



Muzlera, Andrés

Caracterización bioquímica, genética, regulatoria y funcional de lipopéptidos en la rizobacteria biocontrol *Pseudomonas protegens* CHA0



Esta obra está bajo una Licencia Creative Commons Argentina.
Reconocimiento - Compartir Igual 2.5
<https://creativecommons.org/licenses/by-sa/2.5/ar/>

Documento descargado de RIDAA-UNQ Repositorio Institucional Digital de Acceso Abierto de la Universidad Nacional de Quilmes de la Universidad Nacional de Quilmes

Cita recomendada:

Muzlera, A. (2022). *Caracterización bioquímica, genética, regulatoria y funcional de lipopéptidos en la rizobacteria biocontrol Pseudomonas protegens CHA0. (Tesis de doctorado). Universidad Nacional de Quilmes, Bernal, Argentina. Disponible en RIDAA-UNQ Repositorio Institucional Digital de Acceso Abierto de la Universidad Nacional de Quilmes <http://ridaa.unq.edu.ar/handle/20.500.11807/3900>*

Puede encontrar éste y otros documentos en: <https://ridaa.unq.edu.ar>

Caracterización bioquímica, genética, regulatoria y funcional de lipopéptidos en la rizobacteria biocontrol *Pseudomonas protegens* CHA0

TESIS DOCTORAL

Andrés Muzlera

andres.muzlera@gmail.com

Resumen

La siguiente tesis se enfoca en el estudio de la regulación y expresión de lipopéptidos en rizobacteria *Pseudomonas protegens* CHA0. Las orfamidas en particular, son lipopéptidos cíclicos responsables de diversos fenotipos como el biocontrol y la motilidad por swarming entre otros. En este trabajo se describen dos genes que están regulados postranscripcionalmente por la cascada Gac-Rsm, que a su vez están catalogados como posibles reguladores transcripcionales. Se proponen ambos genes como activadores del cluster biosintético de orfamidas, siendo así el vínculo regulatorio entre la cascada Gac-Rsm y la expresión de orfamidas en CHA0. Complementariamente se estudió el impacto de las orfamidas en la competitividad en la colonización radicular en trigo.

A partir de los trabajos realizados con el objetivo de identificar los lipopéptidos producidos por *P. protegens* CHA0, se decidió ampliar el análisis estudiando metabolitos secundarios similares en aislamientos de *Pseudomonas* no caracterizados. El estudio de metabolitos especializados bacterianos propone un desafío significativo, ya que no existe un flujo de trabajo único y establecido. En este trabajo se propone la aplicación del estudio de metabolitos tipo lipopéptidos a partir de espectrometría de masas en tándem, acoplado al análisis por redes moleculares.

Por último, y con el objetivo de profundizar la caracterización de nuevos aislamientos, se decidió desarrollar una metodología para la búsqueda de los mejores marcadores taxonómicos para realizar filiación genética por análisis multi locus de secuencia (MLSA) para el género *Pseudomonas*. A partir de la secuencia de un concatenado de tres regiones internas de genes conservados de aproximadamente 850pb, se logró obtener una discriminación filogenética similar a la obtenida por la metodología hoy aceptada como *gold estándar*, el análisis de identidad promedio de nucleótido (ANI).

Laboratorio de Fisiología y Genética de Bacterias Beneficiosas
para Plantas
Centro de Bioquímica y Microbiología del Suelo

Caracterización bioquímica, genética, regulatoria y
funcional de lipopéptidos en la rizobacteria biocontrol
Pseudomonas protegens CHA0

Autor

Lic. Andrés Muzlera

Director

Dr. Claudio Valverde

Co-Director

Dr. Patricio Sobrero

Tesis para optar por el título de Doctor en Ciencia y Tecnología de la Universidad Nacional de
Quilmes



Índice de contenidos

Agradecimientos	7
Abreviaturas	8
Capítulo 1: La cascada Gac-Rsm controla a nivel post-transcripcional la expresión de reguladores transcripcionales del operón para la biosíntesis de orfamidas en <i>Pseudomonas protegens</i> CHA0	
1.1 Resumen.....	10
1.2 Introducción	10
1.2.1 El género <i>Pseudomonas</i> y su relación con el biocontrol	10
1.2.2 Metabolitos especializados producidos por pseudomonas relacionados con el biocontrol	11
1.2.2.1 Fenazinas	11
1.2.2.2 Pirrolnitrina	12
1.2.2.3 Rizoxinas	12
1.2.2.4 Floroglucinoles	13
1.2.2.5 Acido cianhídrico	13
1.2.2.6 Toxoflavina	14
1.2.2.7 Sideróforos	14
1.2.2.8 Lipopéptidos (LPs)	15
1.2.3 <i>Pseudomonas protegens</i> CHA0 como modelo de biocontrol	17
1.2.4 Regulación genética	18
1.2.5 Distintos niveles de regulación.....	18
1.2.6 Regulación por ARNs codificados en <i>cis</i> . Termo-sensores y <i>riboswitchs</i>	19
1.2.7 Regulación por ARNs codificados en <i>trans</i>	21
1.2.7.1 Mecanismos dependientes de Hfq	21
1.2.7.2 Reguladores globales de la familia CsrA.....	22
1.2.8 Cascada regulatoria Gac-Rsm en <i>Pseudomonas protegens</i> CHA0	25
1.2.9 Experimentos previos, screening bioinformático y fusiones reporteras	26
1.3 Materiales y métodos	30
1.3.1 Cepas bacterianas, plásmidos y oligonucleótidos.....	30

1.3.2	Colapso de gota.....	31
1.3.3	Enriquecimiento de lipopéptidos a partir de sobrenadantes de cultivo.....	31
1.3.4	Análisis y purificación de lipopéptidos por HLPC	31
1.3.5	Ensayo de actividad β -galactosidasa	32
1.3.6	Transformación de <i>E. coli</i> por electroporación	32
1.3.7	Construcción de fusiones reporteras	32
1.3.8	Transferencia de plásmidos a pseudomonas por conjugación triparental	33
1.3.9	Cinéticas de expresión con reporteros fluorescentes.....	33
1.3.10	Expresión y purificación de las proteínas RsmA/E-his6	33
1.3.11	Ensayos de cambio en la movilidad electroforética.....	34
1.4	Resultados y discusión	34
1.4.1	Estudio de la expresión de los genes orfUR y orfDR utilizando reporteros fusionados a <i>gfp</i> y <i>DsRed</i>	34
1.4.2	Estudio sobre regulación directa de las proteínas represoras RsmA/E sobre los reguladores transcripcionales <i>orfUR</i> y <i>orfDR</i>	37
1.4.3	Purificación de las proteínas RsmA y RsmE.....	38
1.4.4	Vínculo regulatorio entre la cascada Gac-Rsm y la biosíntesis de lipopéptidos en <i>P. protegens</i> CHA0.....	41
1.4.4.1	Optimización de la producción de surfactantes.....	42
1.4.4.2	Purificación e identificación de lipopéptidos de <i>P. protegens</i>	43
1.4.5	Modelo regulatorio de la producción de orfamidas por la cascada Gac-Rsm	46
1.5	Conclusión	47
1.6	Bibliografía	48

Capítulo 2: Papel de las orfamidas en la persistencia en el suelo y la competitividad para la colonización radicular en *Pseudomonas protegens* CHA0

2.1	Resumen.....	55
2.2	Introducción	55
2.2.1	Rol de los lipopéptidos en la naturaleza	55
2.2.1.1	Lipopéptidos involucrados en la motilidad	55
2.2.1.2	Lipopéptidos involucrados en la formación de biopelículas	56

2.3	Materiales y métodos	57
2.3.1	Cepas bacterianas y plásmidos	57
2.3.2	Generación de variantes marcadas.....	58
2.3.3	Generación de vector pME7408 (<i>prsmZ-gfp; orfDR'-mCherry</i>).....	59
2.3.4	Persistencia en suelo natural	59
2.3.5	Inoculación	61
2.3.6	Recolección de muestras.....	61
2.3.7	Recuentos en placa	62
2.3.8	Microscopia de fluorescencia.....	62
2.3.9	Análisis estadístico	62
2.4	Resultados y discusión	63
2.4.1	<i>Fitness</i> de <i>P. protegens</i> silvestre (CHA0) en suelo natural.....	63
2.4.2	Primera etapa: Persistencia en suelo libre.....	63
2.4.3	Segunda etapa: Competencia por la colonización radicular	65
2.4.4	Expresión de <i>ofa</i> en la rizosfera.	68
2.5	Conclusiones.....	72
2.6	Bibliografía	73

Capítulo 3: Aplicación de *molecular networking* para la exploración de metabolitos especializados extracelulares de aislamientos autóctonos de *Pseudomonas*

3.1	Resumen.....	77
3.2	Introducción	77
3.2.1	Colección de aislamientos autóctonos de la región agrícola núcleo Argentina. Proyecto BIOSPAS.....	77
3.2.2	Análisis multilocus de secuencia	79
3.2.3	Los productos naturales como objeto de estudio.....	80
3.2.4	Espectrometría de masas aplicadas al descubrimiento de productos naturales	80
3.2.5	Desreplicación como metodología para la identificación rápida de compuestos naturales	82
3.2.6	<i>Molecular Networking</i>	83
3.2.7	<i>Feature base molecular networking</i>	85

3.3	Materiales y métodos	87
3.3.1	PCR y secuenciación	87
3.3.2	Análisis multi locus de secuencia (MLSA).....	88
3.3.3	Antagonismo in vitro contra <i>Pythium ultimum</i>	88
3.3.4	<i>Molecular Networking</i>	88
3.4	Resultados y discusión	89
3.4.1	Extracción e identificación de compuestos surfactantes en aislamientos autóctonos de pseudomonas.....	89
3.4.1.1	RMAN4	90
3.4.1.2	RBAN4.....	90
3.4.1.3	SMMP3	92
3.4.1.4	SVMP4 y SBMP6	98
3.5	Discusión	100
3.6	Conclusión	101
3.7	Bibliografía	102

Capítulo 4: Diseño de un método económico para la tipificación molecular de aislamientos del género *Pseudomonas* con poder de resolución comparable a la filogenómica

4.1	Resumen.....	107
4.2	Introducción	107
4.2.1	El género <i>Pseudomonas</i>	107
4.2.2	Metodologías aplicadas para la taxonomía en el género <i>Pseudomonas</i> en la historia	108
4.2.3	Metodologías aplicadas para la taxonomía en el género <i>Pseudomonas</i> en la actualidad	109
4.3	Materiales y métodos	111
4.3.1	Cepas bacterianas	111
4.3.2	Herramientas y softwares bioinformáticos.....	111
4.3.3	Diseño de oligonucleótidos para PCR.....	111
4.3.4	Amplificaciones por PCR.....	112
4.4	Resultados y discusión	112

4.4.1	Curado de base de datos y obtención de genes <i>core</i> (CG).....	112
4.4.2	Evaluación del potencial resolutivo de cada amplicón	116
4.4.3	Diseño y validación de oligonucleótidos	119
4.5	Conclusión	122
4.6	Bibliografía	123

Abreviaturas

AB: Antibiótico

ADN: ácido desoxirribonucleico

ANI: *average nucleotide identity*

AP: ampicilina

ARN: ácido ribonucleico

CG: *gen core*

Cm: Cloranfenicol

DAPG: 2,4-Diacetylphloroglucinol

dNTP's: desoxi-ribonucleótidos trifosfato

DO600: Densidad óptica

FBMN: *Feature base Molecular Networking*

GNPS: *Global Natura Products Social Molecular Networking*

Gm: Gentamicina

HCN: Ácido cianhídrico

Km: Kanamicina

LP: lipopéptido

MLSA: *multilocus sequence analysis*

MS: espectrometría de masas

MN: *Molecular Networking*

NA: Agar nutritivo (*Nutrient agar*)

NYB: Caldo nutritivo (*Nutrient yeast broth*)

Tc: Tetraciclina

Agradecimientos

A lo largo de estos años me fui cruzando con mucha gente, que de no ser por ellos este trabajo no hubiese sido lo que es, y tal vez no hubiese existido en lo absoluto. Es por eso que intentaré en algunas palabras demostrar mi agradecimiento.

En primer lugar a mi gente, la de mi labo, que son unos cuantos que llegaron, se fueron o todavía están acá. Fue gracias a ellos que aprendí y me reí en igual proporción, gracias a todos por tan buenos momentos! Sin embargo, hay algunos individuos que por cuestiones de cotidianeidad hemos compartido más momentos y creo que ameritan unas líneas de agradecimiento:

A Fede por ser El gran compañero que es, siempre dispuesto a todo, nunca un no, el famoso amigo de fierro. 10 horas seguidas viéndonos las jetas en el labo pero igual siempre teníamos de que hablar, tanto que casi nos pasábamos de estación.

A Melani, la pequeña, capaz de entrar en casi cualquier objeto cuadrado. Una gran amiga y compañera que no puedo mencionar sin recordar miles de grandes y graciosos momentos como el debate diario de cuando comer una maldita pera!

A Bel, por su buena onda, siempre con disparates, imposible saber con que va a salir pero seguro que es para reírse, gracias por contaminar el Spotify de la gente y por pasearme por todo Quilmes en busca de un técnico. Siempre la juntada es mejor cuando estas.

A Beti, gran compañera de la cual aprendí muchísimo: sobre todo organización, y buenas prácticas. Gracias por tu amistad y las mejores cosas dulces que ha visto la UNQ.

A Anto, siempre con calma, sabiendo que es lo que hay que hacer. Gracias por enseñarme tanto, más de lo que imaginas seguramente. Un gran amigo y un excelente compañero.

A Pato, por su amistad, pero antes que nada por su dirección. Gracias por las buenas ideas, la disposición, la inquietud, y por siempre de alguna forma u otra marcar el camino.

A Mati, por las risas, los consejos y los mejores asados de la historia. Un gran amigo que me deja la UNQ, loco como el solo, que sin proponérselo me ha enseñado tanto.

Y por último a Claudio, por darme la oportunidad de formar parte de este laboratorio, y por ser el responsable de sentar las bases de mi formación, siempre enseñando con el ejemplo. Por su infinita paciencia y voluntad.

A mi gente de la cursada, que ya vamos 10 años (con varios algunos más), yendo por el mismo camino, por momentos más cercanos y por momentos no tanto, pero siempre de alguna manera juntos. A Yami, Juli, Cami, Jair, Viky y Juli. Gracias por los consejos, el aguante, los puchos, los asados!

A la gente de mi ex labo 10, aunque nunca sentí que realmente me fui, y eso fue gracias a ustedes. A Yami, por ser mi primer contacto con la investigación, y desde el primer día hasta hoy siempre empujándome para adelante. A Toti por ser el amigazo que es, por los asados los buenos días de pileta y las buenas charlas.

A mis amigos de la vida, a Agus por ser siempre el primero en estar, para lo bueno y lo malo. Gracias por las charlas interminables y todo el aguante! A Leito y Tute, lo mejor que me dejó la escuela técnica por lejos, gracias por la mejor de las amistades.

Y por último a mi familia, a mi vieja por ser la responsable n°1 de todo esto, diga lo que diga me voy a quedar corto. Gracias por las comidas, el ánimo, el empujón, la lista es interminable. Tu laburo de madre, siempre constante y sutil, por momentos pasa desapercibido y no recibe a veces el agradecimiento que se merece, como decís vos: gracias, gracias, gracias. A mi viejo por ser mi otro pilar y por inculcarme de pequeño esa curiosidad exasperante por las cosas, que definitivamente con los años se transformó en mi profesión. A mi hermano, mi persona favorita en este mundo, mi mejor amigo, cerca o lejos siempre a mi lado. A Naty, mi todo, gracias por compartir tu vida conmigo y dejarme compartir la mía con vos. Sos la persona que me conoce, que me entiende, que me acompaña, que me hace reír, que me sostiene, en especial este último tramo. Gracias por hacerme una mejor persona. Y por último, a Emilia, que aunque todavía no hayas respirado tu primera bocanada de aire, ya sos lo más importante de mi vida.

Capítulo 1

La cascada Gac-Rsm controla a nivel post-transcripcional la expresión de reguladores transcripcionales del operón para la biosíntesis de orfamidas en *Pseudomonas protegens* CHA0.

1.1 Resumen

Previamente en el laboratorio se realizó una búsqueda para identificar nuevos posibles ARNm bajo el control de la cascada Gac-Rsm de *Pseudomonas protegens* CHA0 que podrían servir para dilucidar funciones aún desconocidas involucradas en el biocontrol de patógenos de plantas y/o en procesos celulares requeridos para el *fitness* en ambientes naturales. Esta búsqueda consistió en un *screening* computacional que concluyó en un catálogo de 43 posibles nuevos ARNm *target* de la cascada Gac-Rsm. Este capítulo describe específicamente el estudio de dos ARNm catalogados como posibles objetivos (*targets*) regulatorios, que están físicamente vinculados al grupo de genes biosintéticos de orfamidas y cuyos marcos de lectura abiertos predichos presentan homología con los reguladores transcripcionales de tipo LuxR homólogos de lipopéptidos cíclicos en pseudomonas relacionadas. El posible control de la cascada Gac-Rsm sobre la producción de orfamidas se estudió mediante dos aproximaciones paralelas. Con un enfoque químico, se purificaron y analizaron los lipopéptidos de distintos mutantes isogénicos de CHA0 deficientes en genes de la cascada. Por otro lado, con una mirada molecular, se estudió la expresión mediante el uso de fusiones génicas reporteras para ambos reguladores transcripcionales, y se confirmaron interacciones *in vitro* utilizando proteínas y oligorribonucleótidos sintéticos. De esta manera demostramos que en la cepa CHA0, la producción de orfamidas depende estrictamente de la cascada Gac-Rsm y que ambos ARNm que codifican los reguladores transcripcionales están bajo el control directo de las proteínas RsmA/E. Estos resultados revelan un control jerárquico sobre la expresión de genes biosintéticos de orfamidas con el control transcripcional final subordinado al sistema regulador postranscripcional global Gac-Rsm.

1.2 Introducción

1.2.1 El género *Pseudomonas* y su relación con el biocontrol

La rizosfera alberga una variedad de microorganismos beneficiosos de los cuales las pseudomonas son muy apreciadas por su potencial biocontrolador y promotor del crecimiento vegetal. Las especies de *Pseudomonas* fluorescentes se consideran jugadores clave en la agricultura para la supresión efectiva de enfermedades en plantas asociadas con fitopatógenos fúngicos y bacterianos y para la promoción del crecimiento de las plantas. Algunas pseudomonas mejoran el crecimiento de las plantas al producir factores de crecimiento, solubilización de zinc, fósforo, potasio, degradación de materia orgánica y fijación de nitrógeno (Beneduzi et al. 2012; Gray and Smith 2005; Hayat et al. 2010). Se sabe que integrantes del género colonizan con éxito superficies y tejidos internos de raíces y tallos en altas densidades celulares, para su supervivencia y adaptación en diversos nichos ambientales (Welbaum et al. 2004). Además, las pseudomonas son capaces de sintetizar una batería de metabolitos especializados que le otorgan un significativo potencial biocontrolador. Esta capacidad se atribuye a la producción de muchos tipos de metabolitos con propiedades antimicrobianas, antimitóticas, nematocidas

(Haas and Défago 2005) y fitotóxicas (Bender et al. 1999). Además, muchas especies de *Pseudomonas* también son conocidas por la producción de sideróforos captadores de hierro, enzimas hidrolíticas extracelulares, compuestos orgánicos volátiles y hormonas promotoras del crecimiento vegetal. Muchos de estos compuestos se generan a través de complejas rutas metabólicas y están involucrados en la supresión e inhibición competitiva de patógenos de plantas (Al-Hinai et al. 2010; Shanmugaiah, Mathivanan, and Varghese 2010; Shahid et al. 2017).

