	99
WP_130927067.1	<i>Pseudomonas</i> sp. Sample_14	66	77	99
WP_114886589.1	<i>Pseudomonas kribbensis</i>	68	78	99
WP_085666962.1	unclassified <i>Pseudomonas</i>	68	79	99
WP_186568116.1	<i>Pseudomonas</i> sp. SWRI54	65	76	99
WP_130911346.1	<i>Pseudomonas</i> sp. Sample_9	63	75	99
WP_137215479.1	<i>Pseudomonas</i> sp. CFBP13508	64	75	99

En un trabajo previo, Sharp et al. (2017) identificaron distintas familias de nucleasas DNasa (HNH DNasa, no HNH DNasa) y RNasas (tRNasa y rRNasa) a partir de 3094 dominios citotóxicos de bacteriocinas de diversas especies bacterianas. En este estudio, se realizó un análisis de las secuencias aminoacídicas de los dos dominios DNasa conocidos (HNH y no HNH) para determinar motivos conservados en estas familias. Se encontró que las bacteriocinas con el dominio nucleasa HNH conservan el motivo  $HHX_{14}NX_8HX_8H$ , mientras que las no HNH conservan el motivo  $MX_{46}GX_2LSF$ . Una búsqueda extensiva de dichos motivos conllevó a la identificación de un gran número de bacteriocinas DNasa hipotéticas en los genomas bacterianos de las bases de datos, algunas de ellas nuevas, y a la observación que éstas son exclusivas de las  $\gamma$ -proteobacterias (Sharp et al. 2017). Es importante destacar que en este estudio no se reporta ninguna proteína homóloga a la piocina S14.

Considerando los antecedentes entregados y analizados por Sharp et al. (2017), para identificar a cuál de estos tipos corresponden las bacteriocinas hipotéticas de *Pseudomonas* sp. I1, se planteó buscar los motivos nucleasas (HNH y no HNH) en los tres dominios DNasa, mediante inspección manual y alineamientos. Este análisis

se complementó con una comparación de la predicción de estructura secundaria de dichos dominios utilizando el programa Jpred4, con las predicciones de estructura secundaria de dominios canónicos HNH y no HNH. Como ejemplo de estructura secundaria de las HNH DNasa, la cual conserva dos láminas beta y una alfa hélice (Cascales et al 2007), se utilizó la colicina E9 de *Escherichia coli*, y para la no HNH DNasa se utilizó la secuencia del dominio citotóxico de una piocina tipo no HNH (WP\_016394378.1) (Sharp et al. 2017), cuya estructura secundaria fue predicha en este estudio, la cual contiene dos hélices alfa y dos láminas beta (Figura 14A). A partir de estos análisis, se encontró que la colicina y colicina E9 de *Pseudomonas* sp. I1 poseen el motivo y la estructura secundaria de las HNH DNasa, mientras que en la piocina S14 no se encontró ninguno de los motivos consenso conocidos y presenta una estructura secundaria predicha que consiste en cuatro hélices alfa, que es diferente a las estructuras de los dominios DNasa descritos previamente en bacteriocinas (Figura 14B). Por otro lado, esta estructura secundaria es similar a la predicha para el dominio DNasa identificado en carocina D de *P. carotovorum* (Figura 14B, piocina S14 y Figura 14C), lo cual apoya la hipótesis de la conservación de una función DNasa similar al de la carocina D para el dominio citotóxico de la piocina S14. El hecho que el dominio DNasa hipotético de la piocina S14 no posea los motivos conservados por las DNasas HNH y no HNH, así como la predicción de una estructura secundaria diferente, sugiere que esta piocina representa un nuevo dominio nucleasa.

**A**

**Colicina-E9- *Escherichia coli***

VGGRK**VYELHH**DKPISQGGEVYDMD**N**IRVTTPKR**HIDIHRGK**



**Piocina tipo no HHN DNasa de *Pseudomonas* WP\_016394378.1**

**M**NAKELDVYLAGLGKPHDDLVEHGVLSPGSLIEIYPGSLSLYRDLFP**GV**EL**SF**WAEDQKFEKINISLSKTALSGVV**FQHE**



**B**

**Colicina E9- *Pseudomonas* sp. I1**

GENARYE**IHH**IENIQHGGA**VYDIDN**LRVATPKR**HVEIHKEDRQKS**



**Colicina - *Pseudomonas* sp. I1**

TRGV**FEIHH**VERIAEGGAVYD**VDN**LRVNTPK**NHISQH**KRQ



**Piocina S14 – *Pseudomonas* sp. I1**

**H**IARLDQSD**ANE**IMIQR**L**EKILSGHWMLTDT**DL**RYYT**HE**IRE**L**ERFRAMGL**SDE**FKPESGSPAW**NNA**HTAT**LE**DYKLGSE**DYLL**YSPDAVQAVNSQIDRIYEQL**LK**GT**YQ**



**C**

**Carocina D – *Pectobacterium carotovorum***

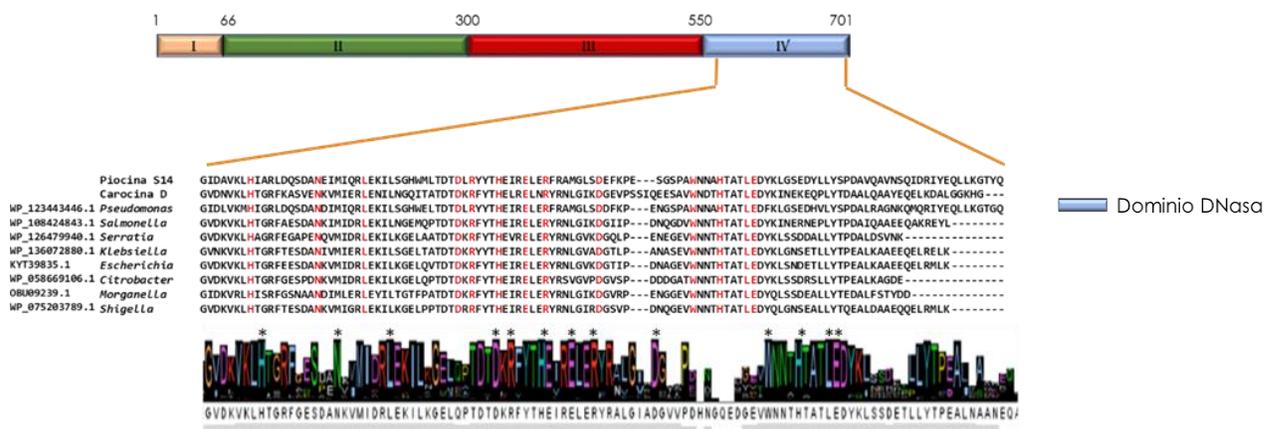
**H**TGRFKAS**VEN**KVMIER**L**ENILNGQITATDT**DK**RFY**THE**LR**EL**NR**YR**NLGI**KD**GEVPSSIQEESAV**W**ND**HT**AT**LE**DYKINEKEQPLYTDAALQAAYEQELKDALGG**KHG**



**Figura 14. Predicción de estructura secundaria del dominio C-terminal de las bacteriocinas. A.**

Se representa la secuencia primaria y la predicción de la estructura secundaria del motivo nucleasa HHN de la colicina E9 de *E. coli* y no HHN de la piocina (WP\_016394378.1) de *Pseudomonas*, respectivamente. Los aminoácidos del motivo conservado se indican en rojo en ambos casos. **B.** Colicina E9, colicina y la piocina S14 de *Pseudomonas* sp. I1 con el motivo conservado (letras en color rojo) y su respectiva estructura secundaria en flechas láminas beta (verde) y hélices alfa representadas en líneas (rojo). **C.** Aminoácidos conservados (letras en color rojo) y estructura secundaria (línea roja) de *P. carotovorum*. El motivo conservado en la piocina S14 y carocina D proviene de un análisis posterior.

Debido a que la piocina S14 no conserva motivos DNasa que hayan sido descritos, se realizó una búsqueda extensiva en bases de datos mediante BLASTP de la secuencia de aminoácidos que corresponde al dominio C-terminal de piocina S14 (aminoácidos 550 al 701), para identificar proteínas que conserven al dominio DNasa hipotético y poder realizar una comparación entre ellas. Para seleccionar *hits* positivos, se utilizó como criterio de selección mínimo el valor del exponente del número E correspondiente al alineamiento del dominio DNasa del homólogo completo más lejano a la piocina S14 determinado anteriormente (proteína WP\_137215479.1, exponente del número E igual a -49). En total, se encontraron 953 proteínas conservadas en diversas especies de  $\gamma$ -proteobacterias con este dominio. Dentro de estos homólogos podrían encontrarse nuevas bacteriocinas que no han sido identificadas y que conservan este dominio C-terminal. Se realizó un alineamiento múltiple de las secuencias primarias del dominio C-terminal de estos 953 homólogos, incluyendo la piocina S14, utilizando ClustalO desde Jalview. Se determinó la conservación de 13 aminoácidos en el dominio DNasa: H592, N602, L609, D623, R625, H629, E633, R637, D645, W656, H659, L663 y E664 (Figura 15). Estos aminoácidos 100% conservados podrían representar aminoácidos esenciales para la función de este dominio. Estos resultados revelan que el dominio DNasa de la piocina S14 posiblemente pertenezca a una nueva familia.



**Figura 15. Alineamiento de las 953 proteínas que conservan el dominio C-terminal de la piocina S14.** Se tomó como referencia una secuencia de cada género bacteriano para mostrar en el gráfico. Se muestra la secuencia primaria de una región seleccionada del dominio DNasa y se indican los aminoácidos conservados (rojo) y un logo donde se observa el consenso de los 953 homólogos. Los asteriscos (\*) señalan cada aminoácido que tiene un 100% de conservación.

## 8. DISCUSIÓN

La búsqueda de vías biosintéticas para metabolitos especializados y bacteriocinas en bacterias es de gran importancia debido a su capacidad bioactiva, siendo relevantes en el ámbito de la salud, por su potencial para combatir bacterias patógenas resistentes a antimicrobianos y también en el área de alimentación por contribuir en la conservación y seguridad de alimentos. Así como algunos estudios también han determinado su capacidad para suprimir enfermedades causadas por patógenos que afectan a plantas (Lozano et al. 2017). Este trabajo exploró el potencial la cepa *Pseudomonas* sp. I1, de codificar proteínas asociadas a la producción de compuestos inhibidores de amplio espectro de inhibición de patógenos, pero cuyos mecanismos de inhibición son desconocidos.

### 8.1. Identificación de metabolitos especializados y bacteriocinas mediante minería genómica

El análisis del genoma de la bacteria inhibidora *Pseudomonas* sp. I1 mediante los programas antiSMASH y BAGEL4 reveló la presencia de tres diferentes tipos de vías con potencial antimicrobiano (Tabla 2): vías tipo NRPS, betalactona y bacteriocinas. Para los tres casos, varios ejemplos de moléculas derivadas con actividad antimicrobiana han sido descritos en bibliografía. Por ejemplo, en *Streptomyces viridochromeogenes* se identificó una NRPS que produce laspartomicinas que tienen un efecto antibiótico que actúa contra patógenos Gram positivo como MRSA (Nikolouli et al. 2012). Otro ejemplo es obtenido desde el genoma de *Enterococcus faecium* E86, donde se detectaron bacteriocinas de importancia como la enterocina P y enterocina TW21. Destaca la posible aplicación de la enterocina P como un agente alternativo a los antibióticos clásicos contra *Enterococcus* resistentes a la vancomicina (Miceli de Farias et al. 2019). También se han reportado betalactonas con actividad antibiótica, como la himeglusina (Skaff et al. 2012). Por lo tanto, encontrar estos *clusters* en *Pseudomonas* sp. I1 sugiere que esta especie bacteriana tiene la capacidad de producir un variado espectro de moléculas inhibidoras.