Los mecanismos de los microorganismos benéficos para el control biológico de enfermedades en las plantas son sofisticados, pero principalmente incluyen efectos directos sobre los patógenos blanco. Los metabolitos especializados tienen actividades antibióticas que se dirigen directamente a los patógenos de las plantas y, por lo tanto, alteran su crecimiento y desarrollo aliviando aún más los síntomas causados sobre las plantas. Estos tipos de antagonismo directo juegan un papel importante en el control de patógenos de plantas. La capacidad de antibiosis de un microorganismo, es decir, su capacidad de producir metabolitos activos, se usa comúnmente para detectar agentes de control biológico (Pliego et al. 2011; Perneel et al. 2007). Las actividades antagónicas producidas por microorganismos benéficos para las plantas se han estudiado ampliamente. Por ejemplo, el potencial antagonista contra los hongos patógenos de plantas por parte de *Bacillus spp.* se atribuyen en gran medida a la producción de lipopéptidos (Cawoy et al. 2015).

1.2.2 Metabolitos especializados producidos por pseudomonas relacionados con el biocontrol

Las especies de *Pseudomonas* no se limitan a ser agentes potenciales de control biológico contra enfermedades en los cultivos (Höfte and Altier 2010), sino que también son grandes reservorios de genes para la biosíntesis de diversos metabolitos especializados (Gross and Loper 2009). Estos metabolitos están involucrados en la antibiosis, la formación de biopelículas u otras funciones indispensables para las propias pseudomonas, y estos metabolitos pueden desempeñar funciones en las interacciones entre especies, la adaptación ambiental y la autodefensa (Loper et al. 2012).

1.2.2.1 Fenazinas

Dentro de los metabolitos de interés producidos por pseudomonas podemos encontrar a las fenazinas, compuestos heterocíclicos de característicos colores brillantes. Las fenazinas constituyen un amplio grupo de compuestos con más de cincuenta derivados (Pierson et al. 2010). Desde su descubrimiento, la búsqueda de nuevos derivados de fenazina nunca se interrumpió y se han identificado diversos derivados con propiedades características. Se demostró que el ácido 5-metilfenazina-1-carboxílico (Figura 1) de *P. putida* PUW5 induce la detención del ciclo celular en G1, la apoptosis y la citotoxicidad selectiva hacia las líneas celulares de cáncer de pulmón (A549) y mama (MDA MB-231) (Kennedy et al. 2015). Mehnaz et al. (2013) informaron sobre la producción de 2,8-dihidroxifenazina a partir de *P. aurantiaca* PB-St2, el cual demostró una actividad antibacteriana moderada frente *Arthrobacter crystallopoietes*.

Cezairliyan et al. (2013) reportaron la toxicidad de la 1-hidroxifenazina y el ácido fenazina-1-carboxílico (Figura 1) frente a *C. elegans*. Además, la implementación de nuevas estrategias para la biosíntesis y el ensamblaje de anillos de derivados de fenazina para mejorar las actividades antimicrobianas y citotóxicas está emergiendo como una nueva herramienta biotecnológica. La síntesis de biomiméticos del ácido fenazina 1,6-dicarboxílico (Figura 1), los procedimientos en un solo recipiente (*one-pot*) para reacciones de múltiples componentes, la síntesis completa de estreptofenazina A, dermacozinas A y el aislamiento e identificación de fenazinas cepa-específicas, han ayudado a comprender la bioquímica de la fenazina en sus intrincadas rutas de biosíntesis (Guttenberger et al. 2017).

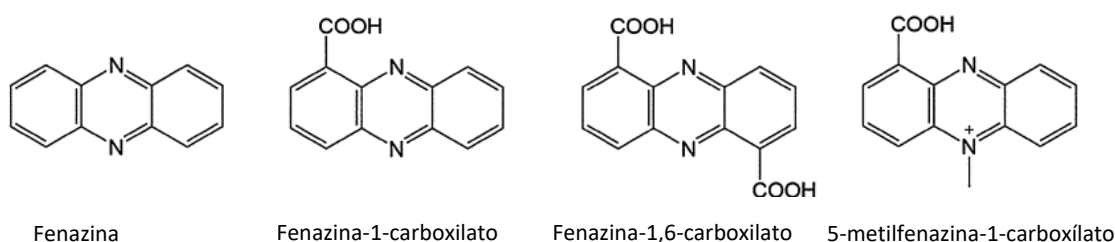


Figura 1. Estructura química de la fenazina y dos de sus derivados

1.2.2.2 Pirrolnitrina

La pirrolnitrina (Figura 2) es un metabolito secundario antifúngico halogenado, producido en general por cepas de *P. chlororaphis* subsp. *aureofaciens*, *P. chlororaphis* subsp. *chlororaphis*, *P. chlororaphis* subsp. *aurantiaca*, *P. fluorescens*, *P. protegens* y *P. putida* (Hashimoto y Hattori 1966). Su mecanismo antifúngico por el cual suprime el crecimiento de diferentes hongos patógenos de plantas fue reportado hace ya varios años y se basa en la inhibición de la cadena transportadora de electrones (Tripathi and Gottlieb 1969). Mas recientemente, se destacó su actividad nematocida contra *C. elegans* (Nandi et al. 2015).

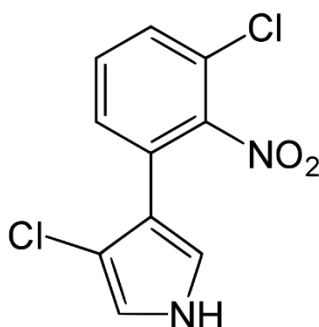


Figura 2. Estructura química de la pirrolnitrina

1.2.2.3 Rizoxinas

La rizoxina se identificó originalmente a partir de *Burkholderia rhizoxina* y *Rhizopus microsporus*, sin embargo, posteriormente se demostró que *P. protegens* Pf-5 tiene la capacidad de producir cinco análogos de rizoxina estructuralmente diferentes con fuertes actividades antifúngicas

contra dos importantes fitopatógenos, *Botrytis cinerea* y *Phytophthora ramorum* (Loper et al. 2008) y con actividad insecticida y letal contra la mosca de la fruta *Drosophila melanogaster* (Loper et al. 2016).

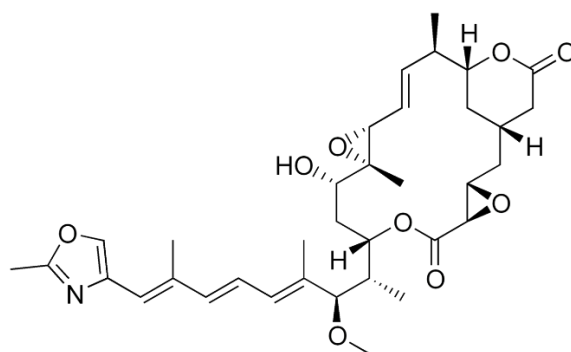


Figura 3. Estructura química de la rizoxina

1.2.2.4 Floroglucinoles

El 2,4-diacetilfloroglucinol (DAPG) (Figura 4) producido por *P. protegens* es el compuesto más conocido de una familia de moléculas relacionadas que incluye monoacetilfloroglucinol y productos de condensación del DAPG y monoacetilfloroglucinol. *P. protegens* además es capaz de producir pioluteorina y ácido cianhídrico, que junto con el DAPG causan daño a la membrana de *Pythium spp.* y es particularmente inhibitorio para las zoosporas de este oomiceto (Mavrodi et al. 2001; Maurhofer et al. 1994a). En altas concentraciones el DAPG puede ser fitotóxico (de Souza et al. 2003).

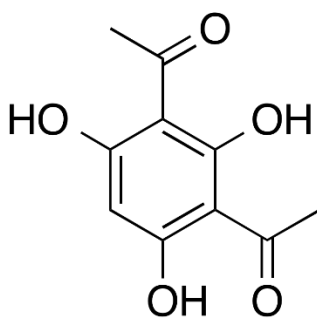


Figura 4. Estructura química del 2,4-diacetilfloroglucinol (DAPG)

1.2.2.5 Ácido cianhídrico

El ácido cianhídrico (HCN) es un compuesto volátil secretado por muchas especies de *Pseudomonas*, incluido el patógeno oportunista humano *P. aeruginosa* (Gilchrist et al. 2013). El HCN secretado juega un papel importante en la supresión del desarrollo de patógenos fúngicos como, por ejemplo, en el antagonismo contra *M. oryzae* 70-15. *P. protegens* CHA0 es capaz de inhibir el crecimiento de *M. oryzae*, mientras que su mutante deficiente en la producción de HCN pierde esta capacidad (Spence et al. 2014).

1.2.2.6 Toxoflavina

La toxoflavina (Figura 5) se descubrió a partir de especies de *Burkholderia* patógenas para plantas en 1934 (Van Veen y Mertens, 1934). La toxoflavina de *Burkholderia glumae* funciona como una fitotoxina que causa marchitez bacteriana en muchos cultivos, como tomate, sésamo, berenjena, pimiento picante, etc. (Jeong et al. 2003; Koh et al. 2011). Hace algunos años además, se demostró que *P. protegens* Pf-5 alberga un grupo de genes para la biosíntesis de toxoflavina, y un análisis de mutagénesis mostró que la toxoflavina contribuye a la actividad antagónica contra la bacteria patógena de plantas *P. syringae* DC3000 (Philmus et al. 2015).

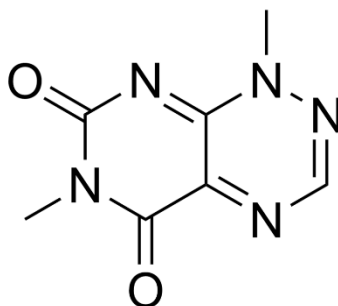


Figura 5. Estructura química de la toxoflavina

1.2.2.7 Sideróforos

Los sideróforos son moléculas que tienen la capacidad de unirse al hierro y transportarlo al interior de las células. En general, la biosíntesis de los sideróforos se lleva a cabo por sintasas de péptidos no ribosomales (NRPS) (Miethke, Marahiel, and reviews 2007; Hoegy, Mislin, and Schalk 2014). Las pseudomonas son potentes productoras de sideróforos y estos participan activamente en el control de patógenos de plantas (Buyer, Wright, and Leong 1986). Diversas especies de *Pseudomonas* pueden secretar sideróforos tipo pioverdina, compuestas por un cromóforo tipo dihidroquinolina unido a un resto peptídico (Schalk and Guillon 2013). Los sideróforos producidos por *P. aeruginosa* JAS-25 están involucrados en la actividad antagónica contra *Fusarium oxysporum* f. sp. *ciceri*, *Fusarium udum* y *Aspergillus niger* (Sulochana et al. 2014). El sideróforo purificado de *P. putida* A1 muestra actividad fungicida contra *Fusarium oxysporum* f. sp. *ciceri* y *Helminthosporium oryzae* (Boopathi, Rao, and enzymology 1999). Además, ciertos metabolitos tipo sideróforo son potentes elicitores para aumentar la inmunidad innata en las plantas. Por ejemplo, la pioquelina secretada por *P. aeruginosa* 7NSK2 tiene la capacidad de activar una resistencia sistema inducida frente a *Botrytis cinerea* en tomate (Audenaert et al. 2002).

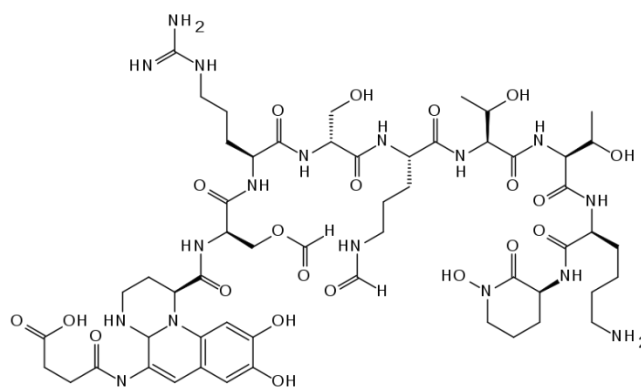


Figura 6. Estructura química de la pioverdina

1.2.2.8 Lipopéptidos (LPs)

Los lipopéptidos cíclicos (cLPs) están compuestos por un esqueleto peptídico cíclico unido a una cola de ácido graso. Los cLPs derivados de especies de *Pseudomonas* se pueden agrupar en varias clases según la longitud y la composición de aminoácidos del resto peptídico, y dentro de cada subclase, los residuos de ácidos grasos también muestran diferencias, por ejemplo, la longitud y la saturación de los ácidos grasos (Roongsawang, Washio, and Morikawa 2010). Los cLPs poseen propiedades hidrofóbicas e hidrofílicas y, por lo tanto, pueden interactuar fácilmente con diferentes membranas celulares y mostrar un amplio espectro de bioactividades contra bacterias, hongos, virus, células de mamíferos, protistas, entre otros. (Raaijmakers et al. 2010).

Dentro de la familia de los cLPs, varios de ellos son fundamentales para la motilidad y la formación de biopelículas. Mutantes con deficiencia en la producción de cLPs perdieron o mostraron una motilidad por *swarming* reducida y formación de biopelículas en comparación con su cepa salvaje (De Bruijn et al. 2007; D'aes et al. 2014). Además, el agregado de cLPs purificados en el medio de cultivo puede restaurar el *swarming* en cepas deficientes de cLPs (D'aes et al. 2014). Las particulares propiedades que otorgan los cLPs podrían ser relevantes para que las bacterias productoras mantengan funciones importantes en su entorno natural.

La biosíntesis de los cLPs, en *Pseudomonas*, está mediada por sintasas de péptidos no ribosomales (NRPS). Los péptidos no ribosomales, a diferencia de los derivados de la traducción de ARNm en los ribosomas, son ensamblados por las NRPS. Los módulos de las NRPS se pueden dividir en dominios catalíticos. Tres dominios son ubicuos en la síntesis de NRP y esenciales para la elongación de péptidos: los dominios responsables de la activación del aminoácido (dominios de adenilación (A)), de la propagación de la cadena peptídica en crecimiento (dominios de tiolación o proteína transportadora de péptidos (PCP)) y de la condensación de los aminoácidos (dominios de condensación (C)). Una cuarta unidad catalítica esencial de NRPS asociada con la liberación del producto es el dominio tioesterasa (TE). El dominio TE está ubicado en el módulo de terminación y cataliza la liberación de péptidos por hidrólisis o macrociclación (Strieker, Tanović, and Marahiel 2010). En la Figura 7 se puede observar un esquema reducido de una NRPS.

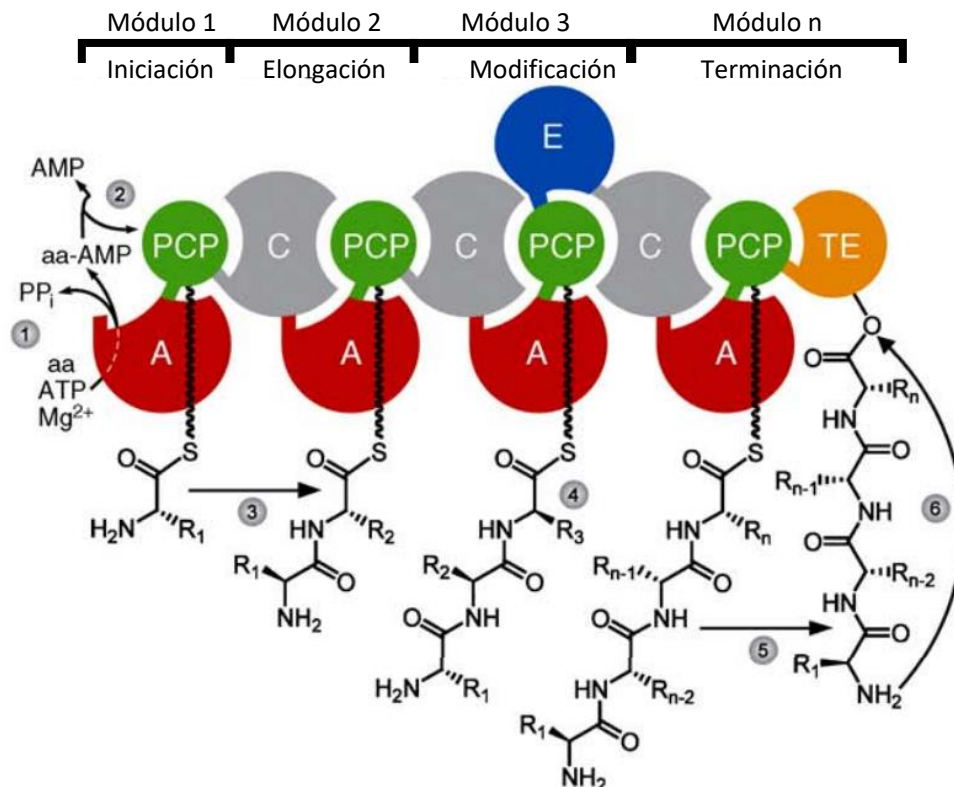


Figura 7. Mecanismo de síntesis de péptidos no ribosómicos (NRP) simplificado. (1) El aminoácido se activa como aminoacil-AMP por el dominio de adenilación. (2) Transferencia del aminoácido al dominio PCP. (3) Condensación de aminoácidos unidos a PCP. (4) Posibilidad de modificaciones de aminoácidos, por ejemplo, por dominios de epimerización. (5) Transesterificación de la cadena peptídica del PCP terminal al dominio TE. (6) Liberación de producto catalizada por TE por hidrólisis o macrociclación. El número de dominios y módulos de modificación total es muy variable. Modificado de (Strieker, Tanović, and Marahiel 2010).

Basándose en la longitud, la composición de la cola de ácidos grasos y el número, tipo y configuración de los aminoácidos en el resto peptídico, los LPs de *Pseudomonas* se clasificaron inicialmente en cuatro grupos principales: viscosina, anfisina, tolaasina y siringomicina (Raaijmakers, De Bruijn, and De Kock 2006). El grupo viscosina alberga LPs con 9 aminoácidos unidos en el extremo N, en la mayoría de los casos, al ácido 3-hidroxidecanoico (3-HDA). Este grupo incluye viscosina, massetolido A, pseudofominas, pseudodesminas, viscosinamida y el principio inductor de la línea blanca (WLIP) (De Bruijn et al. 2007; De Bruijn et al. 2008; Quail et al. 2002; Pedras et al. 2003; Nielsen et al. 1999). El grupo anfisina, compuesto por anfisina, artrofactina, lokisina, folipeptina y tensina, se componen de LPs con péptidos de 11 aminoácidos acoplados a 3-HDA (Koch et al. 2002; Sørensen et al. 2002; Roongsawang et al. 2003; Nielsen et al. 2000). El grupo de las tolaasinas, con los LPs tolasina, siringopeptina, corpeptina y sesillina es más diverso en cuanto a la composición y longitud de la cadena peptídica de sus LPs (19-25 aminoácidos) así como de la cola lipídica (3-HDA o ácido 3-hidroxioctanoico (3-HOA)). El grupo de la siringomicina, al igual que las viscosinas, están conformados por péptidos de 9 aminoácidos, sin embargo, se distinguen por la presencia de aminoácidos inusuales como el

ácido dideshidroaminobutírico (Dhb), el ácido 2,4-diaminobutírico (Dab) y el residuo de treonina clorada C-terminal.