Es importante considerar que existe la posibilidad que algunos *clusters* no sean identificados por programas de minería genómica, ya que no poseen identidad con los *clusters* de bases de datos. Ejemplo de esto es lo reportado en un estudio donde mediante análisis experimentales se identificó en una cepa de *Pseudomonas* de origen ambiental un *cluster* conformado por genes codificantes de una proteína de unión a flavina adenina dinucleótido, una peroxirredoxina y una hidrolasa de ácidos grasos, que estaría involucrado en la inhibición de cepas resistentes a múltiples fármacos, que no fue identificado por programas de minería genómica (Basalla et al. 2019).

Tres de los *clusters* que inicialmente se identificaron como vías tipo NRPS, lokisina, nunamicina y cichopectina (Tabla 2), se encuentran en el límite de los respectivos *contigs*, lo que sugirió que la secuenciación y el ensamblaje no permitieron cubrir *cluster* completos, de acuerdo a los tamaños de otros *loci* similares ya conocidos. Nuestro análisis sugiere que dichos *clusters* no serían independientes, sino que podrían corresponder a fragmentos de una misma vía. Para verificar esta hipótesis se podría realizar una amplificación de las regiones flanqueantes mediante la reacción en cadena de la polimerasa utilizando partidores específicos de los *contigs* 0034, 0039, 0014 y comprobar la amplificación el *cluster* completo.

Finalmente, las bacteriocinas encontradas en *Pseudomonas* sp. I1, según BAGEL4 poseen identidad con la carocina D, colicina E9 y una bacteriocina denominada genéricamente colicina, siendo todas piocinas tipo S.

## **8.2. Identificación de un *cluster* de genes en *Pseudomonas* sp. I1 para la síntesis de lokisina.**

El análisis de los tres *clusters* (lokisina, nunamicina y cichopectina) identificados por antiSMASH, muestra que en conjunto forman un *cluster* completo para la producción de lokisina en el genoma de *Pseudomonas* sp. I1, ya que posee una identidad del 84,5% con el *cluster* de biosíntesis de lokisina de *Pseudomonas protegens* pf-5, el cual está conformado por genes core: *lokA*, *lokB* y *lokC* y genes accesorios: *lokT*, *lokR*, *macA* y *macB*. Estos genes están conservados en el *cluster* identificado en *Pseudomonas* sp. I1, sugiriendo que se trata de una vía que sintetiza un producto

funcional. Este *cluster* está codificado principalmente en los *contigs* 0034, 0039 y 0014. Sin embargo, existe una región repetida de 1101 nucleótidos que en el genoma de *Pseudomonas* sp. I1 sólo se ensambló una vez en los *contigs* 0054 y 0068. Esta región repetida pudiera ser la causa por la que el proceso de ensamblaje del genoma original no logró unir estas regiones en un solo *contig*.

Identificar una lokisina es relevante debido a su capacidad antimicrobiana. En estudios previos se ha reportado que la lokisina de una cepa de *P. koreensis* es capaz de inhibir patógenos fúngicos vegetales como *Phytophthora nicotianae* y *Rhizoctonia solani*, mientras que la lokisina de *Pseudomonas* sp. COR10 controla la enfermedad del arroz causada por *Magnaporthe oryzae* (Gu et al. 2020). En *P. koreensis* 2.74, la lokisina disminuye la enfermedad del tomate causada por *Pythium ultimum* (Hultberg et al. 2010). Debido al porcentaje de conservación a nivel de nucleótidos entre los *clusters* y también a nivel de aminoácidos entre las proteínas codificadas, es probable que *Pseudomonas* sp. I1 produzca lokisina con capacidades muy parecidas a las reportadas en bibliografía. En los casos descritos, la lokisina en general tiene actividad contra fitopatógenos fúngicos, lo cual sugiere que nuestra cepa podría tener potencial antifúngico en plantas. Sin embargo, el papel de la lokisina en la inhibición de patógenos bacterianos no ha sido reportado. Por lo tanto, con los antecedentes aquí aportados, no es posible postular un papel de la vía de biosíntesis de lokisina en la inhibición de los patógenos bacterianos humanos por *Pseudomonas* sp. I1, sin embargo, no se descarta la posibilidad que esta lokisina inhiba el crecimiento de patógenos bacterianos.

### **8.3. Identificación de un *cluster* de genes en *Pseudomonas* sp. I1 para la síntesis de pioverdina**

El descubrimiento de una vía biosintética de pioverdina es relevante debido a que este sideróforo es capaz de secuestrar hierro férrico del medio e introducirlo hacia la bacteria productora mediante la interacción específica con el receptor de pioverdina, provocando una reducción en la biodisponibilidad del hierro para el crecimiento de otros microorganismos (Ringel y Brüser. 2018). Estudios han reportado que las pioverdinas producidas por cepas de *P. fluorescens* pueden inhibir

el crecimiento de varios patógenos que afectan a peces como *Vibrio ordalii*, *Aeromonas salmonicida*, *Lactococcus garvieae*, *Streptococcus iniae* y también a las bacterias fitopatógenas *Erwinia carotovora* y *Ralstonia solanacearum* (Satonyo et al. 2009). Por lo tanto, la pioverdina podría estar implicada en la inhibición del crecimiento de *V. cholerae*, la especie más relacionada con *V. ordalii* entre los patógenos inhibidos por esta cepa, y posiblemente de otros patógenos. Las bacterias sobre las que se realizaron previamente ensayos de inhibición por nuestro grupo también producen sus propios sideróforos para compensar la deprivación de hierro, y algunas como *V. cholerae* y *S. aureus* usan receptores de xeno-sideróforos (utilizan sideróforos producidos por otros organismos) (Holden et al. 2016; Parquet et al 2018; Trapet et al 2016; Wyckoff et al. 2015; Searle et al. 2015). Por lo tanto, es probable que dichas cepas sean capaces de competir con *Pseudomonas* sp. I1 por el hierro y no es posible determinar si la pioverdina en este caso es suficiente para causar la inhibición. Para producir dicho efecto, estas pioverdinas deberían tener más afinidad por el hierro que los sideróforos producidos por las otras especies, pudiendo de esta manera causar la depleción de hierro en las especies inhibidas.

Se ha reportado que las pioverdinas inhiben el crecimiento de patógenos fúngicos, como *Fusarium oxysporum*, el cual causa la enfermedad del marchitamiento de la papa, además contribuye al control biológico de *Gaeumannomyces graminis* que provoca una deficiencia en el crecimiento del trigo y la cebada, e incluso suprimen los fitopatógenos en el maní y el maíz (Ahmed y Holmström 2014; Trapet et al. 2016). Por lo tanto, la pioverdina identificada en *Pseudomonas* sp. I1 podría tener un potencial de biocontrol sobre patógenos fúngicos vegetales.

#### **8.4. Identificación de una vía que codifica hipotéticamente para betalactona**

El *cluster* identificado cuyo producto es betalactona en *Pseudomonas* sp. I1, está conformado por 17 genes, de los cuales solo 2 tienen identidad con la vía de biosíntesis de fengicina, que es la vía más cercana según antiSMASH. Excepto el mencionado BGC, este *cluster* no tiene identidad con ningún otro *cluster* de betalactonas conocidos en la base de datos MiBIG, sin embargo, el *cluster*

identificado en *Pseudomonas* sp. I1 conserva genes accesorios que normalmente conforman un *cluster* para la síntesis de betalactona. Por lo tanto, podría corresponder a un nuevo tipo de betalactona. Se ha reportado que las betalactonas tienen capacidad antifúngica, como es el caso de la fengicina, y antibacteriana como la obafluorina producida por *P. fluorescens* ATCC 39502 que es activa contra patógenos Gram positivo y Gram negativo (Scott et al. 2019; Gulder et al. 2010; Böttcher et al. 2009; Zihahirwa et al. 2017), por lo cual éste se considera un *cluster* relevante ya que podría contribuir al desarrollo de nuevas moléculas con capacidad antimicrobiana y podría dar cuenta de la capacidad antibacteriana observada en nuestros ensayos.

#### **8.5. Caracterización de las bacteriocinas identificadas en *Pseudomonas* sp. I1 por BAGEL4**

El análisis de las tres bacteriocinas identificadas por BAGEL4 determinó que una de ellas corresponde a una bacteriocina nueva, ya que posee una combinación de dominios nueva, ausente en bacteriocinas caracterizadas, a la que se denominó como piocina S14.

El análisis reveló que la piocina S14 parece tener un dominio de translocación similar al de las piocinas S1 y S2 y un dominio citotóxico similar al de carocina D. Se ha reportado que este dominio de carocina D posee actividad DNasa (Roh et al. 2010). De manera interesante, S14 no presenta homología con ningún dominio de unión a receptor conocido. Sin embargo, la predicción de una región helicoidal y una región *coiled coil* en la región II (véase la Figura 12B) de la piocina S14, sugiere que esta región podría conformar un dominio de unión a receptor, ya que dichas cualidades son típicas de los dominios de unión a receptor de las colicinas (Zakharov et al. 2002). Esta región solo conserva identidad con una región similar no caracterizada en la piocina S8. El hecho de no encontrar homólogos con dominios de unión a receptor caracterizados, sugiere que la piocina S14 posee una variante nueva de dominio de unión de receptor. Generalmente, los receptores que reconocen las piocinas S son proteínas de membrana externa que conservan una estructura de barril  $\beta$ , como los receptores de ferri-sideróforo o los receptores de

vitamina B12 (Dingemans et al. 2016). La mayoría de las piocinas conocidas poseen un dominio de unión a receptor que reconocen variantes específicas de los receptores de membrana de la célula diana, en este caso de algunas cepas de *Pseudomonas*. Por lo tanto, el dominio de unión a receptor de la piocina S14 posiblemente tenga la capacidad de detectar otras variaciones de los receptores de membrana de células diana, que normalmente no reconocen los dominios de unión a receptor conocidos, ampliando de esta manera el espectro de cepas dentro del género *Pseudomonas* inhibidas por las piocinas. La región de 66 aminoácidos en el amino terminal correspondería a una región no estructurada común en las piocinas, la cual se encarga de encontrar y unirse al transportador, dicha unión hace que la piocina se reoriente permitiendo que el dominio de translocación se involucre con el transportador (Behrens et al. 2020).

Para las piocinas identificadas colicina E9 y colicina no es posible postular una función hipotética, debido a que no poseen identidad con dominios de bacteriocinas conocidas hacia el extremo amino terminal. Además, estas dos bacteriocinas no conservan región helicoidal y el motivo *coiled coil*, las cuales son características de los dominios de unión a receptor de las colicinas, razón por la cual posiblemente estas secuencias no correspondan a bacteriocinas funcionales. Sin embargo, también podría tratarse de nuevas bacteriocinas. Para probar estas hipótesis se debería realizar análisis experimentales.

De manera reciente, ha cobrado relevancia el estudio de las piocinas tipo S debido a su potencial para uso en terapia antimicrobiana. Por ejemplo, la piocina S8 a una concentración de 10 µg/mL muestra una potente actividad antibacteriana contra los aislados de *P. aeruginosa* resistente a múltiples fármacos (Turano et al. 2020). También se ha determinado que la piocina SD2 tiene la capacidad de resolver infecciones pulmonares causadas por *P. aeruginosa* en un modelo murino. En dicho estudio se infectaron intranasalmente con *P. aeruginosa* PAO1 y se trataron intranasalmente 1 hora después de la inoculación con piocina SD2. Los ratones fueron después sacrificados y se realizaron recuentos bacterianos de los pulmones que mostraron una disminución de *P. aeruginosa* al tratar a los ratones con piocina

SD2, mostrando así que esta piocina tiene una excelente actividad antimicrobiana *in vivo* (McCaughey et al. 2016).