Con la creciente disponibilidad de secuencias de genomas completos de pseudomonas, se han descubierto nuevos y diversos genes de biosíntesis de LPs, lo que llevó a una extensión sustancial de la clasificación química propuesta inicialmente por Raaijmakers, De Bruijn, and De Kock (2006). Por ejemplo, el análisis del genoma de *P. protegens* Pf-5 condujo a la identificación de las orfamidas (Figura 8), cLPs con 10 aminoácidos y una cola de ácido mirístico 3-hidroxi (3-HMA) (Gross et al. 2007). Posteriormente, se identificaron análogos estructurales de orfamidas en *Pseudomonas* CMR12a, una cepa aislada de la rizosfera del taro (*Colocasia esculenta*) (D'aes et al. 2014). Esta cepa también produce sessilina, un LPs que difiere en un solo aminoácido de la tolaasina, la toxina producida por el hongo patógeno *P. tolaasii* (D'aes et al. 2014). Hoy en día existen nuevos intentos de ordenar la gran variedad de lipopéptidos diferentes en familias, basados principalmente en la clasificación de sus NRPS (Girard, Höfte, and Mot 2020).

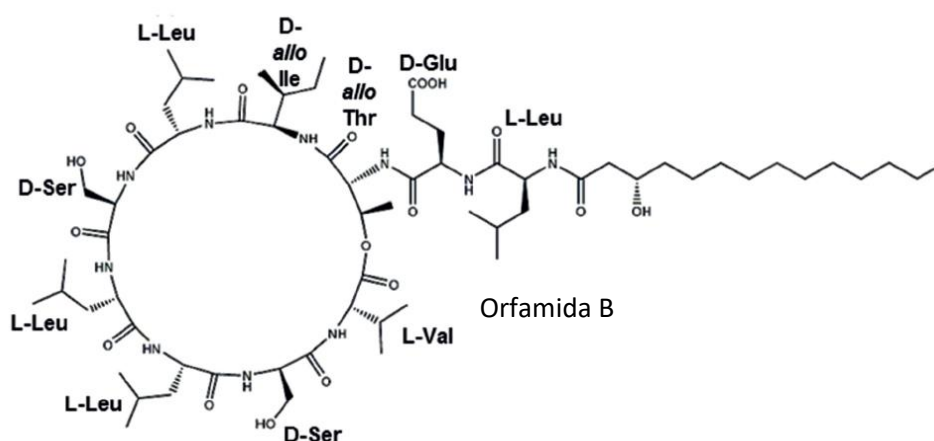


Figura 8. Estructura química de la orfamida B

1.2.3 *Pseudomonas protegens* CHA0 como modelo de biocontrol

Pseudomonas protegens CHA0 fue aislada de raíces de tabaco cultivadas en suelos supresivos en 1986, y la cepa mostró potencial para controlar la pudrición de raíz negra, causada por *Thielaviopsis basicola* (Berk. & Br.) Ferraris, en plantas de tabaco (Stutz, Défago, and Kern 1986). La especie *protegens* se caracterizó como una nueva especie hace ya más de 10 años, abandonando su antigua clasificación como *fluorescens*, siendo CHA0 la cepa tipo (Ramette et al. 2011). Esta cepa manifestó gran potencial biocontrolador en pepino y berro frente a la enfermedad causada por *Pythium ultimum* (Maurhofer et al. 1992; Péchy-Tarr et al. 2005; Valverde et al. 2003; Takeuchi, Yamada, and Haas 2012), en trigo contra la enfermedad causada por *Gaeumannomyces graminis* var. *tritici* (Keel et al. 1992), en tomate contra *Fusarium oxysporum* (Zuber et al. 2003), y mostró inducción de la resistencia sistémica en vid contra *Botrytis cinerea* (Verhagen et al. 2010), y en *Arabidopsis thaliana* contra el oomicete *Peronospora parasitica* (Iavicoli et al. 2003).

P. protegens CHA0 ha sido sujeto de estudio de numerosas investigaciones, principalmente por su capacidad de sintetizar metabolitos especializados que le confieren un fenotipo biocontrolador característico. Sin embargo, no es la única razón por la cual esta bacteria y otras de su género se han transformado en atractivos modelos de estudio. En CHA0 y otras pseudomonas, todos sus metabolitos relacionados con el biocontrol se encuentran bajo la regulación de una cascada global llamada Gac-Rsm, denominada así por sus siglas en inglés *global antibiotic and cyanide control* (Laville et al. 1992) y *regulator of secondary metabolism* (Blumer et al. 1999). Esta cascada se conforma de un sistema de dos componentes, GacS/GacA, ampliamente distribuido en bacterias (BarA/UvrY en *E. coli* (Pernestig, Melefors, and Georgellis 2001)), y un grupo de proteínas y ARNs no codificantes que actúan a nivel postranscripcional. Pero antes de adentrarnos en los detalles de dicho sistema, es necesario establecer diversos conceptos e hitos en la historia de nuestra comprensión sobre la regulación genética.

1.2.4 Regulación genética

El principal motor de la ciencia fue y siempre será la curiosidad humana, pero principalmente la comprensión de la vida misma es motivo de desvelo para muchos de nosotros. Este interrogante infinito parecía tener respuesta al inicio de la era genómica, donde ingenuamente creímos que podríamos comprenderlo si obteníamos el código fuente de la vida, el ADN. Si bien la información genómica nos ha permitido comprender y desarrollar un sinfín de conocimiento y tecnologías, lejos estamos de comprenderlo todo. El estudio de las partes de un sistema es claramente un punto importante y necesario, pero hoy en día sabemos que podemos conocer todas las proteínas, ácidos nucleicos, azúcares, etc. de un organismo, y así y todo no comprender del todo que es lo que está ocurriendo. Es como tener el código de un programa de computadora completo, podemos ver sus rutinas, sus métodos, pero hasta que no coordinemos como van a interactuar unos con los otros, el programa ni siquiera va a iniciar. Estamos en la presencia de un sistema complejo que, según su definición, es aquel sistema que posee más información que la que aporta cada parte independiente. Esta información se esconde en las interacciones, entre todas las moléculas pertenecientes a un organismo, a diferentes e incluso a factores abióticos. Como no podemos abordar el estudio de semejante sistema en su totalidad, nos resignamos a segmentar dicha red, intentando desenmarañarla de a pequeñas porciones, por ejemplo, limitándonos al estudio de la regulación de un gen. El estudio de los circuitos génicos en la respuesta celular está determinado por las interacciones entre los genes que conforman un genoma, y son la clave para comprender el fenotipo que presenta un organismo como respuesta al entorno al que está sometido.

1.2.5 Distintos niveles de regulación

La regulación de la expresión genética no solo define el fenotipo celular, sino que también decide como será expresado dicho fenotipo. Entre la secuencia de un gen y una proteína funcional existen varios puntos de control, donde es posible ajustar el flujo de la expresión génica. En un principio, la mayor parte de las investigaciones se centraron en el estudio del inicio de la transcripción y los factores proteicos que la modulan, sin embargo, hace ya varias décadas

que, debido a la enorme obtención de datos genéticos, bioquímicos, fisiológicos y ómicos sabemos la importancia que tiene la regulación por encima del nivel transcripcional. Aunque en la mayoría de los casos el efecto es leve y sirve para afinar el resultado de la regulación transcripcional, existen situaciones donde las regulaciones postranscripcionales pueden introducir un cambio significativo y convertirse en el control maestro del flujo genético de una determinada vía o proceso.

Los ácidos ribonucleicos cumplen un papel central en las redes postranscripcionales, particularmente cuando se trata de una respuesta rápida y concertada en condiciones de estrés. El término riborregulación refiere a los mecanismos reguladores de la expresión génica que se basan en el funcionamiento de moléculas de ARN (Papenfert, Vogel, and microbe 2010). El control postranscripcional de la expresión génica puede ser determinado por porciones de secuencia del propio ARNm (motivos que actúan en *cis* como *riboswitch* y termo-sensores), por apareamiento de bases entre pequeños ARN reguladores (sARN) no codificantes y ARNm, o por proteínas de unión a ARN que tienen preferencia por secuencias y/o motivos estructurales con sus ARNm objetivo.

1.2.6 Regulación por ARNs codificados en *cis*. Termo-sensores y *riboswitches*.

La estructura de la región no traducida '5 (UTR'5) de un ARNm puede determinar su tasa de traducción. Esto es particularmente relevante para las estructuras de ARN que involucran el sitio de unión al ribosoma (RBS) que contiene la secuencia *Shine-Dalgarno* (SD) y el codón de inicio (AUG), y para las estructuras de tipo bucle (*loop*) que funcionan como atenuadores transcripcionales.

Para muchas bacterias, la temperatura del huésped es una importante señal externa que desencadena la virulencia o la expresión génica relacionada con el estrés. La detección de temperatura puede ocurrir por casi todos los mecanismos de regulación de genes procarióticos, incluido el reconocimiento promotores alterados por cambios en la topología del ADN o la modulación de la actividad del factor de transcripción a través de cambios en la conformación de proteínas (Klinkert and Narberhaus 2009). Es importante destacar que la temperatura también puede influir en el plegamiento de las regiones estructuradas en los ARNm y, por lo tanto, afectar la traducción. Estos llamados "termómetros de ARN" normalmente residen en la región UTR5' de genes sensibles a la temperatura y cambian entre dos estructuras distintas: una conformación "cerrada" formada a baja temperatura, en la que SD y/o AUG están inaccesible a los ribosomas; y la confirmación "abierta" formada a alta temperatura tras la fusión de la estructura inhibitoria alrededor del RBS (Figura 9).

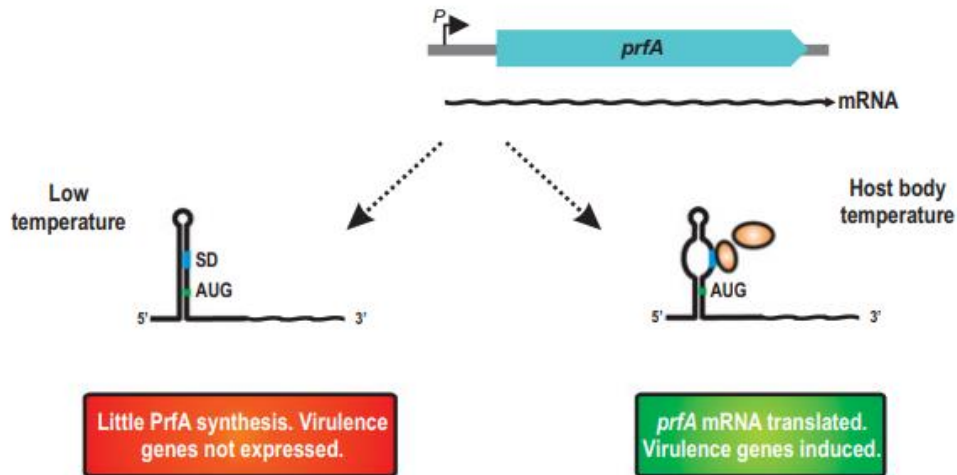


Figura 9. El ARNm del factor de transcripción de virulencia *prfA* de *L. monocytogenes* se controla postranscripcionalmente mediante un termómetro de ARN en su UTR5'. Esta estructura reguladora permite el inicio de la traducción a alta temperatura. Modificado de (Papenfors and Vogel 2010).

Por otro lado, los *riboswitch* detectan señales físicas y químicas, además de la temperatura, y estos incluyen una abundancia de metabolitos químicamente muy diversa (Roth and Breaker 2009), pH e iones metálicos (Dambach and Winkler 2009). Estructuralmente, los *riboswitch* están organizados en dos dominios: la "región del aptámero", que se une al ligando y la "plataforma de expresión" capaz de formar dos estructuras de ARN mutuamente excluyentes dependiendo de si un ligando está unido o no. En la mayoría de los casos, la unión del ligando bloquea la traducción del ARNm, ya sea promoviendo la formación de un terminador transcripcional o mediante el secuestro de RBS como en los termómetros de ARN. En la Figura 10 se esquematiza un ejemplo de un *riboswitch* sensor de di-GMP cíclico.

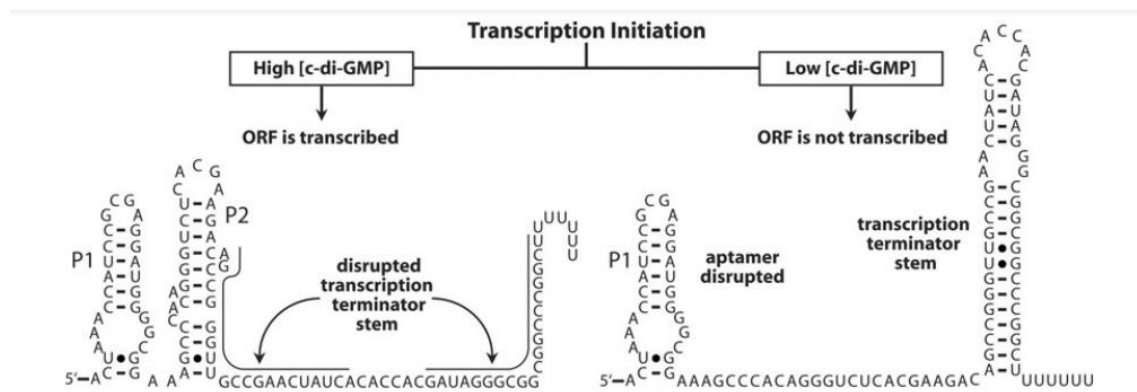


Figura 10. Mecanismo predicho de un *riboswitch* c-di-GMP de *Geobacter sulfurreducens* ubicado en la región UTR'5 de un gen putativo del citocromo C. Cuando las concentraciones de c-di-GMP son altas (izquierda), se predice que la unión del ligando estabilizará la estructura terciaria que involucra a las horquillas (*stem loops*) P1 y P2, lo que debería impedir la formación de un *stem loop* terminador de transcripción intrínseco. Cuando las concentraciones de c-diGMP son bajas (derecha), el aptámero no se unirá al ligando y se formará el *stem loop* terminador en lugar del P2 inestable. Modificado de (Weinberg et al. 2007).

1.2.7 Regulación por ARNs codificados en *trans*

La mayoría de los riboreguladores que modulan la expresión de ARNm mediante mecanismos de apareamiento de bases se transcriben como ARN antisentido codificado en *cis* desde una hebra opuesta a la del ARNm, o como ARNs codificados en *trans* en loci físicamente no vinculados al ARNm que regulan. Los ARNs *trans*-antisentido, generalmente actúan mediante un apareamiento de bases corto e imperfecto y, a menudo, requieren chaperonas de ARN como la proteína Hfq, mientras que los ARN *cis*-antisentido generalmente necesitan una amplia complementariedad de secuencia con el ARN transcrito de manera opuesta (Wagner, Altuvia, and Romby 2002). Existe otro mecanismo de regulación mediado por ARNs codificados en *trans*, que no involucra complementariedad de bases, sino que actúan por interacción con proteínas (Csr/Rsm) regulatorias por un mecanismo comúnmente llamado mimetismo molecular.

1.2.7.1 Mecanismos dependientes de Hfq

La proteína de unión a ARN Hfq es un regulador global necesario para la función de pequeños ARN reguladores que actúan por apareamiento de bases limitado. Hfq estabiliza los ARNs y promueve sus interacciones con los ARNm, lo que conduce a una alteración de la estabilidad y/o traducción de los transcritos (Vogel and Luisi 2011). Hfq se descubrió originalmente como un factor celular necesario para la replicación eficiente del bacteriófago Q β en *E. coli* (Franze de Fernandez, Eoyang, and August 1968). Hfq es un miembro de la gran familia Sm/Lsm de proteínas de unión a ARN que se encuentran en eucariotas y arqueas, y los homólogos de Hfq están presentes en aproximadamente el 50 % de las especies bacterianas (Sun, Zhulin, and Wartell 2002). Hfq se ha relacionado con una gran variedad de procesos celulares, sin embargo, es mejor conocido por su función como emparejador en la regulación postranscripcional al promover la interacción entre los pares de sARN y mARN (Figura 11) (Møller et al. 2002).

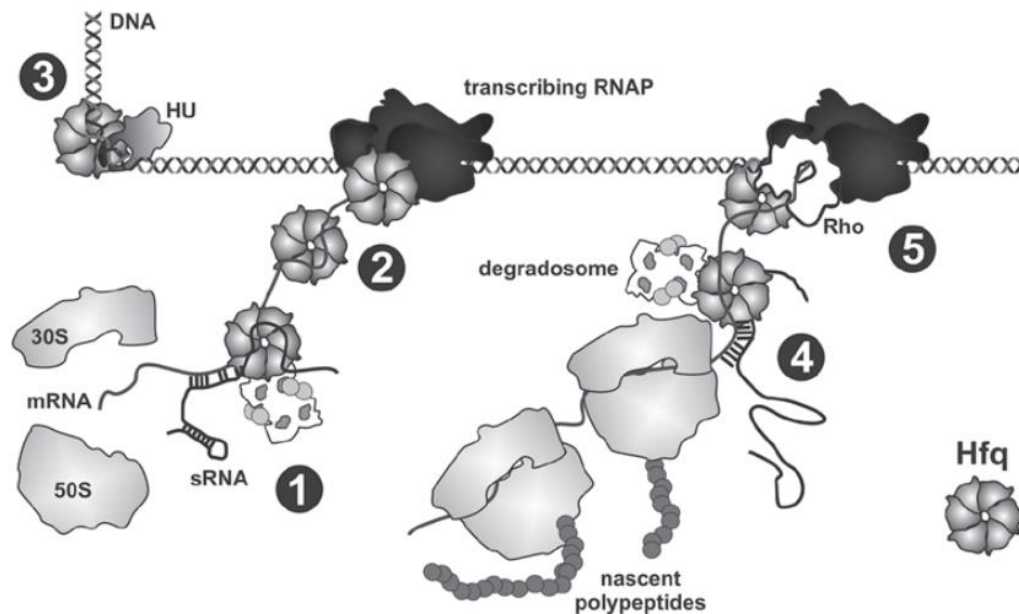


Figura 11. Rol de la proteína reguladora global Hfq en el control de la expresión génica por diferentes mecanismos a diferentes niveles del flujo de información genética. (1) Hfq promueve el emparejamiento antisentido entre un sRNA y un ARNm, generalmente cerca o en el RBS que impide el acceso al ribosoma. Hfq recluta el degradosoma, facilitando la degradación acoplada de ARNm y sRNA y, por lo tanto, silenciando la expresión del ARNm. (2) Como chaperona de ARN, Hfq puede afectar localmente el plegamiento del ARNm y atenuar la pausa transcripcional. El reconocimiento del sustrato de ARN por Hfq puede ocurrir directamente en el sitio de transcripción ya que la holoenzima de la polimerasa de ARN se copurifica con Hfq. (3) Hfq puede unirse al ADN y también copurificar las proteínas tipo histonas. Por lo tanto, podría ser posible que tenga un papel en el plegamiento del ADN y el control transcripcional. (4) El control traduccional mediado por Hfq y sRNA dentro de la región codificante de un mRNA puede desacoplar la traducción de la transcripción y facilitar la terminación dependiente de Rho. (5) En ausencia de control de traducción, Hfq puede controlar la antiterminación a través de su interacción con Rho (Sobrero and Valverde 2012).

1.2.7.2 Reguladores globales de la familia CsrA

CsrA significa regulador de almacenamiento de carbono A (*Carbon Storage Regulator A*) y se descubrió hace 30 años en un *screening* de mutantes Tn5 de *E. coli* como un factor regulador de los genes de biosíntesis de glucógeno, y rápidamente se reveló su función como regulador global de la expresión génica (Romeo et al. 1993). Los miembros de la familia CsrA son polipéptidos bastante bien conservados de longitud relativamente corta (65–75 aminoácidos) que funcionan como homodímeros. Aún no se ha demostrado la heterodimerización en bacterias que codifican más de un parólogo, pero su existencia no se ha descartado (Sobrero and Valverde 2020). CsrA es capaz de unirse a motivos específicos de diversos ARNs, caracterizados por pequeños *stem loops* que poseen el trinucleótido GGA en su extremo. Un residuo de arginina fuertemente conservado es fundamental para el reconocimiento de la primera G del trinucleótido, y su reemplazo anula la unión reguladora de las proteínas CsrA a sus ARNs objetivos (Schubert et al.