Por lo tanto, la identificación de la piocina S14 es importante y amplía el número de candidatos para el desarrollo terapéutico. Debido a que usualmente las piocinas son utilizadas por especies de *Pseudomonas* para inhibir el crecimiento de otras *Pseudomonas*, es posible que esta bacteriocina esté involucrada en la inhibición del crecimiento de la cepa *P. aeruginosa* que fue utilizada previamente para ensayos de inhibición realizados por nuestro grupo de investigación. Sin embargo, esta posibilidad, así como la determinación de otros posibles blancos de la piocina S14, requiere aproximaciones experimentales para probarse.

#### **8.6. Identificación de homólogos de la piocina S14 y motivos conservados del dominio citotóxico C-terminal.**

Los homólogos de la piocina S14 se encontraron exclusivamente en bacterias del género *Pseudomonas* y además son homólogos que no se han caracterizado como bacteriocinas dentro de este género. Cuando se realizó la búsqueda de motivos conservados HNH y no HNH, y la predicción de estructura secundaria que conservan estos dominios en las bacteriocinas identificadas por BAGEL4 en *Pseudomonas* sp. I1, se evidenció que la piocina S14 no conserva dichos motivos y que no posee la estructura secundaria típica de las nucleasas descritas. La piocina S14 tiene una estructura secundaria similar al dominio DNasa identificado en la carocina D de *P. carotovorum*. Por lo tanto, es posible que esta piocina posea un nuevo dominio nucleasa. La búsqueda de homólogos del C-terminal de la piocina S14, determinó que existen diversas proteínas con este dominio codificadas en diversas especies bacterianas de  $\gamma$ -proteobacterias, lo que coincide con las bacteriocinas nucleasas HNH y no HNH que también son exclusivas de esta familia. Posteriormente, mediante alineamiento de las 953 proteínas con este dominio se identificaron los aminoácidos 100% conservados en el dominio. Estos análisis indican que podrían representar aminoácidos esenciales para la función de este nuevo dominio DNasa hipotético. Dentro de los homólogos encontrados que conservan este dominio C-terminal podrían encontrarse nuevas bacteriocinas, por

lo que es necesario un análisis de la secuencia de dichas proteínas para identificar otras posibles combinaciones novedosas de dominios de bacteriocinas en este grupo.

En este estudio se encontraron posibles determinantes de inhibición en el genoma de *Pseudomonas* sp. I1, como son lokisina y pioverdina, ejemplos de estas moléculas se conocen y fueron caracterizadas por su potencial antimicrobiano en otras especies bacterianas. Además, se identificó una betalactona que posiblemente corresponda a un nuevo tipo, y una bacteriocina, la piocina S14 que posiblemente sea una bacteriocina nueva que aún no ha sido caracterizada. Para confirmar los resultados predictivos obtenidos en este estudio, se requerirá realizar experimentos con la finalidad de corroborar la capacidad inhibitoria de cada uno de los determinantes identificados.

Dentro de los *clusters* identificados, la piocina S14 sería la mas relevante por las predicciones obtenidas en este estudio y por la capacidad antimicrobiana que poseen las piocinas según la bibliografía disponible. Por lo tanto, la caracterización experimental de esta bacteriocina sería uno de los objetivos más inmediato posterior a este estudio. Para verificar la funcionalidad de dicha piocina se podría realizar el clonamiento del operón que codifica la piocina S14 y su gen de inmunidad amplificado estos genes mediante PCR utilizando partidores que incluyan sitios de reconocimiento de endonucleasas de restricción compatibles para ligar el producto amplificado en un plásmido apropiado como el pBBR1 MCS1. Los plásmidos recombinantes serían secuenciados y transformados en *E. coli* DH5 $\alpha$ . Para evaluar el potencial inhibitorio se obtendría un extracto total de proteínas de las bacterias transformadas tanto con el plásmido recombinante que sintetice la piocina S14 o con el vector solo, sería utilizado en ensayos de inhibición en placas de agar LB sembradas con diferentes cepas de *Pseudomonas* patógenas (género bacteriano generalmente diana de piocinas), sobre las cuales se añadirá un disco de papel impregnado con el extracto. Los resultados se analizarían evaluando la formación de un halo de inhibición.



## 9. CONCLUSIONES

Con base en los resultados encontrados en este estudio se obtuvieron las siguientes conclusiones:

- Mediante análisis del genoma de *Pseudomonas* sp. I1 utilizando programas de minería genómica se identificó la presencia de tres tipos de *clusters* con potencial antimicrobiano: *clusters* para vías tipo NRPS, betalactonas y bacteriocinas. Este resultado sugiere que esta especie bacteriana tiene la capacidad de producir un variado espectro de moléculas inhibidoras.
- Utilizando el programa antiSMAH se identificaron *clusters* tipo NRPS que codifican hipotéticamente a lokisina y pioverdina que posiblemente tengan potencial inhibitorio. Adicionalmente se identificó una vía que sintetizaría una betalactona.
- Con BAGEL4 se identificaron tres *clusters* que codifican para bacteriocinas hipotéticas, una de ellas, la piocina S14, que posee una nueva combinación de dominios, con homólogos conservados en otras especies de *Pseudomonas*.
- La piocina S14 no posee los dominios HNH y no-HNH descritos, contiene otros aminoácidos conservados lo que sugiere un nuevo dominio nucleasa.

## 10. BIBLIOGRAFÍA

Agrawal S, Acharya D, Adholeya A, Barrow C, Deshmukh S. Nonribosomal peptides from marine microbes and their antimicrobial and anticancer potential. 2017. *Frontiers in Pharmacology*. 8: 828.

Agrawal S, Adholeya A, Deshmukh S. The pharmacological potential of non ribosomal peptides from marine sponge and tunicates. 2016. *Frontiers in Pharmacology*. 7: 333.