2007). Según la región del ARNm a la que se une CsrA, el resultado de la interacción puede ser: (a) la represión de la traducción del ARNm al bloquear a la unidad ribosómica pequeña; (b) activación de la traducción tras el reordenamiento estructural de la región UTR5' y la exposición del RBS; (c) regulación de la traducción mediante el replegamiento de la región UTR5' de modo que un sARN pueda obtener acceso por emparejamiento de bases y evitar la entrada al ribosoma; (d) modulación de la degradación del ARNm favoreciendo el acceso de las ribonucleasas a sus sitios blanco; (e) modulación de la terminación de la transcripción (Figura 12).

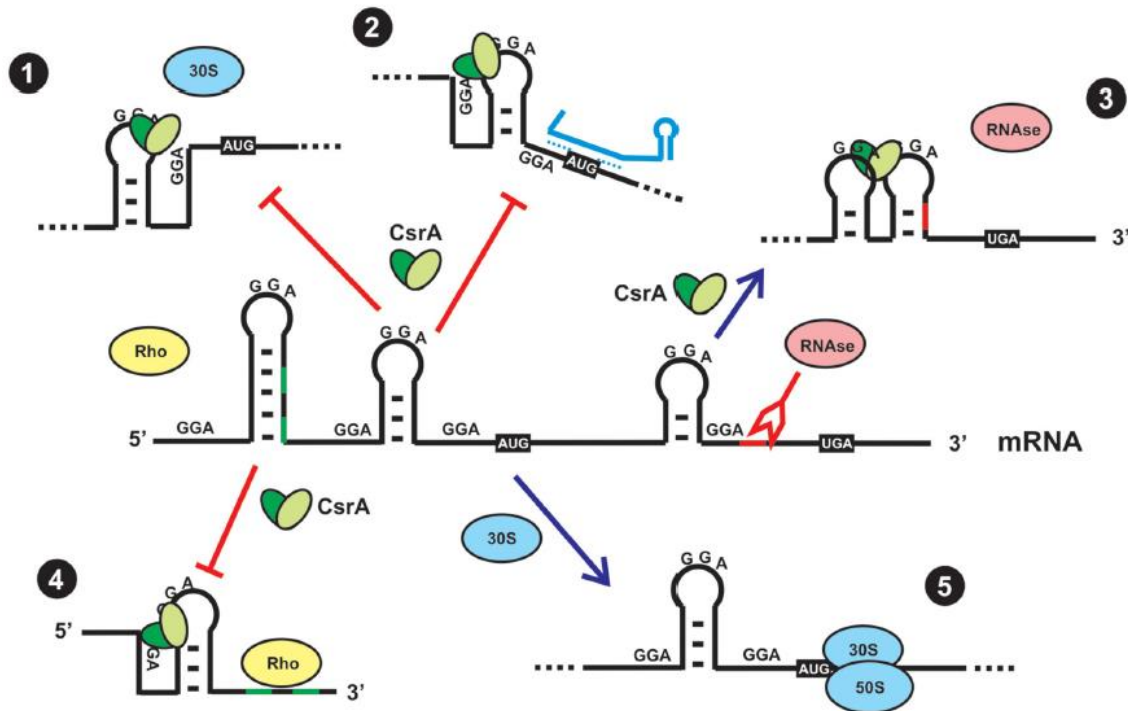


Figura 12. Efectos reguladores directos e indirectos de la unión de CsrA a un ARNm. 1) Represión traduccional directa. En la mayoría de los casos, CsrA se une alrededor o dentro de la región de SD y, bloquea la unión de la subunidad ribosómica pequeña y su participación en el inicio de la traducción. (2) Represión traduccional indirecta. CsrA se une a sitios específicos dentro del UTR5' e induce un cambio conformacional que expone la región SD al apareamiento de base con un sARN regulador, lo que impide el acceso al ribosoma. (3) Control indirecto de la estabilidad del ARNm. CsrA se une a un sitio objetivo que, al replegarse, impide el acceso de una ribonucleasa. (4) Control indirecto de la terminación transcripcional. CsrA se une a un sitio específico en el ARNm e induce un cambio conformacional que expone las secuencias de reconocimiento de Rho que ahora pueden ser utilizadas por la proteína Rho para detener la elongación del ARNm. (5) En ausencia de CsrA (ya sea debido a una mutación que suprime su función o debido al secuestro por antagonistas de CsrA, como se muestra en la Figura 13), un ribosoma puede reconocer la secuencia de SD e iniciar la traducción del ARNm. (Sobrero and Valverde 2020)

Teniendo en cuenta que la abundancia intracelular de la proteína CsrA es relativamente alta, (Ishihama et al. 2008), es lógico pensar que numerosos ARNm se encuentren bloqueados por CsrA, esperando ser liberados ante una señal regulatoria que orqueste una activación coordinada. Hasta la fecha, se han descrito dos mecanismos mutuamente excluyentes para

antagonizar la actividad de las proteínas CsrA: 1) mimetismo molecular por sARN (Figura 13A); 2) interacción alostérica con proteínas, como FliW (Figura 13B). Este último se ha descrito en especies para las que aún no se han encontrado antagonistas de sARN, como *Bacillus subtilis* y *Campylobacter jejuni* (Altegoer, Rensing, and Bange 2016; Dugar et al. 2016). Sin embargo, la estrategia de antagonismo más generalizada para aliviar el control postranscripcional de las proteínas de la familia CsrA es a través de la expresión de ARNs imitadores moleculares (Babitzke et al. 2019). Esta clase de moléculas de ARN reguladoras que no codifican proteínas se comportan como “esponjas” de proteínas que forman complejos ribonucleoproteicos que alivian temporalmente el efecto regulador que las proteínas CsrA sobre sus objetivos de ARNm (Figura 13A). Para lograr esto, los sARN miméticos moleculares ofrecen múltiples motivos estructurales y de secuencia formados por horquillas cortas que exponen pentanucleótidos ANGGA desapareados, es decir, el objetivo molecular preferido de las proteínas CsrA (Figura 14).

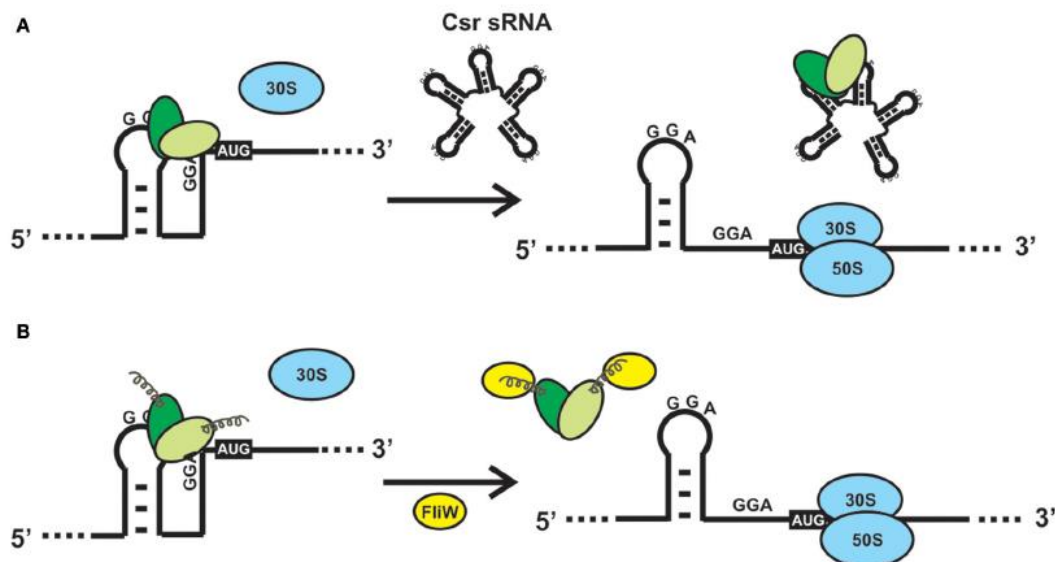


Figura 13. Mecanismos alternativos para aliviar los efectos directos o indirectos de la unión de CsrA a un ARNm. (A) Secuestro de CsrA por sARNs de imitación molecular. Los sARNs de la familia Csr/Rsm actúan como esponjas proteicas al ofrecer múltiples secuencias/motivos estructurales de unión a CsrA. La disponibilidad de dímeros de CsrA para unirse a múltiples ARNm está modulada por la concentración intracelular de uno o más de estos sARN miméticos moleculares. (B) CsrA se puede desplazar de sus sustratos de ARNm tras la interacción de una proteína como FliW de *B. subtilis* con la región C-terminal extendida (Altegoer, Rensing, and Bange 2016). Tal interacción provoca un cambio conformacional en el dímero de la proteína que reduce la afinidad por el ARN. (Sobrero and Valverde 2020)

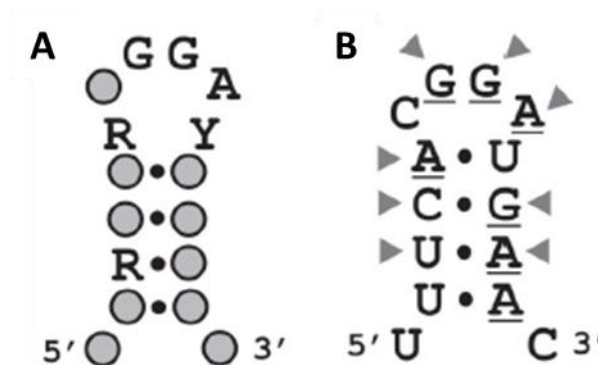


Figura 14. (A) Secuencia consenso/motivo estructural de unión de por CsrA *in vivo*, como se detecta. R = G o A, Y = C o U. (B) Estructura secundaria predicha de la región de 14 nucleótidos que abarca la secuencia SD (nucleótidos subrayados) del UTR5' del ARNm de *hcnA* de *P. protegens* CHA0, un *target* validado experimentalmente de la proteína homóloga de CsrA, RsmE (Lapouge et al. 2007; Schubert et al. 2007).

E. coli sintetiza dos sARN miméticos denominados CsrB (con 18 motivos de unión a CsrA) y CsrC (con 11 motivos de unión a CsrA) que son capaces de secuestrar CsrA, y así invertir el sentido de su regulación. La expresión de los ARN se encuentra coordinada por el sistema de dos componentes BarA-UvrY, que responde a la presencia de acetato o formiato y también a cambios de pH (Mondragón et al. 2006). El resultado de la activación en cascada es la regulación de la biosíntesis de polímeros de carbono, la motilidad y la formación de biopelículas (Babitzke et al. 2019). Esta cascada regulatoria se encuentra ampliamente distribuida en bacterias y es de particular importancia para este trabajo el estudio de sus homólogos en *Pseudomonas*.

1.2.8 Cascada regulatoria Gac-Rsm en *Pseudomonas protegens* CHA0

Como se mencionó previamente, las pseudomonas poseen un sistema de regulación postranscripcional homólogo al de Csr en enterobacterias. Existen numerosos parálogos de CsrA denominados Rsm y es posible encontrar entre uno y cinco homólogos dentro de la misma bacteria, siendo RsmA el más distribuido. Lo mismo ocurre con los ARNs miméticos denominados rsm. Particularmente en CHA0 podemos encontrar dos proteínas tipo CsrA, denominadas RsmA y RsmE; y tres ARNs pequeños denominados *rsmX*, *rsmY* y *rsmZ*.

Tras la activación, el sistema de dos componentes GacS/GacA activa la transcripción de los sARN denominados *csrB*, *csrC* (en bacterias entéricas), y *rsmX*, *rsmY* y *rsmZ* (en pseudomonas y bacterias relacionadas) (Figura 15). La expresión de los genes sARN y, por lo tanto, la de muchos genes *target* aumenta fuertemente con el aumento de la densidad de población celular en *P. protegens* (Valverde et al. 2003). Se sospecha que el regulador de respuesta GacA activado (fosforilado) se une a un elemento conservado río arriba denominado caja GacA (consenso TGTAAGN₆CTTACA) en los promotores de los genes sARN mencionados anteriormente. Al producir los sARN, la vía de transducción de señales Gac/Rsm aumenta la producción de numerosas proteínas cuya producción es reprimida por las proteínas RsmA/CsrA (Figura 15). Existen factores adicionales que influyen en la cascada Gac-Rsm, se han reportado dos quinasas

sensores que modulan la actividad de GacS de forma directa: LadS, que estimula la actividad de GacS, y RetS, que es un regulador negativo (Figura 15) (Ventre et al. 2006; Chambonnier et al. 2016). Mientras que los componentes maestros de la cascada GacS-GacA están ampliamente distribuidos dentro de γ -proteobacterias, *ladS* y *retS* muestran una ocurrencia más esporádica. Ambos genes que codifican LadS y RetS están presentes junto con los homólogos de GacS y GacA solo en las especies de *Pseudomonas*, *Lysobacter* y *Alcanivorax* (Sobrero and Valverde 2020).

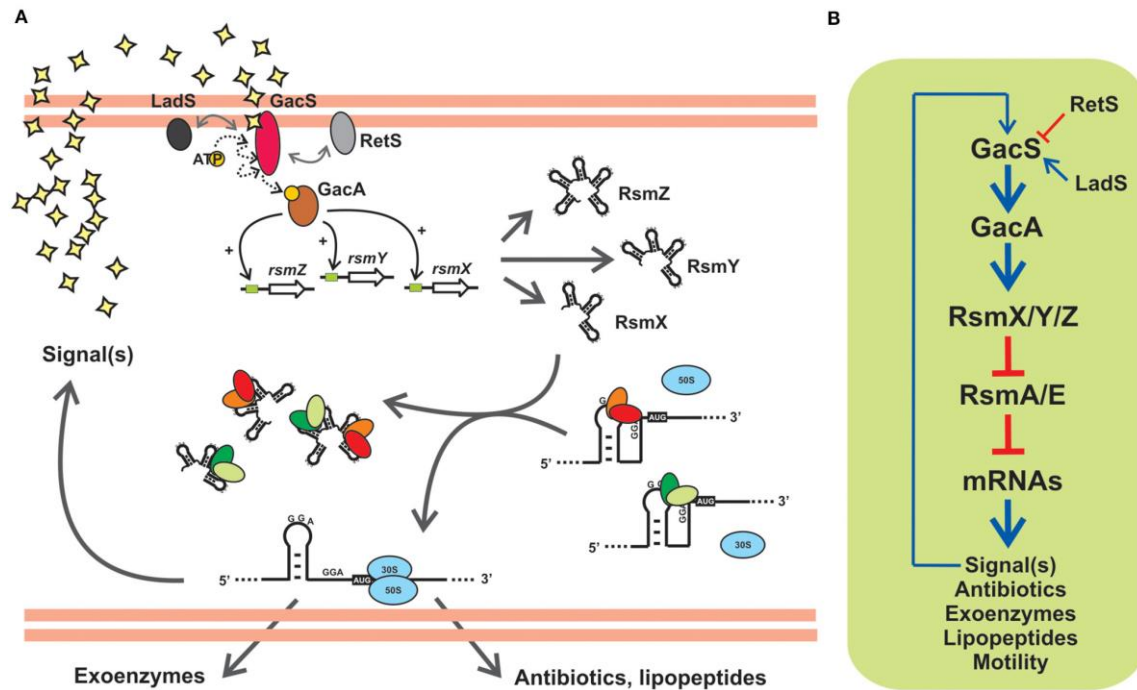


Figura 15. Esquema de la cascada postranscripcional de Gac-Rsm de la cepa *P. protegens* CHA0, la cual representa una vía arquetípica de transducción de señales que controla la actividad de las proteínas CsrA con ARNs imitadores moleculares en γ -proteobacteria. (A) En la cepa CHA0, el sistema de dos componentes GacS-GacA responde a señales de autoinducción no caracterizadas hasta el momento, y promueve la traducción de varios ARNm involucrados en la formación de exoproductos y el control de patógenos de raíces de plantas (antibióticos, lipopéptidos, enzimas líticas extracelulares), al aliviar el bloqueo de traducción causado por los dos homólogos de CsrA, RsmA y RsmE. Esto se logra activando la transcripción de los genes *rsmX*, *rsmY* y *rsmZ*, aumentando así la abundancia intracelular de los tres ARN imitadores RsmX, RsmY y RsmZ, que secuestran RsmA y RsmE. Dos quinasas sensoras huérfanas accesorias, RetS y LadS, modulan la actividad de GacS. (B) Representación lineal esquemática de la vía reguladora Gac-Rsm de la cepa *P. protegens* CHA0. La biosíntesis de un compuesto de señal difusible también está bajo el control de la cascada, como una serie de otros fenotipos que contribuyen a la capacidad antifúngica de esta cepa, y representa una entrada de retroalimentación positiva para la vía. (Sobrero and Valverde 2020).

1.2.9 Experimentos previos, screening bioinformático y fusiones reporteras

Previamente en nuestro laboratorio, se llevó a cabo una búsqueda bioinformática con el objetivo de identificar posibles nuevos objetivos (*targets*) regulatorios directos de la cascada Gac-Rsm. Para eso se seleccionaron regiones de 300pb circundantes al codón de inicio (250 pb río arriba del codón de inicio y 50pb río abajo), de todos los marcos abiertos de lectura (ORF) anotados en

el genoma de CHA0. Sobre esas regiones se buscaron los motivos de unión a las proteínas RsmA/E predichos por Schubert et al. (2007). Como resultado de este análisis se obtuvieron 116 posibles targets, dentro de los cuales se encontraron algunos ya validados experimentalmente como *hcnA* y *aprA* (Blumer et al. 1999). La lista principal se redujo tras aplicar los siguientes tres criterios de filtro: (1) el sitio de unión RsmA/E putativo debe estar ubicado dentro o cerca de la secuencia SD; (2) la secuencia del sitio debe plegarse en una estructura de *stem loop* que exponga el trinucleótido GGA en el *loop*; (3) el motivo de secuencia debe conservarse evolutivamente en las regiones UTR5' de los genes ortólogos en pseudomonas relacionadas filogenéticamente. Tras aplicar estos filtros, se obtuvo una lista de 45 posibles *targets* de ARNm para las proteínas RsmA/E. A raíz de este trabajo se lograron identificar dos genes que flanquean un *cluster* que posiblemente sintetice un lipopéptido y que poseen un posible sitio unión para RsmA/E. Además, según la información obtenida en base de datos por alineamiento de secuencias, ambas proteínas presentaron homología con la familia de reguladores transcripcionales del tipo *luxR*. En base a su posición relativa del *cluster* se los llamo OrfUR (Upstream Regulator) y OrfDR (Downstream Regulator). No se encontró ningún posible sitio de regulación por RsmA/E en los genes involucrados en la biosíntesis, por lo que se descartó una regulación directa y se formuló la primera hipótesis: los genes *orfUR* y *orfDR* están involucrados en la activación transcripcional de los genes biosintéticos, siendo así el vínculo entre la cascada Gac-Rsm y la biosíntesis de LPs. Para dar sustento a dicha afirmación, lo primero que se realizó fue un análisis de sintenia entre los *clusters* de lipopéptidos de otras pseudomonas, bajo la premisa de que, si están involucrados en la regulación, ambos genes deberían estar conservados junto con el resto de los genes biosintéticos. Dentro de la Figura 16 se pueden observar diferentes *clusters* dedicados a la biosíntesis de LPs en pseudomonas, en los que todos los genes se encuentran dentro de un mismo locus cromosomal. En todos los casos existe la conservación de dos reguladores transcripcionales del tipo *luxR* flanqueando cada *cluster*. Para mayor detalle sobre este trabajo referir a Sobrero et al. (2017).