Ahmed E, Holmström SJ. Siderophores in environmental research: roles and applications. 2014. *Microbial Biotechnology*. 7(3): 196-208.

Aslam B, Wang W, Arshad MI, Khurshid M, Muzammil S, Rasool M, Nisar M, Alvi R, Aslam M, Qamar M, Salamat M, Baloch Z. Antibiotic resistance: a rundown of a global crisis. 2018. *Infection and Drug Resistance*. 11: 1645-1658.

Babalola OO, Ayangbenro AS. Draft genome sequence of *Pseudomonas koreensis* strain AB36, isolated from gold mining soil. 2019. *Microbiology Resource Announcements*. 8 (20): e00175-19

Basalla J, Chatterjee P, Burgess E, Khan M, Verbrugge E, Wiegmann D, LiPuma J, Wildschutte H. Loci encoding compounds potentially active against drug-resistant pathogens amidst a decreasing pool of novel antibiotics. 2019. *Applied and Environmental Microbiology*. 85(23): e01438-19.

Barry A, Fuchs P, Brown S. *In vitro* activities of Daptomycin against 2,789 clinical isolates from 11 North American medical centers. 2001. *Antimicrobial Agents and Chemotherapy*. 45 (6): 1919-1922.

Begrem S, Ivaniuk F, Gigout-Chevalier F, Kolypczuk L, Bonnetot S, Leroi F, Grovel O, Delbarre-Ladrat C, Passerini D. New insight into antimicrobial compounds from food and marine-sourced *Carnobacterium* species through phenotype and genome analyses. 2020. *Microorganisms*. 8(7): 1093.

Behrens HM, Lowe ED, Gault J, Housden NG, Kaminska R, Weber T, Thompson, Mislin G, Schalk D, Robinson C, Kleanthous C. Pyocin S5 import into *Pseudomonas aeruginosa* reveals a generic mode of bacteriocin transport. 2020. mBio. 11 (2): e03230-19.

Böttcher T, Sieber S.  $\beta$ -Lactones as privileged structures for the active-site labeling of versatile bacterial enzyme classes. 2008. Angewandte Chemie. 47: 4600-4603.

Cascales E, Buchanan SK, Duché D, Kleanthous C, Lloubès R, Postle K, Riley M, Slatin S, Cavard D. Colicin biology. 2007. Microbiology and Molecular Biology Reviews. 71 (1): 158-229.

Caulier S, Nannan C, Gillis A, Licciardi F, Bragard C, Mahillon J. Overview of the antimicrobial compounds produced by members of the *Bacillus subtilis* group. 2019. Frontiers in Microbiology.10: 302.

Chanos P, Mygind T. Co-culture inducible bacteriocin production in lactic acid bacteria. 2016. Applied Microbiology and Biotechnology. 100: 4297–4308.

Chavali A, Rhee S. Bioinformatics tools for the identification of gene *clusters* that biosynthesize specialized metabolites.2018. Briefings in Bioinformatics. 19(5): 1022–1034.

de Jong A, van Hijum S, Bijlsma J, Kok J, Kuipers O. BAGEL: a web-based bacteriocin genome mining tool. 2006. Nucleic Acids Research. 34: W273–W279.

Dingemans J, Ghequire M, Craggs M, De Mot R, Cornelis P. Identification and functional analysis of a bacteriocin, pyocin S6, with ribonuclease activity from a *Pseudomonas aeruginosa* cystic fibrosis clinical isolate. 2016. Microbiology Open. 5 (3): 413–423.

Dolejšová T, Sokol A, Bosák J, Šmajš D, Konopásek I, Mikušová G, Fišer R. Colicin U from *Shigella boydii* forms voltage-dependent pores. 2019. Journal of Bacteriology. 201(24): e00493-19.

Duquesne S, Destoumieux-Garzón D, Peduzzi J, Rebuffat S. Microcins, gene-encoded antibacterial peptides from enterobacteria. 2007. *Natural Products*. 24: 708–734.

Felnagle E, Jackson E, Chan Y, Podevels A, Berti A, McMahon M, Thomas M. Nonribosomal peptide synthetases involved in the production of medically relevant natural products. 2008. *Molecular Pharmaceutics*. 5(2): 191–211.

Genilloud O. Natural products discovery and potential for new antibiotics. 2019. *Current Opinion in Microbiology*. 51: 81–87.

Ghazaryan L, Tonoyan L, Ashhab A, Soares M, Gillor O. The role of stress in colicin regulation. 2014. *Archives of Microbiology*. 196: 753–764.

Ghequire M, Mot R. Ribosomally encoded antibacterial proteins and peptides from *Pseudomonas*. 2014. *FEMS Microbiology Reviews*. 38: 523–568

Gillor O, Nigro L, Riley N. Genetically engineered bacteriocins and their potential as the next generation of antimicrobials. 2005. *Current Pharmaceutical Design*. 11: 1067-1075.

Gillor O, Kirkup BC, Riley MA. Colicins and microcins: the next generation antimicrobials. 2004. *Advances in Applied Microbiology*. 54: 129–146.

Gross H, Stockwell VO, Henkels MD, Nowak-Thompson B, Loper JE, Gerwick W. The genomisotopic approach: A systematic method to isolate products of orphan biosynthetic gene *clusters*. 2007. *Chemistry & Biology*. 14 (1): 53-63.

Gu Y, Ma YN, Wang J, Xia Z, Wei HL. Genomic insights into a plant growthpromoting *Pseudomonas koreensis* strain with cyclic lipopeptide-mediated antifungal activity. 2020. *Microbiology Open*. 9(9): e1092.

Gulder TA, Moore BS. Salinosporamide natural products: Potent 20 S proteasome inhibitors as promising cancer chemotherapeutics. 2010. *Angewandte Chemie*. 49: 9346-9367.

Hamdache A, Lamarti A, Aleu J, Collado I. Non-peptide metabolites from the genus *Bacillus*. 2011. *Journal of Natural Products*. 74(4): 893–899.