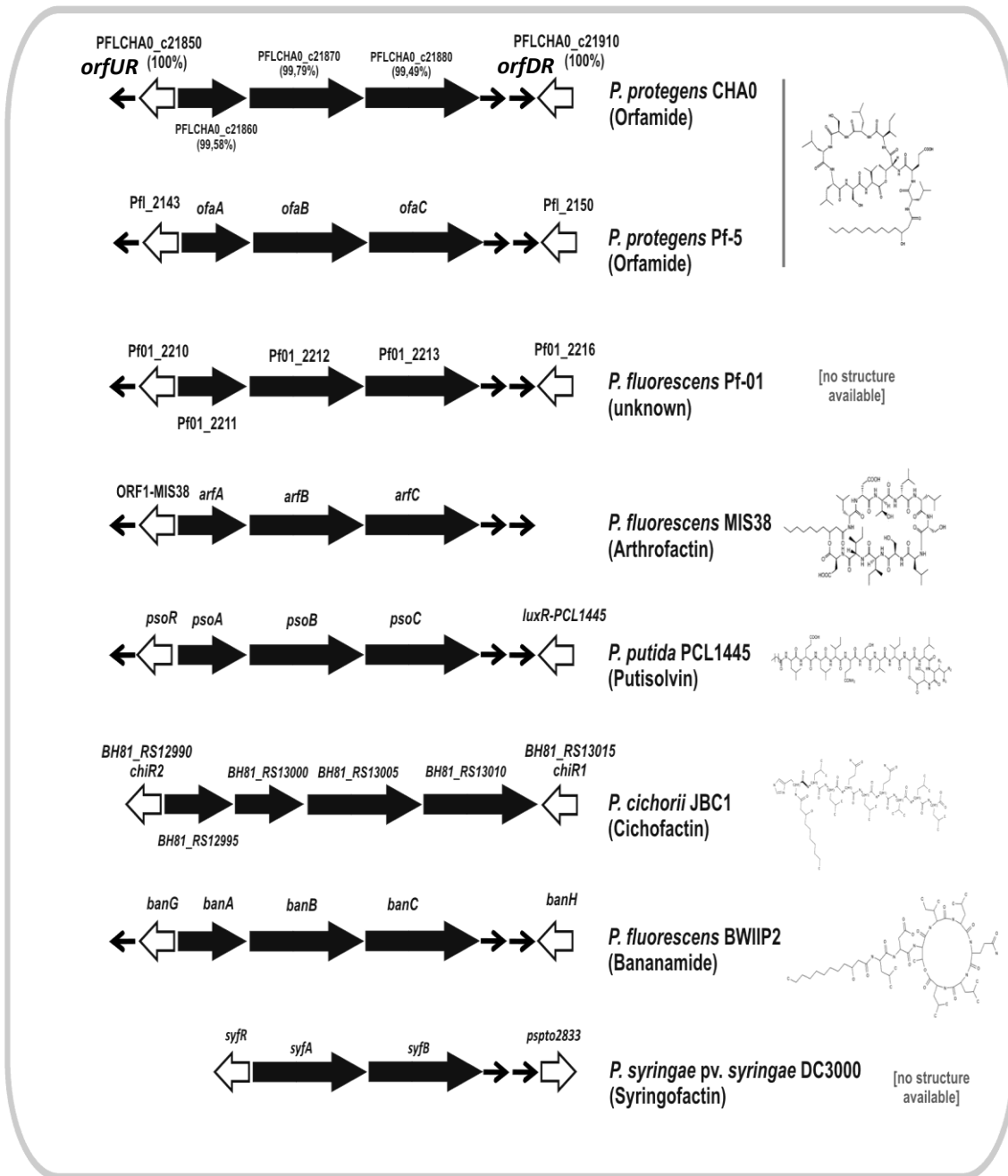


Figura 16. Sintenia del locus biosintético cLPs en pseudomonas relacionadas. Los genes implicados en la biosíntesis de los cLPs correspondientes se indican con flechas negras sólidas, mientras que los loci que codifican los reguladores transcripcionales predichos o verificados experimentalmente de los genes biosintéticos se indican con flechas blancas abiertas. Esta figura representa los grupos de genes cLPs codificados dentro de un solo locus. Modificado de Sobrero et al. (2017)

Para determinar el patrón de expresión de los reguladores *orfUR* y *orfDR* nuestro laboratorio contaba con dos vectores que poseen una fusión traduccional de la región líder de cada gen y el ORF de *lacZ* desprovisto de sus primeros 7 codones (*'lacZ*), bajo el control de los promotores naturales de cada gen *orfUR* u *orfDR* (pME7015 y pME7016, respectivamente). Ambos genes presentaron un patrón de expresión típico de regulación por Gac/Rsm Figura 17, es decir, baja

expresión en ausencia de sensor GacS funcional o de los ARNs miméticos RsmX/Y/Z; y por el contrario una sobre expresión en mutantes Δ RsmA/E.

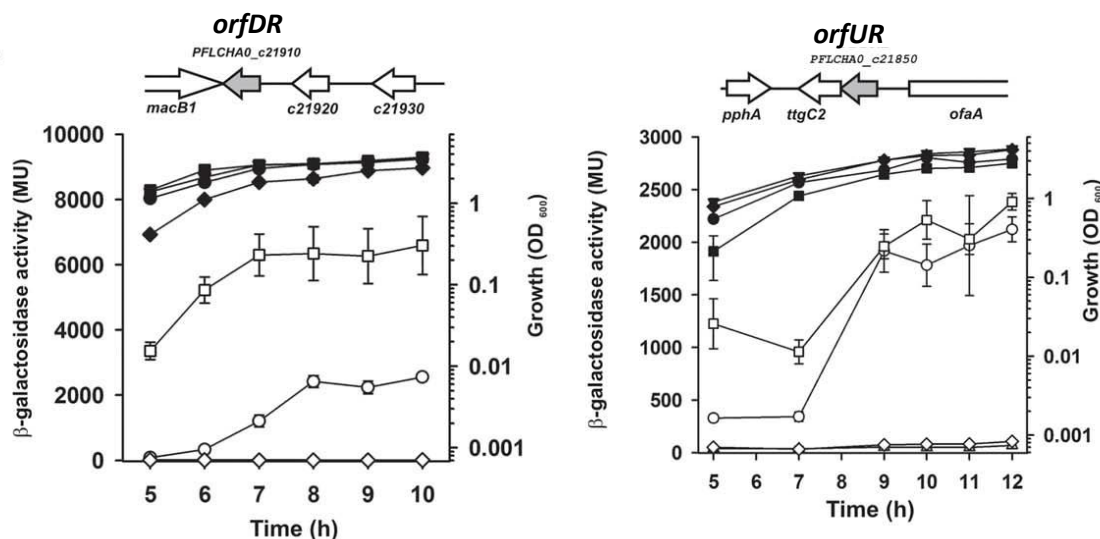


Figura 17. Control de la expresión de los dos genes que codifican posibles reguladores transcripcionales de la familia *luxR*, *orfDR* y *orfUR*, por el sistema Gac-Rsm de *P. protegens* CHA0. Impacto de las mutaciones *gacS*, *rsmA/rsmE* y *rsmX/Y/Z* en la expresión de un plásmido *orfDR*'-'*lacZ* (izquierda) u *orfUR*'-'*lacZ* (derecha). Modificado de Sobrero et al. (2017)

En base a toda la evidencia mencionada, se decidió estudiar el control de la expresión de los posibles reguladores transcripcionales OrfUR y OrfDR por parte de la cascada Gac-Rsm, bajo la hipótesis que ambos genes son activadores transcripcionales de los genes biosintéticos del *cluster* dedicado a la síntesis de LPs en CHA0.

1.3 Hipótesis de trabajo

En *Pseudomonas protegens* CHA0, la cascada regulatoria Gac-Rsm controla a nivel postranscripcional la expresión de numerosos genes que le confieren un fenotipo biocontrolador, por medio de la expresión de diversos metabolitos especializados. Particularmente los lipopéptidos, responden a una regulación por Gac-Rsm pero los genes encargados de su biosíntesis no poseen motivos los secuenciales necesarios para ser blancos regulatorios. Por lo tanto, existirán genes regulados por Gac-Rsm involucrados en la activación de los genes biosintéticos, siendo así el vínculo entre la cascada Gac-Rsm y la biosíntesis de lipopéptidos.

1.4 Objetivos específicos

- Determinar el patrón de expresión de los genes *orfUR* y *orfDR*
- Confirmar la regulación directa de los genes *orfUR* y *orfDR* por la cascada Gac-Rsm
- Estudiar el patrón de expresión de los lipopéptidos de CHA0 y determinar su naturaleza química

1.5 Materiales y métodos

1.5.1 Cepas bacterianas, plásmidos y oligonucleótidos

Las cepas bacterianas, plásmidos y oligonucleótidos empleados en este capítulo se listan en la Tabla 1. *E. Coli* se creció a 37°C y 200 rpm en caldo NYB (caldo nutritivo, 25 g/L; extracto de levadura, 5 g/L), todas las pseudomonas se crecieron en NYB a 28°C y 200 rpm de forma rutinaria excepto que se detalle lo contrario. Las cepas se conservaron, en condiciones de esterilidad, en una solución acuosa de glicerol 20% final en ultrafreezer (-80°C). En los casos requeridos, los antibióticos fueron suministrados en las siguientes concentraciones: para *E. coli*, ampicilina (Ap) 100 µg/ml; kanamicina (Km) 50 µg/ml; gentamicina (Gm) 10 µg/ml; cloranfenicol 10 µg/ml; tetraciclina (Tc) 50 µg/ml. Para pseudomonas, kanamicina (Km) 50 µg/ml; gentamicina (Gm) 10 µg/ml; tetraciclina (Tc) 125 µg/ml.

Los cultivos de pseudomonas destinados la obtención de lipopéptidos fueron crecidos en medio PSFM (triptona 10 g/L, peptona 10 g/L, K₂HPO₄ 1,5 g/L, MgSO₄·7H₂O 1,5 g/L, glicerol 19 g/L). Los medios utilizados para evaluar la producción óptima de lipopéptidos fueron: King's B (peptona 20 g/L, K₂HPO₄ 1,5 g/L, MgSO₄·7H₂O 1,5 g/L, glicerol 7 g/L), DBM glucosa ((NH₄)₂SO₄ 1 g/L, Na₂HPO₄ 8.83 g/L, MgSO₄ 0.1 g/L, KH₂PO₄ 2 g/L, NaH₂PO₄ 1.15 g/L, glucosa 0.5 g/L), DBM manitol ((NH₄)₂SO₄ 1 g/L, Na₂HPO₄ 8.83 g/L, MgSO₄ 0.1 g/L, KH₂PO₄ 2 g/L, NaH₂PO₄ 1.15 g/L, manitol 0.5 g/L).

Los cultivos de pseudomonas crecidos en placas multiwell destinados a cinéticas de crecimiento con medidas de fluorescencia fueron crecidos en GCM (Casaminoácidos 5 g/L, K₂HPO₄ 0,75 g/L, MgSO₄·7H₂O 0,75 g/L, glicerol 20 g/L).

Tabla 1. Cepas bacterianas, plásmidos y oligonucleótidos utilizados en este capítulo

Microorganismo/ oligonucleótido	Genotipo característica o secuencia				Referencia
<i>Pseudomonas protegens</i>					
CHA0	Cepa Silvestre				(Stutz, Défago, and Kern 1986)
CHA19	$\Delta gacS$				(Zuber et al. 2003)
CHA89	$gacA::\Omega Km^R$				(Blumer et al. 1999)
CHA1009	$rsmA::\Omega Km$ $rsmE::\Omega Hg$; Km^R				(Reimmann et al. 2005)
CHA1144	$\Delta rsmX \Delta rsmY \Delta rsmZ$				(Kay, Dubuis, and Haas 2005)
<i>Oomicete</i>					
<i>Pythium ultimum</i> 67-1	Oomicete fitopatógeno				(Maurhofer et al. 1994b)
<i>Escherichia coli</i>					
DH5α	$recA1$	$endA1$	$hsdR17$	$deoR$ $thi-1$ $supE44$	Stock de laboratorio
	$gyrA96$ $relA1\Delta(lacZYA-argF)$ $U169$ ($\Phi 80d lacZ\Delta M15$)				
HB101	Cepa <i>helper</i> ; F- λ - $hsdS20(rB^- mB^-)$ $recA13$ $leuB6(Am)$ $araC14 \Delta(gpt-proA)62$ $lacY1$ $galK2(Oc)$ $xyl-5$ $mtl-1$ $thiE1$ $rpsL20(Sm^R)$ $glnX44(AS)$				(Boyer and Roulland-Dussoix 1969)

Plásmidos

pBluescript SK(-)	Vector de clonado, ColE1-replicón; Ap ^R	Stratagen
pRK600	Vector <i>helper</i> para conjugación; ori(ColE1), RK2(mob+ tra+); Cm ^R	(Kessler et al. 1992)
pME6032	<i>lacI</i> ^q -P _{trac} Vector de expresión; Tc ^R	(Heeb, Blumer, and Haas 2002b)
pME6078	<i>rsmA</i> ^{CHA0} -His ₆ en pME6032	(Heeb, Blumer, and Haas 2002b)
pME7010	<i>rsmE</i> ^{CHA0} -His ₆ en pME6032	(Sobrero et al. 2017)
pME6000	Vector de clonado, pBBR1MCS derivative; Tc ^R	(Maurhofer et al. 1998)
pME6102	<i>orfUR</i> '-' <i>egfp</i> ; <i>orfDR</i> '-' <i>DsRed</i> . Tc ^R , Km ^R	Este trabajo
pME7018	<i>orfUR</i> ^{CHA0} '-' <i>lacZ</i> . Tc ^R	(Sobrero et al. 2017)
pME7019	<i>orfDR</i> ^{CHA0} '-' <i>lacZ</i> . Tc ^R	(Sobrero et al. 2017)

Oligonucleótidos**de ARN**

orfUR silvestre	CACAUUCAUGGAUGAAAUGA	(Sobrero et al. 2017)
orfDR silvestre	CUCGAUCAAGGAUGAUAAAGG	(Sobrero et al. 2017)
orfUR^{AUGGA→AUGAA}	CACAUUCAUGGAUGAAAUGA	(Sobrero et al. 2017)

1.5.2 Colapso de gota

Para detectar la presencia de compuestos surfactantes como lipopéptidos en sobrenadantes de cultivo, a través de su capacidad para reducir la tensión superficial, se depositó una gota de sobrenadante sobre una superficie de Parafilm®, siguiendo la metodología propuesta por De Bruijn, Raaijmakers, and microbiology (2009). Para mejorar la visualización se agregó colorante Coomassie blue a las gotas.

1.5.3 Enriquecimiento de lipopéptidos a partir de sobrenadantes de cultivo

Para generar fracciones de sobrenadantes de cultivo enriquecidas en lipopéptidos se utilizó la metodología planteada por Alvarez et al. (2012). Se realizaron cultivos de 250 ml en PSFM para cada cepa a 28°C y 200rpm, y se colectaron las células luego de 48hs. Se centrifugó a 8.000 rpm por 15min y se guardó el sobrenadante. Se agregó HCl 37% hasta pH = 2 y se dejó en reposo a 4°C por 1 hora. Luego se centrifugó a 10.000 rpm y se conservó el pellet. Se agregaron 2 ml de metanol, se resuspendió el pellet y se filtró a través de membranas estériles de 0.2 µm. Las muestras fueron conservadas a -20°C hasta su uso.

1.5.4 Análisis y purificación de lipopéptidos por HPLC

Se inyectaron los precipitados de lipopéptidos en un HPLC (Infinity LC Grad, Agilent) utilizando una columna Analytical Zorbax C18 (4.6 mm × 150 mm, Agilent) y un detector ultravioleta (Agilent). La temperatura se mantuvo a 28°C. La separación se llevó a cabo utilizando una mezcla de solventes conformada por acetonitrilo y agua acidificada con 0.1% ácido fórmico a un flujo de 1 µl/min. Durante la elución el porcentaje de acetonitrilo fue aumentando de la siguiente manera: 10% por 0–10 min, 35–80% por 10–45 min, 80% por 45–55 min y 80–100% por 55–65 min. La identidad de los lipopéptidos en las fracciones eluidas de interés (orfamidas) fue

confirmada por espectrometría de masas MALDI-TOF en el CEQUIBIEM (Centro de Estudios Químicos y Biológicos por Espectrometría de Masa) de FCEN-UBA.

1.5.5 Ensayo de actividad β -galactosidasa

Se realizó el ensayo de actividad para β -galactosidasa propuesto por Miller (1972). Para cada medida se tomó 1 ml de cultivo, se midió la absorbancia a 600nm, se centrifugó a máxima velocidad por 1 min y el pellet se resuspendió en solución salina estéril (NaCl 0.9%). Se incubó la suspensión por 20 min con 45 μ l de tolueno a 37°C. Para el ensayo colorimétrico se prepararon reacciones conteniendo 200 μ l de suspensión bacteriana y 800 μ l de buffer Z⁺ (Na₂HPO₄ 8.5 g/L, NaH₂PO₄·H₂O 5.5 g/L, KCl 0.75 g/L, MgSO₄·7H₂O 125 mg/L, 0.1% 2-mercaptoethanol V/V). Luego de termostatar a 37 °C por 5 min, se inició la reacción con el agregado de 200 μ l de ONPG (o-nitrofenil- β -galactósido) 4 mg/ml en agua destilada y se detuvo por el agregado de 500 μ l de una solución 1M de Na₂CO₃. Se registró la absorbancia a 420 nm y se calcularon las unidades Miller a partir de la siguiente formula: $Miller\ units\ (MU) = 1000 * Abs_{420} / t * V * DO_{600}$, donde t = tiempo de reacción, V= volumen de la muestra.

1.5.6 Transformación de *E. coli* por electroporación

Las células de *E. coli* fueron transformadas según procedimientos descriptos (Crick 1970). Para preparar las células electrocompetentes se inoculó un Erlenmeyer de 250 ml conteniendo 100 ml de medio NYB. Una vez que el cultivo alcanzó una DO₆₀₀ de 0.7, las células se colectaron por centrifugación a 3000 rpm por 15 min a 4°C y se lavaron al menos tres veces con una solución fría de glicerol 10% (v/v) en H₂O milliQ. Para la electroporación, 50 μ l de células electrocompetentes se mezclaron con 0.05-0.5 μ g de ADN en una cubeta 0.1 cm y se electroporó en un Bio-RAD *gene pulser* con un pulso de voltaje de 2.5 kV, 25 μ F y 200 Ω . Luego de la transformación, la mezcla se diluyó en 1 ml de medio NYB y se incubó durante 1 h a 37°C y 200 rpm. Por último, la mezcla se plaqueo en medio NA solido con los antibióticos correspondientes.

1.5.7 Construcción de fusiones reporteras

Ambas construcciones se diseñaron *in-silico*, y fueron sintetizadas por Genscript® en el vector pUC57. Cada construcción posee su propio promotor y terminador, la región líder del ARNm y los primeros 7 codones del ORF, fusionados a *egfp* o *DsRed*. Además, para facilitar sub-clonados, se agregaron sitios de restricción para la enzima *HindIII* flanqueando la construcción *orfUR*'-*egfp*, y los sitios *SpeI* flanqueando la construcción *orfDR*'-*DsRed*.

Para la construcción del vector pME6102 se realizó una doble digestión *KpnI*-*SacI* sobre el vector pUC57 que contenía las fusiones *orfUR*'-*egfp* y *orfDR*'-*DsRed* escindiendo ambas en un único fragmento. Luego se clonó en el vector pME6000 en los mismos sitios y se transformó en *E. coli*. Para su movilización a *pseudomonas* se realizó una conjugación triparental.

1.5.8 Transferencia de plásmidos a pseudomonas por conjugación triparental

Como cepa dadora se utilizó *E. coli* DH5 α electroporada con cada uno de los diferentes plásmidos a transferir a la cepa aceptora de pseudomonas correspondiente. Para la conjugación triparental, se crecieron cultivos de la cepa dadora, la aceptora y una *helper* a 35°C hasta la fase exponencial de crecimiento (DO_{600} = 0.5-0.7) y se mezclaron 500 μ l de cada cultivo. Se centrifugó la mezcla a 2000 rpm por 4 min. El pellet se resuspendió en H₂O milliQ, se depositó sobre el borde de una placa de medio NA sin antibióticos y se incubó a 37°C o.n. Luego, la mezcla de conjugación se resuspendió en medio NYB líquido y se plaquearon diferentes diluciones en M9 citrato (citrato 0.4%, MgSO₄ 0.5 g/L, CaCl₂ 111 mg/L, Na₂HPO₄ 6.78 g/L, KH₂PO₄ 3 g/L, NaCl 0.5 g/L, NH₄Cl 1 g/L, agar 15 g/L) suplementado con los antibióticos correctos para llevar a cabo la selección deseada.

1.5.9 Cinéticas de expresión con reporteros fluorescentes

Para las cinéticas de expresión de cepas portadores de plásmidos reporteros con proteínas fluorescentes se realizaron cultivos en placas de 48 pocillos en medio GCM. El crecimiento y las medidas se llevaron a cabo utilizando un lector de placas multiwell Cytation 5 Agilent, a 28°C y 320 rpm de agitación doble orbital. Para estimar el crecimiento se midió densidad óptica a 600nm y para registrar fluorescencia se utilizaron monocromadores con los siguientes parámetros: eGFP ex=485, em=528; DsRed ex= 554, em= 586.

1.5.10 Expresión y purificación de las proteínas RsmA/E-his6

Se crecieron cultivos de 200ml de *E. coli* conteniendo el plásmido pME6078 o pME7010 hasta DO_{600} = 0.7 y se agregó IPTG (Isopropil- β -D-1-tiogalactopiranosido) hasta una concentración de 1 mM. Las células fueron colectadas a una DO_{600} = 1.4-1.6 y se lavaron en frío con 20 ml de buffer de lisis (NaCl 300 mM, imidazol 10 mM en buffer fosfato 50 mM, pH=8). Se resuspendió en el mismo buffer, se agregó lisozima 1 mg/ml y se incubó 45 min en hielo. Con un sonicador de punta de diámetro 6 mm (Sonics Vibra CellPower 130 Watts) se realizaron 10 tandas de sonication con 10 segundos de descanso en hielo. El lisado se centrifugó a 10.000 rpm por 30 min a 4°C. Se colectó el sobrenadante, se mezcló con 1 ml de resina Ni-NTA Agarosa (Qiagen®) al 50% y se cargó en una columna de 5ml. Luego de eluir toda la solución, se realizaron dos lavados con 4ml de buffer de lavado (NaCl 300 mM, imidazol 20 mM en buffer fosfato 50 mM, pH=8). Por último, las proteínas se eluyeron con 4 descargas de 500 μ l de buffer de elución (NaCl 300 mM, imidazol 250 mM en buffer fosfato 50 mM, pH=8) y fueron dializadas frente a un buffer tris-acetato 10 mM.

Para la visualización de las proteínas se realizó SDS-PAGE en geles de 15% acrilamida. Luego para la confirmación se realizó un Western blot, transfiriendo las proteínas del gel a una membrana de PVDF (Immobilon P; Millipore), durante 45 min a 4°C aplicando una corriente de 50 mA. La membrana se bloqueó con una solución de 5% leche descremada en buffer TBS (Tris 20 mM, NaCl 8 g/L, Tween 20 0.1%, pH = 7.6). Se incubó con un anticuerpo policlonal producido contra

RsmA de *Yersinia enterocolitica* (Reimann et al. 2005) o.n. a 4°C, se realizó un lavado con buffer TBS y se incubó con un anticuerpo policlonal de cabra anti-inmunoglobulina de conejo conjugada con fosfatasa alcalina (DakoCytomation/ Dako Agilent®). Finalmente se agregó una solución de Lumi-Phos 530 (Lumigen®) y se expuso a una placa fotográfica, que fue posteriormente revelada manualmente en cuarto oscuro.

1.5.11 Ensayos de cambio en la movilidad electroforética

Los ensayos de cambio en la movilidad electroforética, se realizaron según lo descrito en (Valverde 2009). Los oligoribonucleótidos sintéticos se renaturalizaron en presencia de MgCl₂ 10mM incubando por 10 min a 65°C seguido de un enfriamiento gradual hasta temperatura ambiente. Los cambios en la movilidad electroforética en geles de acrilamida nativos (15% acrilamida en TBE), producto de la formación de complejos ribonucleoproteicos, se detectaron por medio de una tinción con GelRed®.

1.6 Resultados y discusión

1.6.1 Estudio de la expresión de los genes *orfUR* y *orfDR* utilizando reporteros fusionados a *gfp* y *DsRed*

Como se mencionó previamente, nuestro laboratorio contaba con dos vectores que contienen funciones reporteras traduccionales de los genes *orfUR* u *orfDR* y '*lacZ*' (Figura 17). Si bien fue posible analizar el patrón de expresión de ambos genes, decidimos modificar las construcciones reemplazando el reportero '*lacZ*' por ORFs de proteínas fluorescentes para sortear las siguientes limitaciones: 1) el ensayo de actividad para determinar las unidades Miller consume más tiempo y reactivos que las determinaciones de fluorescencia; 2) la medición de actividad β-galactosidasa implica un método destructivo; 3) no es posible cuantificar ambas fusiones en simultáneo en el mismo contexto genético (cepa). 4) por último y más importante, las fusiones génicas a reporteros fluorescentes permiten su detección directa por microscopía.

Las nuevas construcciones se diseñaron *in-silico* y fueron sintetizadas y clonadas por Genscript® en el vector pUC57, luego fueron sub-clonadas en el vector pME6000, para dar origen al vector pME6102 (Tabla 1), y se movilizaron por conjugación triparental en las cepas CHA0, CHA19, CHA89, CHA1009 y CHA1144. Se realizaron cinéticas de crecimiento con el objetivo de verificar el correcto funcionamiento del sistema reportero. En la Figura 18, se observa la expresión de *orfUR*'-'*egfp* relativizada a la DO₆₀₀ de las cepas ensayadas, durante una cinética de 30 horas. Al igual que previamente para las fusiones traduccionales a '*lacZ*', se observó el patrón característico de un gen regulado por la cascada Gac-Rsm. La cepa silvestre expresó *egfp* cuando el cultivo comenzó a saturarse, mientras que la mutante CHA1009 lo hizo en forma prematura y

con mayor nivel de expresión. El resto de los mutantes incapaces de activar la cascada no expresaron *egfp* en niveles significativos.

En la Figura 19A, se observa la expresión de *orfDR'*-*DsRed* relativizada a DO_{600} , de la misma cinética previamente mencionada. Se observó el mismo patrón de expresión que para *orfUR'*-*gfp*, pero la expresión fue mucho más tardía. Este resultado no fue coherente con lo observado con las fusiones a *'lacZ*, donde si bien la expresión de cada gen indica tener un patrón característico de regulación traduccional por la cascada Gac-Rsm, se expresan dentro del mismo rango temporal. Cabe mencionar que los tiempos de expresión no son comparables con las cinéticas realizadas con los reporteros de *'lacZ*, ya que el sistema de cultivo es diferente. Estos cultivos son crecidos en placas multiwells, donde la agitación y la aireación no son tan eficientes como en un Erlenmeyer, y por lo tanto se obtienen velocidades de crecimiento menores. Para intentar comprender esta diferencia se repitió la cinética solo para ver DsRed, pero con una duración total de 65 horas. Como se puede ver en Figura 19B, DsRed parece seguir expresándose muchas horas pasadas la fase estacionaria, lo cual tampoco se corresponde a lo observado previamente, donde el gen *orfDR* alcanza su máximo de expresión, y se mantiene constante durante la fase estacionaria. Esto nos hizo plantearnos si realmente era un problema de expresión del gen, o de la proteína fluorescente. A diferencia de eGFP, DsRed es una proteína tetramérica, que requiere mayores tiempos para lograr su maduración.

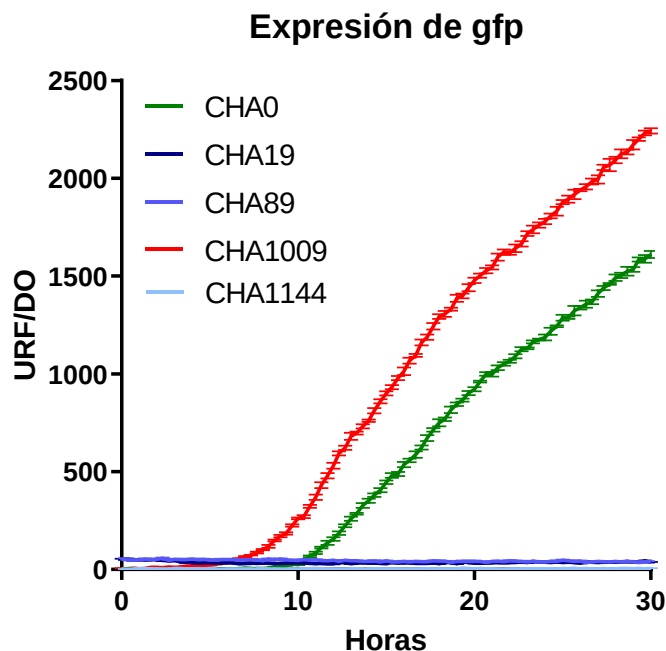


Figura 18. Cinética de crecimiento de las cepas de CHA0 (silvestre), CHA19(*gacA*:: Ω Km), CHA89 (Δ *gacS*), CHA1009 (*rsmA*:: Ω Km *rsmE*:: Ω Hg), CHA1144 (Δ *rsmXYZ*). Las cepas fueron crecidas en placas de 48 pocillos. Se relativizó la medida de fluorescencia correspondiente a eGFP sobre la DO_{600} .

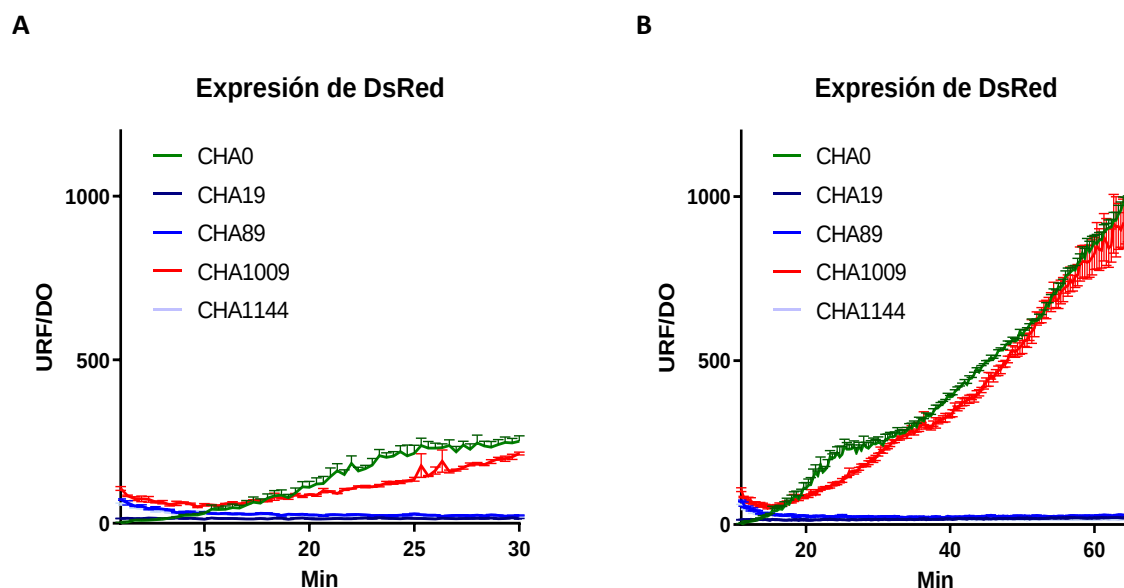


Figura 19 Cinética de crecimiento de las cepas de CHA0 (silvestre), CHA19(*gacA*:: Ω Km), CHA89 (Δ *gacS*), CHA1009 (*rsmA*:: Ω Km *rsmE*:: Ω Hg), CHA1144 (Δ *rsmXYZ*). Las cepas fueron crecidas en placas de 48 pocillos. Se relativizó la medida de fluorescencia correspondiente a DsRed sobre la DO_{600} . La Figura B es la repetición de la cinética de la Figura A, pero con una duración de 65 horas. Las primeras horas fueron eliminadas debido a que el inóculo contiene una alta acumulación de la proteína DsRed, la cual genera mucho error.

Para evaluar si la expresión tardía de DsRed se trató de un problema de expresión o de maduración, se realizaron dos cinéticas de crecimiento con las cepas CHA0, CHA19, CHA1009 y CHA1144, conteniendo el vector reportero pME7019 (*orfDR'*-*'lacZ*) y paralelamente otra cinética con las mismas cepas conteniendo el vector reportero pME6102 (*orfUR'*-*'egfp*; *orfDR'*-*'DsRed*). La cinética se realizó durante 8h comenzando a medir a partir de la hora 9 (fase exponencial tardía). Como se puede ver en la Figura 20A, el reportero de *'lacZ* a las 9 horas de cultivo ya se encontraba en su máximo de expresión, y se mantuvo en un nivel relativamente constante. Por el contrario, en la Figura 20B, se pudo observar fluorescencia roja a partir de la hora 14. Ambos cultivos fueron crecidos de igual manera, y comparten regiones regulatorias, por lo tanto, deben expresarse de igual manera. La diferencia observada se debe entonces al tiempo de maduración de la proteína DsRed, y no a un retardo en la expresión. La variante original de DsRed puede necesitar 21h aproximadamente para lograr su maduración a 28°C (Hakkila et al. 2002). Si bien la variante utilizada en esta fusión fue la DsRed-express, la cual tiene menor tiempo de maduración, en las condiciones de nuestro cultivo, puede demorarse hasta 8h la aparición de fluorescencia. Además, la proteína DsRed parece tener una vida media significativamente mayor que eGFP; si vemos la Figura 19B, la intensidad de fluorescencia no se estabiliza pasadas las 65 horas, sino que se acumula indefinidamente.

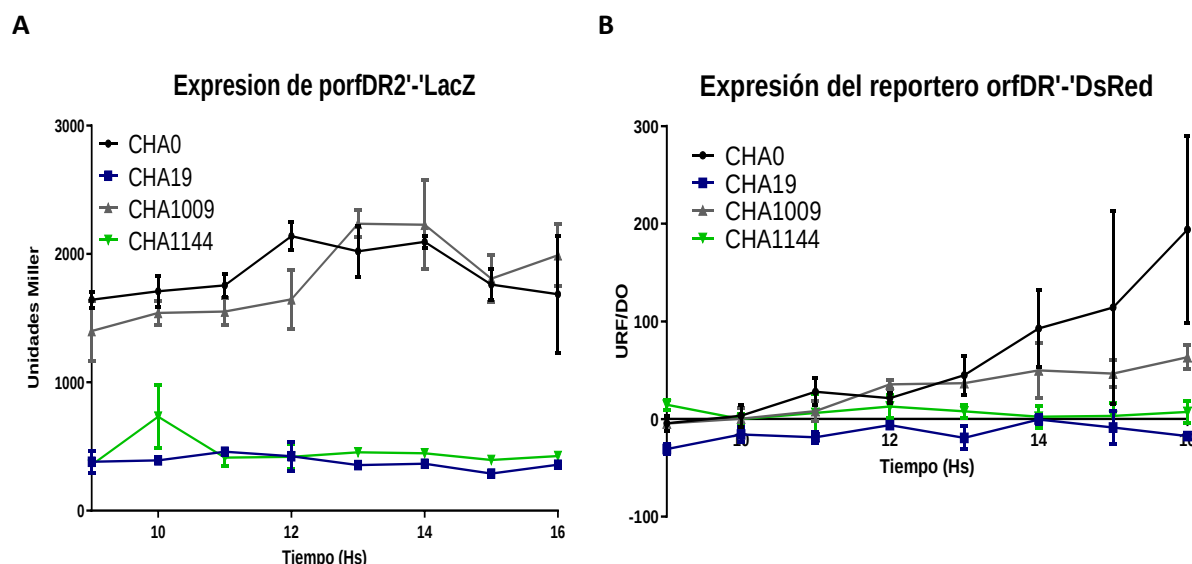


Figura 20. Cinética de crecimiento de las cepas de CHA0 (silvestre), CHA19(*gacA*:: Ω Km), CHA1009 (*rsmA*:: Ω Km *rsmE*:: Ω Hg), CHA1144 (Δ *rsmXYZ*). Las cepas fueron crecidas en Erlenmeyer, la toma de muestras comenzó en la hora 9, y se tomaron medidas cada una hora. En el panel A se utilizó el reportero *orfDR'*-*lacZ*, y se realizó un ensayo de β -galactosidasa. En el Panel B se utilizó el reportero *orfDR'*-*DsRed* y se registró el valor de fluorescencia correspondiente a DsRed. Las medidas de fluorescencia o unidades Miller fueron relativizadas sobre la DO_{600} .

Las fusiones *orfUR'*-*egfp* y *orfDR'*-*DsRed* resultaron ser de utilidad para reportar la expresión de ambos genes *in vivo*, sin embargo la proteína DsRed presentó la desventaja de tener mayor tiempo de maduración, y elevada estabilidad, lo cual no la convierte en buen candidato para utilizarla como reportero *in vivo* de la expresión de un gen. Para solucionar dichos problemas se planteó reemplazo del ORF de *DsRed* por *mCherry*, como se describe en el Capítulo 2. Sin embargo, la fusión *orfDR'*-*DsRed* aún tiene el potencial de utilizarse como marca fluorescente en ensayos de microscopía.

1.6.2 Estudio sobre regulación directa de las proteínas represoras RsmA/E sobre los reguladores transcripcionales *orfUR* y *orfDR*

Luego de realizar un análisis *in-silico* sobre posibles genes regulados por el sistema Gac-Rsm, los genes *orfUR* y *orfDR* fueron postulados como candidatos. Se encontraron posible sitios de unión de las proteínas represoras RsmA y RsmE sobre la región 5'-UTR, solapando el sitio de unión a ribosomas predicho por Schubert et al. (2007) (Figura 21). El posible control de la traducción de estos mRNAs por parte de las proteínas RsmA y RsmE, había sido previamente validada mediante el análisis de la actividad de fusiones reporteras traduccionales a *lacZ* (Figura 17) (Sobrero and Valverde 2020). Para sumar evidencia independiente a la hipótesis de la regulación directa se realizó un ensayo de interacción entre las proteínas RsmA/E y fragmentos de ARNs sintéticos que contenían el sitio predicho de unión.