Holden V, Breen P, Houle S, Dozois C. *Klebsiella pneumoniae* siderophores induce inflammation, bacterial dissemination, and HIF-1 $\alpha$  stabilization during pneumonia. 2016. *mBio*. 7(5): e01397-16.

Hultberg M, Alsberg T, Khalil S, Alsanius B. Suppression of disease in tomato infected by *Pythium ultimum* with a biosurfactant produced by *Pseudomonas koreensis*. 2010. *BioControl*. 55: 435 - 444.

Jakes K. The Colicin E1 TolC Box: Identification of a domain required for colicin E1 cytotoxicity and TolC binding. 2016. *Journal of Bacteriology*. 199(1): e00412-16.

Jeong H, Rha E, Kim H, Lee S. Complete genome sequence of the soil bacterium *Pseudomonas kribbensis* strain 46-2T. 2018. *Microbiology Resource Announcements*. 7(21): e01161-18.

Katz M, Hover B, Brady S. Culture-independent discovery of natural products from soil metagenomes. 2016. *Journal of Industrial Microbiology and Biotechnology*. 43: 129–141.

Keller N. Fungal secondary metabolism: regulation, function and drug discovery. 2019. *Nature Reviews. Microbiology*. 17(3): 167–180.

Koumoutsi A, Chen X, Henne A, Liesegang H, Hitzeroth G, Franke P, Vater J, Borriss R. Structural and functional characterization of gene *clusters* directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. 2004. *Journal of Bacteriology*. 186(4): 1084–1096.

Kreitler D, Gemmell E, Schaffer J, Wencewicz T. The structural basis of N-acyl- $\alpha$ -amino- $\beta$ -lactone formation catalyzed by a nonribosomal peptide synthetase. 2019. *Nature Communications*. 10: 3432

Langa S, Arqués JL, Medina M, Landete JM. Coproduction of colicin V and lactic acid bacteria bacteriocins in *Lactococci* and *Enterococci* strains of biotechnological interest. 2017. Journal of Applied Microbiology. 122 (5): 1159-1167.

Le Govic Y, Papon N, Le Gal S, Bouchara J, Vandeputte P. Non-ribosomal peptide synthetase gene clusters in the human pathogenic fungus *Scedosporium apiospermum*. 2019. Frontiers in Microbiology. 10: 2062.

Lin H, Hu S, Liu R, Chen P, Ge C, Zhu B, Guo L. Genome sequence of *Pseudomonas koreensis* CRS05-R5, an antagonistic bacterium isolated from rice paddy field. 2016. Frontiers in Microbiology. 7: 1756.

Lozano GL, Bravo JI, Handelsman J. Draft genome sequence of *Pseudomonas koreensis* CI12, a *Bacillus cereus* “hitchhiker” from the soybean rhizosphere. 2017. Genome Announcements. 5 (26): e00570-17.

Lozano G, Park H, Bravo J, Armstrong E, Denu J, Stabb E, Broderick N, Crawford J, Handelsman J. Bacterial analogs of plant tetrahydropyridine alkaloids mediate microbial interactions in a rhizosphere model system. 2019. Applied and Environmental Microbiology. 85(10): e03058-18.

McCaughey L, Josts I, Grinter R, White P, Byron O, Tucker N, Matthews J, Kleanthous C, Whitchurch C, Walker D. Discovery, characterization and *in vivo* activity of pyocin SD2, a protein antibiotic from *Pseudomonas aeruginosa*. 2016. Biochemical Journal. 473(15): 2345-58.

Meade E, Slattery MA, Garvey M. Bacteriocins, potent antimicrobial peptides and the fight against multi drug resistant species: Resistance is futile. 2020. Antibiotics. 85(10): e03058-18.

Medina E, Pieper DH. Tackling threats and future problems of multidrug-resistant bacteria. 2016. Current Topics in Microbiology and Immunology. 398: 3-33.

Miceli de Farias F, Silva F, Nascimento de Sousa S, Salustiano SL, Lemos M, Mattos R, Carmo de Freire M. Draft genome sequence of *Enterococcus faecium* E86, a strain producing broad-spectrum antimicrobial peptides: Description of a novel bacteriocin immunity protein and a novel sequence type. 2019. Journal of Global Antimicrobial Resistance. 17: 195-197.

Miller B, Gulick A. Structural biology of nonribosomal peptide synthetases. 2016. Methods in Molecular Biology. 1401: 3–29.

Milshteyn A, Schneider J, Brady S. Mining the metabiome: identifying novel natural products from microbial communities. 2014. Chemistry & Biology. 21(9): 1211–1223.

Moon C, Zhang X, Matthijs S, Schäfer M, Budzikiewicz H, Rainey P. Genomic, genetic and structural analysis of pyoverdine-mediated iron acquisition in the plant growth-promoting bacterium *Pseudomonas fluorescens* SBW25. 2008. BMC Microbiology. 8: 7.

Morehead MS, Scarbrough C. Emergence of Global Antibiotic Resistance. 2018. Primary Care. 45 (3): 467-484.

Nikolouli K, Mossialos D. Bioactive compounds synthesized by non-ribosomal peptide synthetases and type-I polyketide synthases discovered through genome-mining and metagenomics. 2012. Biotechnology Letters. 34: 1393-1403.

Parquet M, Savage K, Allan D, Davidson R, Holbein B. Novel iron-chelator DIBI inhibits *Staphylococcus aureus* growth, suppresses experimental MRSA infection in mice and enhances the activities of diverse antibiotics in vitro. 2018. Frontiers in Microbiology. 9: 1811.

Poey ME, Azpiroz F, Lavin. Comparative analysis of chromosome-encoded microcins. 2006. Antimicrobial Agents and Chemotherapy. 50: 1411–1418.

Radaic A, de Jesus MB, Kapila YL. Bacterial anti-microbial peptides and nano-sized drug delivery systems: The state of the art toward improved bacteriocins. 2020. Journal of Controlled Release. 321: 100-118.