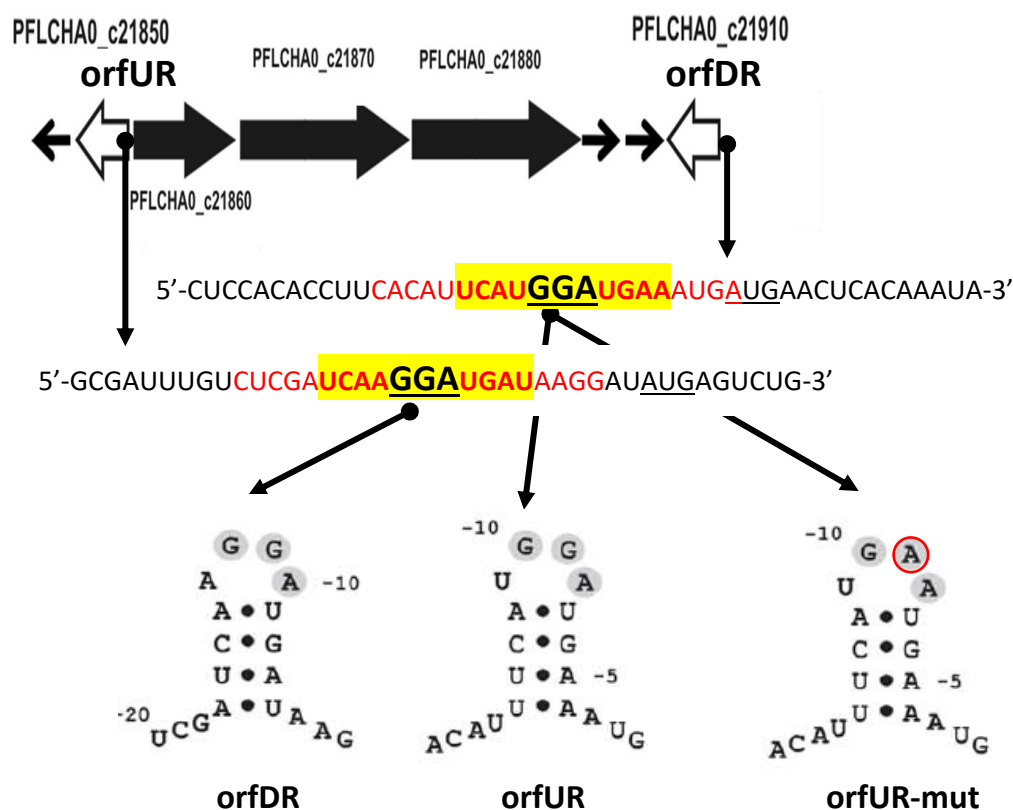


Figura 21. Esquema de las secuencias utilizadas para los ensayos de interacción. En la parte superior se encuentra el *cluster* de biosíntesis de lipopéptidos de CHA0. En la parte central se identifican fragmentos de la región líder de los ARNm de *orfUR* y *orfDR*. En rojo se resaltan las secuencias diseñadas para realizar ensayos de interacción *in vitro*. La región con fondo amarillo corresponde al sitio de unión de las proteínas represoras RsmA/E (Schubert et al. 2007), y dentro de esta región se encuentra el triplete GGA, indispensable para la interacción con RsmA/E según la predicción del *loop* formado en esta región. Los tripletes AUG subrayados corresponden al codón de inicio de la traducción. Las estructuras secundarias de los ARNs fueron tomadas de Sobrero et al. (2017).

1.6.3 Purificación de las proteínas RsmA y RsmE

Se purificaron las proteínas etiquetadas RsmA-His₆ y RsmE-His₆ según el protocolo mencionado en Materiales y métodos. Se confirmó la presencia de las proteínas por SDS-PAGE seguido de un *Western blot* revelado con anticuerpos específicos (Figura 22).

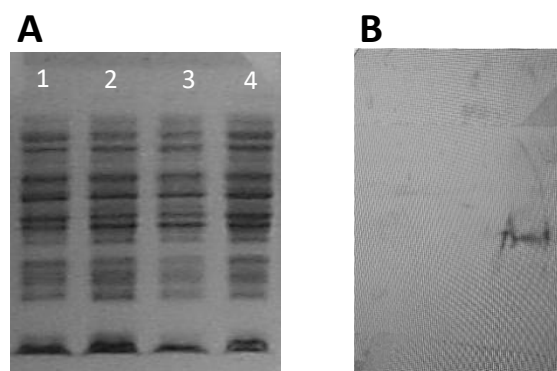


Figura 22 .A) SDS-PAGE. 1) y 2) Lisado de *E. coli* conteniendo el plásmido pME6078 sin inducir e inducido. 3) y 4) Lisado de *E. coli* conteniendo el plásmido pME7010 sin inducir e inducido. B) Western blot de RsmA-His6 purificada.

Se diseñaron dos oligorribonucleótidos simple cadena de 20 nucleótidos, conteniendo la región predicha de unión a las proteínas RsmA/E de los mensajeros *orfUR* y *orfDR* (Tabla 2, Figura 21). Para el *orfUR* se sintetizó también un oligo conteniendo una sustitución de una base que está involucrada en la formación del *loop* necesario para la interacción con las proteínas RsmA/E (Figura 21).

Tabla 2

Origen del oligo de RNA	Secuencia
<i>orfUR</i>	5'-CACATT CATGGATGAA ATGA-3'
<i>orfUR</i> -mutado	5'-CACATT CATGA A ATGAAATGA-3'
<i>orfDR</i>	5'-CTCGAT CAAGGATGATA AGG-3'

Las regiones remarcadas en rojo corresponden al sitio predicho de unión de RsmA (Schubert et al. 2007). El remarcado amarillo corresponde a la base modificada de la secuencia silvestre

Cada oligorribonucleótido fue incubado con RsmA-His₆ o RsmE-His₆ en concentraciones crecientes, y se analizó el retardo en la corrida electroforética en geles de acrilamida nativos (Figura 23). Todos los geles fueron teñidos con el colorante GelRed®, el cual es un intercalante de ácidos nucleicos, por lo tanto, las bandas visibles corresponden a moléculas de ARN. La última calle de cada gel corresponde a una muestra de RsmA-His₆ o RsmE-His₆, donde se pueden ver dos bandas marcadas con * que corresponden a ARNs co-purificados con las proteínas, como fue descrito por Valverde et al. (2003) para la purificación de RsmA. Si bien fueron expresadas en *E. coli*, existen cientos de ARNs que son regulados naturalmente por la proteína homóloga CsrA, que también serán capaces de interactuar con las variantes His6 de los homólogos RsmA/E.

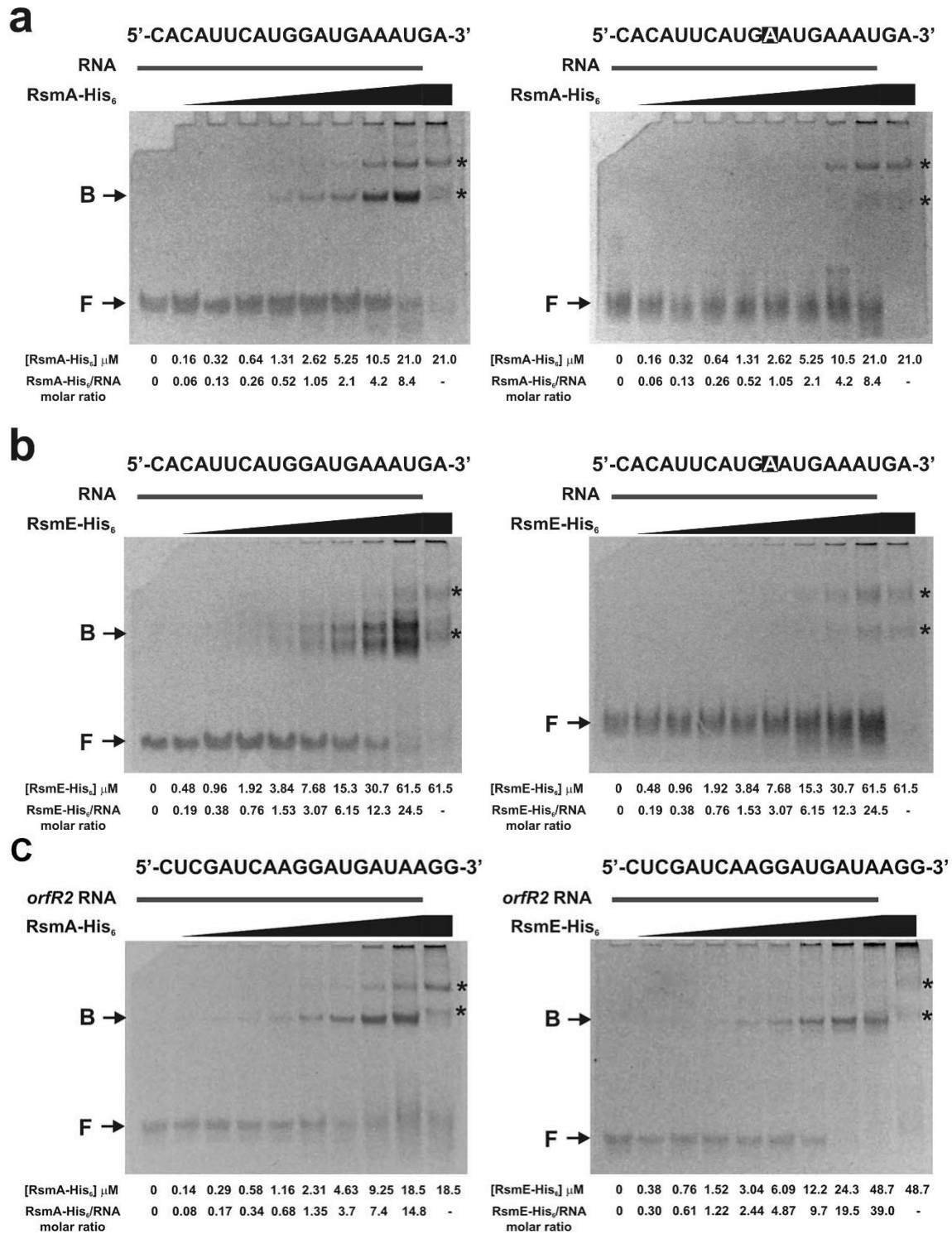


Figura 23. Se incubaron diferentes concentraciones de RsmA-His6 o RsmE-His6 con cada oligonucleótido de ARN (que se muestra arriba del panel) previo a la corrida en geles nativos. Se realizó un teñido con GelRed. A. *orfUR* silvestre (izquierda) o mutante (derecha) vs RsmA-His6. B. *orfUR* silvestre (izquierda) o mutante (derecha) vs RsmE-His6. La base mutada se encuentra sombreada. C. *orfDR* silvestre (izquierda) vs RsmA-His6 o RsmE-His6 (derecha). F y B indican las posiciones de las respectivas especies de ARN libres (F) o unidas (B).

El asterisco en la derecha de los geles señala los ARNs co-purificados con la proteína recombinante cargados en la calle sin oligonucleótidos de ARN. Figura tomada de Sobrero et al. (2017)

En el panel A de la Figura 23, se enfrenta el oligo de la región 5'UTR de *orfUR* silvestre a la izquierda, y el mutado a la derecha, con la proteína RsmA. A concentraciones intermedias (1.31 μ M) de proteína, se puede ver que el oligo silvestre cambia su posición en el gel a una de mayor peso molecular, que corresponde al complejo ARN-proteína (B). A medida que se aumenta la concentración de proteína aumenta la intensidad de la banda, secuestrando mayor cantidad de ARN, y a su vez disminuye la intensidad de la banda correspondiente al ARN libre (F). Por el contrario, a medida que se aumenta la concentración de proteína frente al oligo mutado, ninguna de las bandas modifica su posición, evidenciando que no existe interacción en las condiciones de este ensayo. Lo mismo ocurre cuando enfrentamos los mismos oligos contra la proteína RsmE (panel B Figura 23). En el panel C Figura 23, se encuentran los geles del enfrentamiento entre *orfDR* y RsmA a la izquierda, y *orfDR* y RsmE a la derecha. Se pudo observar nuevamente el complejo RNA-proteína a concentraciones similares.

En vista de los resultados obtenidos, se pudo determinar que las proteínas RsmA/E son capaces de interaccionar de forma directa con las regiones líderes de los ARNm de los *orfUR* y *orfDR*. Además, la integridad del triplete GGA resultó ser determinante para dicha interacción.

1.6.4 Vinculo regulatorio entre la cascada Gac-Rsm y la biosíntesis de lipopéptidos en *P. protegens* CHAO

Anteriormente, se determinó el patrón de expresión de los genes *orfUR* y *orfDR* mediante fusiones traduccionales a '*lacZ*', observando que existía una regulación positiva por la cascada Gac-Rsm. Además, fue posible determinar que las proteínas RsmA/E ejercen un control directo sobre sus ARNm, modulando su traducción. Por otro lado, ambos genes poseen una alta homología con reguladores de la familia LuxR y se encuentran flanqueando un *cluster* de biosíntesis de lipopéptidos (Figura 16). En combinación, todos estos resultados derivaron en la hipótesis de que los genes *orfUR* y *orfDR* se encontraban involucrados en la regulación de lipopéptidos en *P. protegens* CHAO. Para comprobar si la regulación por Gac-Rsm podría estar controlando la biosíntesis de lipopéptidos se evaluó la presencia de surfactantes en el sobrenadante de cultivo de diferentes mutantes isogénicos de CHAO deficientes en genes de la cascada. Debido a su naturaleza química, los lipopéptidos son compuestos tensioactivos capaces de ser detectados por medio del colapso de una gota de sobrenadante de cultivo sobre una superficie de Parafilm®. Las gotas de sobrenadantes de cultivos de mutantes deficientes en la activación del sistema GacA/GacS o que no poseen los ARNs miméticos no colapsaron, mientras que, en la cepa silvestre, o en el mutante carente de las proteínas represoras RsmA/E la gota si colapso. Una vez confirmado que la respuesta era coincidente al patrón de expresión de *orfUR* y *orfDR* (Figura 24), se procedió a la purificación, caracterización y estudio de la producción de lipopéptidos en CHAO.

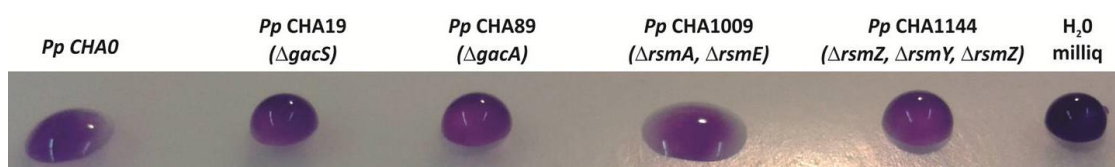


Figura 24. Colapso de gota de sobrenadante de las distintos mutantes isogénicos de CHA0. Se agregó colorante Coomassie blue para resaltar las gotas.

1.6.4.1 Optimización de la producción de surfactantes

El primer paso consistió en identificar y caracterizar el metabolito responsable de la actividad surfactante del sobrenadante de cultivo de CHA0. Previamente al paso de aislamiento, se seleccionó un medio de cultivo óptimo para la producción de surfactante. Se utilizaron 4 medios de cultivos diferentes para la producción de surfactantes estimando su concentración a través de la actividad de reducción de la tensión superficial del medio de cultivo por el método de colapso de gota. Se generaron cultivos de *P. protegens* CHA0 en los medios King' s B, PSFM, DMB glucosa y DBM manitol, se tomó su sobrenadante y se realizaron diluciones seriadas en agua destilada hasta obtener la máxima dilución posible sin perder el colapso de gota. Esta medida se relativizó a DO_{600} del cultivo y se utilizó para estimar la producción de surfactantes (Figura 25). De aquí en adelante todos los cultivos destinados a la producción de surfactantes se realizaron en PSFM ya que el sobrenadante de cultivo presentó la mayor actividad tensioactiva en comparación con el resto. En otros medios de cultivo previamente utilizados, pudimos detectar que la actividad tensioactiva se perdía cuando CHA0 sufría cualquier mutación que interrumpiese el funcionamiento de la cascada Gac-Rsm. Para confirmar que dicho fenotipo no se vea modificado por el cambio de medio de cultivo se analizó la actividad tensioactiva de distintos mutantes afectados en componentes de la cascada Gac-Rsm creciendo en PSFM (Figura 26). Si comparamos con la Figura 18 y Figura 19, el patrón de actividad surfactante resultó coherente con la expresión de los reguladores *orfUR* y *orfDR*, es decir, es necesario que ambos reguladores se expresen para que el sobrenadante de cultivo presente actividad surfactante.

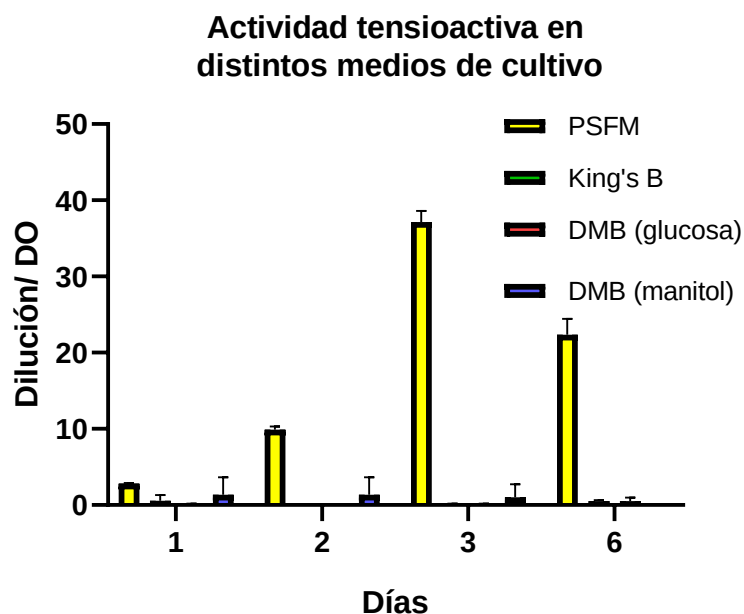


Figura 25. Dilución máxima de sobrenadante de cultivo obtenida previa a perder la actividad surfactante, relativizado a la DO del cultivo.

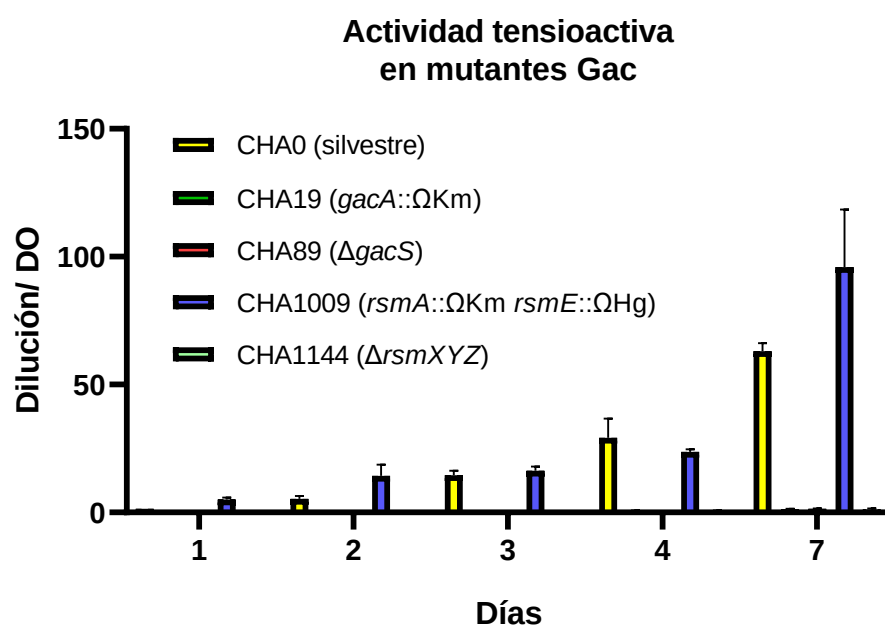


Figura 26. Actividad tensioactiva en sobrenadante de cultivo de CHA0 (silvestre), CHA19(*gacA::ΩKm*), CHA89 ($\Delta gacS$), CHA1009 (*rsmA::ΩKm rsmE::ΩHg*), CHA1144 ($\Delta rsmXYZ$) creciendo en PSFM.