Ringel M, Brüser T. The biosynthesis of pyoverdines. 2018. *Microbial Cell*. 5 (10): 424-437.

Robinson SL, Wackett LP. Rings of power: Enzymatic routes to  $\beta$ -lactones. 2019. *Chemistry and Biology*. 4: 323-345.

Roh E, Park T, Kim M, Lee S, Ryu S, Chang-Sik O, Rhee S, Kim D, Park B, Heu S. Characterization of a new bacteriocin, carocin D, from *Pectobacterium carotovorum* subsp. *carotovorum* Pcc21. 2010. *Applied and Environmental Microbiology*. 76(22): 7541–7549.

Rudolf JD, Chang CY. Terpene synthases in disguise: enzymology, structure, and opportunities of non-canonical terpene synthases. 2020. *Natural Product Reports*. 37 (3): 425–463.

Satonyo G, Valencia-Cantero E, Orozco M, Peña J, Farias F. Role of siderophores in antagonistic activity of *Pseudomonas fluorescens* ZUM80 against plant fungi. 2009. *Terra Latinoamericana*. 28: 53-60.

Scott T, Batey S, Wiencek P, Chandra G, Alt S. Immunity-guided identification of threonyl-tRNA synthetase as the molecular target of obafluorin, a  $\beta$ -lactone antibiotic. 2019. *ACS Chemical Biology*. 12: 2663–2671.

Searle L, Méric G, Porcelli I, Sheppard S, Lucchini S. Variation in siderophore biosynthetic gene distribution and production across environmental and faecal populations of *Escherichia coli*. 2015. *Plos One*. 10(3): e0117906.

Sélem-Mojica N, Aguilar C, Gutiérrez-García K, Martínez-Guerrero CE, Barona-Gómez F. EvoMining reveals the origin and fate of natural product biosynthetic enzymes. 2019. *Microbial Genomics*. 5(12): e000260.

Sharp C, Bray J, Housden NG, Maiden MCJ, Kleanthous C. Diversity and distribution of nuclease bacteriocins in bacterial genomes revealed using Hidden Markov Models. 2017. *Plos Computational Biology*. 13(7): e1005652.

Skaff D, Ramyar K, McWhorter W, Barta M, Geisbrecht B, Miziorko H. Biochemical and structural basis for inhibition of *Enterococcus faecalis* hydroxymethylglutaryl-CoA synthase, *mvaS*, by hymeglusin 2012. *Biochemistry*. 51 (23): 4713–4722.

Simons A, Alhanout K, Duval RE. Bacteriocins, antimicrobial peptides from bacterial origin: Overview of their biology and their impact against multidrug-resistant bacteria. 2020. *Microorganisms*. 8(5): 639.

Subedi D, Kumar A, Willcox M. Overview of mechanisms of antibiotic resistance in *Pseudomonas aeruginosa*: an ocular perspective. 2018. *Clinical and Experimental Optometry*. 101: 162-171.

Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. 2018. *The Lancet Infectious Diseases*. 18 (3): 318-327.

Trapet P, Avoscan L, Klinguer A, Pateyron S, Citerne S, Chervin C, Mazurier S, Lemanceau P, Wendehenne D, Besson-Bard A. The *Pseudomonas fluorescens* siderophore pyoverdine weakens *Arabidopsis thaliana* defense in favor of growth in iron-deficient conditions. 2016. *Plant Physiology*. 171(1): 675–693.

Turano H, Gomes F, Domingos R, Degenhardt M, Oliveira C, Garratt R, Lincopan N, Netto L. Molecular structure and functional analysis of Pyocin S8 from *Pseudomonas aeruginosa* reveals the essential requirement of a glutamate residue in the H-N-H motif for DNase activity. 2020. *Journal of Bacteriology*. 202(21): e00346-20.

Walker D, Georgiou T, Pommer A, Walker D, Moore G, Kleanthous C, James R. Mutagenic scan of the H-N-H motif of colicin E9: implications for the mechanistic enzymology of colicins, homing enzymes and apoptotic endonucleases. 2002. *Nucleic Acids Research*. 30(14): 3225–3234.

Ward A, Allenby N. Genome mining for the search and discovery of bioactive compounds: the *Streptomyces* paradigm. 2018. *FEMS Microbiology Letters*. 365: 24.

Wyckoff E, Allred B, Raymond K, Payne S. Catechol siderophore transport by *Vibrio cholerae*. 2015. *Journal of Bacteriology*. 197(17): 2840–2849.

Xu L, Ye KX, Dai WH, Sun C, Xu LH, Han BN. Comparative genomic insights into secondary metabolism biosynthetic gene *cluster* distributions of marine *Streptomyces*. 2019. *Marine Drugs*. 17(9): 498.

Yan Q, Lopes L, Shaffer T, Kidarsa T, Vining O, Philmus B, Song C, Stockwell V, Raaijmakers J, McPhail K, Andreote F, Chang C, Loper J. Secondary metabolism and interspecific competition affect accumulation of spontaneous mutants in the GacS-GacA regulatory system in *Pseudomonas protegens*. 2018. *mBio*. 9(1): e01845-17.

Zakharov S, Cramer W. Colicin crystal structures: pathways and mechanisms for colicin insertion into membranes. 2002. *Biochimica et Biophysica Acta (BBA) Biomembranes*. 1565: 333-346.

Zheng S, Sonomoto K. Diversified transporters and pathways for bacteriocin secretion in Gram-positive bacteria. 2018. *Applied Microbiology and Biotechnology*. 102 (10): 4243-53.

Ziemert N, Alanjary M, Weber T. The evolution of genome mining in microbes a review. 2016. *Natural Product Reports*. 33 (8): 988-1005.

Zihalirwa P, Argüelles A, Franzil L, Steels S, Ongena M. Stimulation of fengycin-type antifungal lipopeptides in *Bacillus amyloliquefaciens* in the presence of the maize fungal pathogen *Rhizomucor variabilis*. 2017. *Frontiers in Microbiology*. 8: 850.