1.6.4.2 Purificación e identificación de lipopéptidos de *P. protegens*

Se crecieron cultivos de las siguientes cepas en PSFM para realizar una extracción de lipopéptidos de sobrenadantes como se detalla en el apartado MyM: *P. protegens* CHA0, CHA19, CHA89, CHA1009, CHA1144. Todos los extractos fueron resuspendidos en metanol y analizados por HPLC. Como se muestra en la Figura 27, los cromatogramas de las cepas CHA19 (*gacA::ΩKm*),

CHA89 ($\Delta gacS$), y CHA1144 ($\Delta rsmXYZ$) no poseen los picos marcados con flechas (min 48-65), que podrían corresponder a lipopéptidos. Estos picos son observables en la cepa silvestre CHA0 y son significativamente más abundantes en la cepa CHA1009 ($rsmA::\Omega Km$ $rsmE::\Omega Hg$), en la cual los genes regulados por la cascada Gac-Rsm se encuentran sobreexpresados a nivel traduccional. Esto resultó coherente con los ensayos de colapso de gota, y confirmatorio para asociar la expresión de surfactantes en CHA0 con la cascada Gac-Rsm.

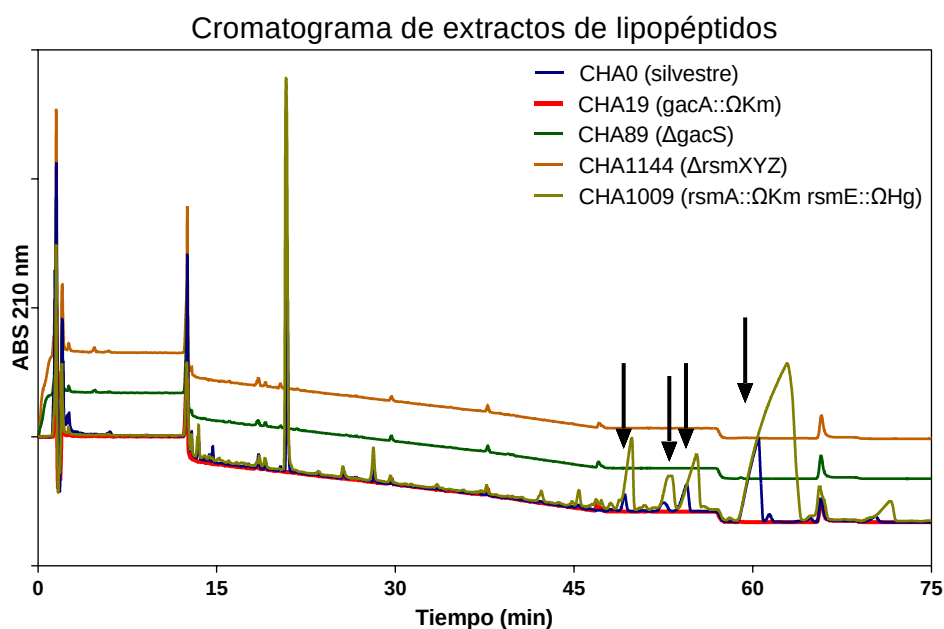


Figura 27. Perfil de extractos de lipopéptidos por HPLC de las cepas CHA0 (silvestre), CHA19 ($gacA::\Omega Km$), CHA89 ($\Delta gacS$), CHA1009 ($rsmA::\Omega Km$ $rsmE::\Omega Hg$), CHA1144 ($\Delta rsmXYZ$). Las flechas negras indican los picos que fueron colectados para futuros análisis.

Durante el análisis por HPLC, se colectaron 4 fracciones correspondientes a los supuestos lipopéptidos de CHA0 y de CHA1009, con el objetivo de dilucidar su estructura química. En ese momento Ma et al. (2016) reportó un análisis químico sobre lipopéptidos de distintas *pseudomonas*, entre ellas CHA0, confirmando que ésta produce el lipopéptido orfamida en sus isoformas A, B, C y G (Tabla 3). Por lo tanto, decidimos analizar el pico mayoritario por espectrometría de masas para confirmar que esta fuera orfamida. Se utilizó la fracción colectada de la mutante $rsmA::\Omega Km$ $rsmE::\Omega Hg$ (CHA1009), y no de CHA0, ya que la doble mutante producía significativamente mayor cantidad de lipopéptidos (Figura 27). Dicha muestra fue nuevamente separada por HLPC, dividida en tres fracciones que luego fueron analizadas por espectrometría de masas, confirmando según el peso del ion molecular, que se trataba de la orfamida A, y en menor proporción orfamida G (Figura 28). El ion molecular de la orfamida A $[M+H]^+$ tienen una masa teórica de 1295.7 Da, sin embargo, el ion puede asociarse con Na^+ o K^+ , presentando una masa de 1317.6 y 1333.7 respectivamente, la cual coincide casi perfectamente con los dos iones mayoritarios en los paneles A y B de la Figura 28. El espectro del panel C corresponde a la orfamida G, ionizada con Na^+ , la cual tiene una masa teórica de 1337.9 que coincide a la perfección con el ion mayoritario de nuestro espectro.

Tabla 3. Estructura de las distintas especies de orfamidas aislados de diferentes pseudomonas.

Compound	Fatty acid residue	Amino acid sequence									
		1	2	3	4	5	6	7	8	9	10
Orfamide A	C14:0-OH (3)	L-Leu	D-Glu	D-aThr	D-Ile	L-Leu	D-Ser	L-Leu	L-Leu	D-Ser	L-Val
Orfamide B	C14:0-OH (3)	L-Leu	D-Glu	D-aThr	D-Val	L-Leu	D-Ser	L-Leu	L-Leu	D-Ser	L-Val
Orfamide C	C12:0-OH (3)	L-Leu	D-Glu	D-aThr	D-Ile	L-Leu	D-Ser	L-Leu	L-Leu	D-Ser	L-Val
Orfamide D	C12:0-OH (3)	L-Leu	D-Glu	D-aThr	D-Val	L-Leu	D-Ser	L-Leu	L-Leu	D-Ser	L-Val
Orfamide E	C14:1-OH (3)	L-Leu	D-Glu	D-aThr	D-Val	L-Leu	D-Ser	L-Leu	L-Leu	D-Ser	L-Val
Orfamide F	C16:1-OH (3)	L-Leu	D-Glu	D-aThr	D-Val	L-Leu	D-Ser	L-Leu	L-Leu	D-Ser	L-Val
Orfamide G	C16:0-OH (3)	L-Leu	D-Glu	D-aThr	D-Val	L-Leu	D-Ser	L-Leu	L-Leu	D-Ser	L-Val

P. protegens CHA0 produce principalmente la orfamida A, y en menos proporción la B, C y G. Modificado de Ma et al. (2016).

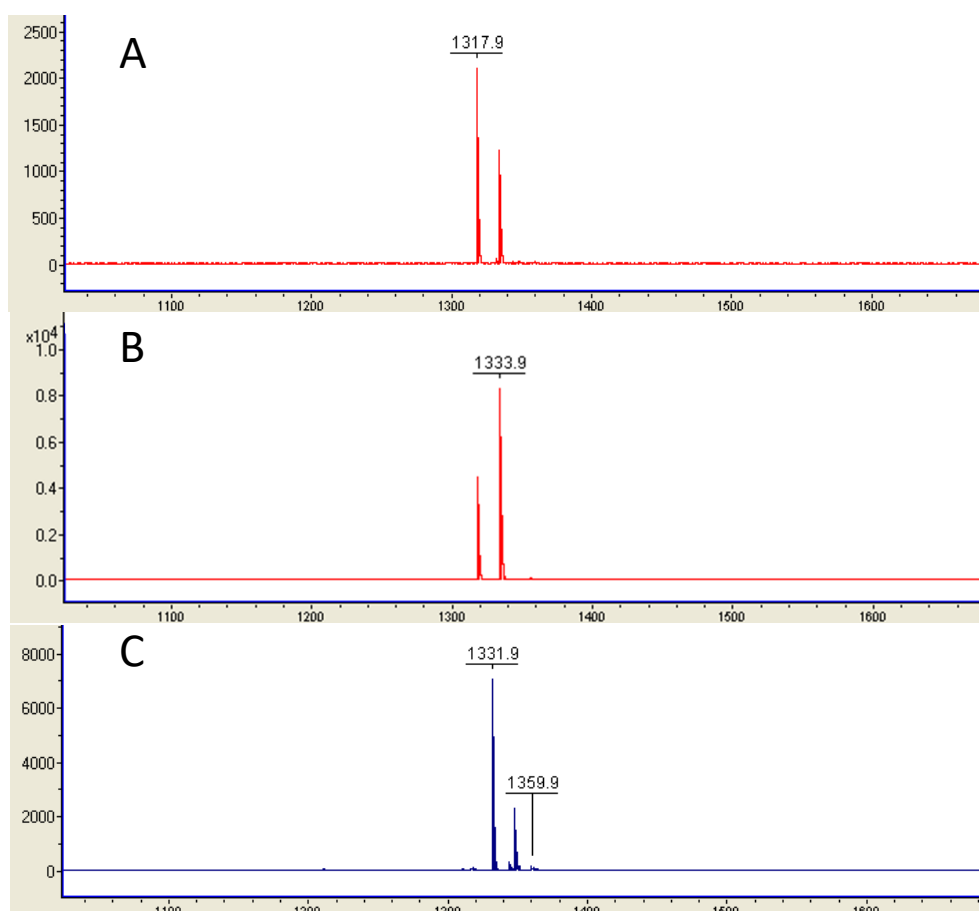


Figura 28. Espectro de masas de la fracción del pico mayoritario del cromatograma de *P. protegens* CHA1009. La muestra nuevamente se dividió por HPLC en tres y así se obtuvieron los espectros A, B y C.

De esta manera quedó establecido que CHA0 efectivamente produce orfamidas cuando crece en cultivo puro, pero no lo hace cuando se interrumpe alguno de los genes que codifican elementos esenciales para el correcto funcionamiento de la cascada Gac-Rsm, como las proteínas GacS o GacA, o los ARNs imitadores moleculares RsmX, RsmY y RsmZ (antagonistas de RsmA y RsmE). En forma consistente, observamos una mayor acumulación de orfamidas cuando

se interrumpen los genes para las proteínas represoras RsmA/E. En conjunto, estos resultados indican que la cascada Gac-Rsm promueve la biosíntesis de orfamidas en CHA0.

1.6.5 Modelo regulatorio de la producción de orfamidas por la cascada Gac-Rsm

En resumen, teniendo en cuenta los perfiles de expresión de las fusiones reporteras de los genes *orfUR* y *orfDR* (Figura 18 y Figura 19) y los hallazgos en los ensayos de cambio de movilidad de ARN (Figura 23), en su conjunto los resultados respaldan firmemente la hipótesis que las secuencias y motivos identificados dentro de las regiones líderes de los ARNm de *orfUR* y *orfDR* constituyen los sitios de unión para las proteínas represoras de la traducción RsmA y RsmE, y que éstos son fundamentales para la activación de la traducción de los ARNm por parte de la cascada Gac-Rsm operativa (Figura 29).

La participación de la cascada Gac-Rsm en el control de la producción de lipopéptidos en *Pseudomonas* probióticas de plantas fue reportada hace más de una década en *Pseudomonas* sp. DSS73 en relación con la producción de amfisina (Koch et al. 2002) y más recientemente con respecto a la síntesis de esclerosina en *Pseudomonas* sp. DF41 (Berry et al. 2012), el lipopéptido responsable del principio inductor de la línea blanca (*white-line-inducing principle*, WLIP) en *P. putida* RW10S2 (Rokni-Zadeh et al. 2012) y massetolido en *P. fluorescens* SS101 (Song et al. 2015). En la especie *P. protegens*, se reportó la necesidad de un sistema Gac-Rsm funcional para la producción de orfamidas en la cepa Pf-5; para la cual la inactivación de *gacA* resultó en la reducción de 64 veces en la producción de orfamida A (Hassan et al. 2010), y mutaciones espontáneas en *gacA* o *gacS* produjeron una disminución en la síntesis de orfamida A (Song et al. 2016). Ambos estudios también informaron una regulación a la baja del ortólogo *orfUR* Pfl_2150 (Hassan et al. 2010; Song et al. 2016) y del ortólogo *orfDR* Pfl_2143 (Song et al. 2016) en ausencia de una proteína GacA funcional, sugiriendo que el control podría ser una consecuencia de una mayor degradación de ARNm después del bloqueo de la traducción por parte de RsmA y RsmE. Nuestros resultados (Figura 18, Figura 19, Figura 23) proporcionan la primera evidencia de un control directo de Gac-Rsm sobre los genes que codifican los dos reguladores transcripcionales de la biosíntesis de lipopéptidos en *Pseudomonas*. Además, es interesante resaltar que el control traduccional de la cascada Gac-Rsm sobre la expresión del operón *ofaABC* resulta jerárquicamente más relevante que el control transcripcional. En general, la regulación transcripcional ocurre en una primera instancia, siendo cuantitativamente la más importante, y luego ocurre la regulación postranscripcional ejerciendo un control más ajustado, contrariamente a lo observado en la Figura 29. Es probable entonces, que la inversión de jerarquías en esta cascada, sea producto de la necesidad de coordinar la biosíntesis de orfamidas en conjunto con el resto de los fenotipos controlados por Gac-Rsm, que se van a co-expresar en las condiciones en las que la cascada esté operativa (sección 1.2.3) (Lapouge et al. 2008).

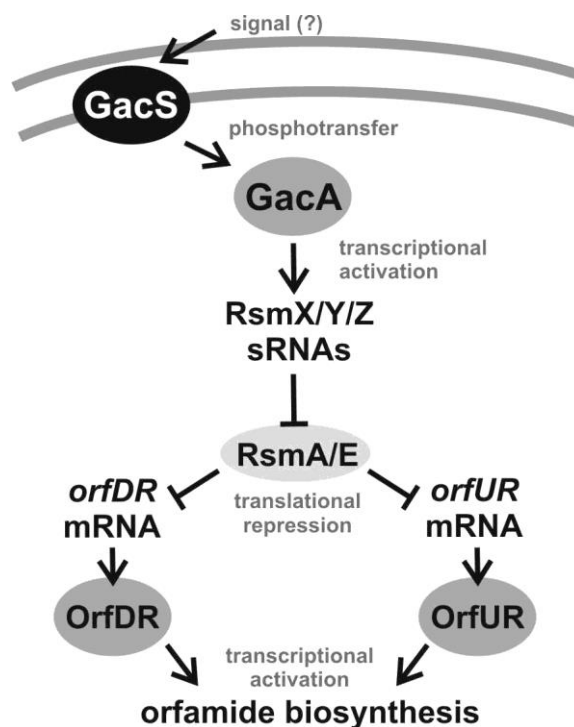


Figura 29. Modelo del control jerárquico por la cascada Gac-Rsm de los reguladores transcripcionales del cluster de genes biosintéticos de orfamidas en *P. protegens* CHA0. Figura tomada de Sobrero et al. (2017)

1.7 Conclusión

En *P. protegens* CHA0, el control postranscripcional de la expresión génica a través de la vía Gac-Rsm tiene un gran impacto en el metabolismo secundario y, en consecuencia, en varios fenotipos relacionados con la capacidad de esta bacteria del suelo para ejercer su poder biocontrolador frente a diferentes fitopatógenos. y regular su interacción con eucariotas (Haas and Défago 2005; Lapouge et al. 2008). Previamente en el laboratorio, se llevó a cabo una búsqueda *in-silico* de nuevos genes y procesos que pueden estar sujetos al control de la cascada Gac-Rsm sobre la base de restricciones secuenciales, estructurales y evolutivas para el sitio de unión RsmA/E. Dicho enfoque detectó un conjunto de 43 nuevos posibles *targets* de la vía Gac-Rsm. Dos de los más destacados candidatos descubiertos correspondían a un par de reguladores transcripcionales de tipo LuxR (*orfUR* y *orfDR*) codificados en las inmediaciones del *cluster* de genes encargado de la producción de orfamida A (Sobrero et al. 2017). A partir de estos resultados, en este capítulo se proporcionó evidencia sólida de un control directo de Gac-Rsm sobre la traducción de los ARNm de *orfUR* y *orfDR* (Figura 18, Figura 19 y Figura 23) y, en consecuencia, sobre la producción de orfamidas en *P. protegens* CHA0 (Figura 27). De esta forma, hemos demostrado que la producción de orfamidas es otro rasgo que forma parte del regulón Gac-Rsm en *P. protegens* CHA0.

1.8 Bibliografía

- Al-Hinai, A., et al. (2010). "Isolation and characterization of *Pseudomonas aeruginosa* with antagonistic activity against *Pythium aphanidermatum*." 653-660.
- Altegoer, F., et al. (2016). "Structural basis for the CsrA-dependent modulation of translation initiation by an ancient regulatory protein." 113(36): 10168-10173.
- Alvarez, F., et al. (2012). "The plant-associated *Bacillus amyloliquefaciens* strains MEP218 and ARP23 capable of producing the cyclic lipopeptides iturin or surfactin and fengycin are effective in biocontrol of sclerotinia stem rot disease." 112(1): 159-174.
- Audenaert, K., et al. (2002). "Induction of systemic resistance to *Botrytis cinerea* in tomato by *Pseudomonas aeruginosa* 7NSK2: role of salicylic acid, pyochelin, and pyocyanin." 15(11): 1147-1156.
- Babitzke, P., et al. (2019). "Posttranscription initiation control of gene expression mediated by bacterial RNA-binding proteins." 73: 43-67.
- Bender, C. L., et al. (1999). "Pseudomonas syringae phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases." 63(2): 266-292.
- Beneduzi, A., et al. (2012). "Plant growth-promoting rhizobacteria (PGPR): their potential as antagonists and biocontrol agents." 35: 1044-1051.
- Berry, C. L., et al. (2012). "Chemical and biological characterization of sclerosin, an antifungal lipopeptide." 58(8): 1027-1034.
- Blumer, C., et al. (1999). "Global GacA-steered control of cyanide and exoprotease production in *Pseudomonas fluorescens* involves specific ribosome binding sites." 96(24): 14073-14078.
- Boopathi, E., et al. (1999). "A siderophore from *Pseudomonas putida* type A1: structural and biological characterization." 1435(1-2): 30-40.
- Boyer, H. W. and D. J. J. o. m. b. Roulland-Dussoix (1969). "A complementation analysis of the restriction and modification of DNA in *Escherichia coli*." 41(3): 459-472.
- Buyer, J. S., et al. (1986). "Structure of pseudobactin A214, a siderophore from a bean-deleterious *Pseudomonas*." 25(19): 5492-5499.
- Cawoy, H., et al. (2015). "Lipopeptides as main ingredients for inhibition of fungal phytopathogens by *Bacillus subtilis*/amyloliquefaciens." 8(2): 281-295.
- Cezairliyan, B., et al. (2013). "Identification of *Pseudomonas aeruginosa* phenazines that kill *Caenorhabditis elegans*." 9(1): e1003101.
- Chambonnier, G., et al. (2016). "The hybrid histidine kinase LadS forms a multicomponent signal transduction system with the GacS/GacA two-component system in *Pseudomonas aeruginosa*." 12(5): e1006032.
- Crick, F. J. N. (1970). "Central dogma of molecular biology." 227(5258): 561-563.
- D'aes, J., et al. (2014). "To settle or to move? The interplay between two classes of cyclic lipopeptides in the biocontrol strain *Pseudomonas* CMR 12a." 16(7): 2282-2300.
- Dambach, M. D. and W. C. J. C. o. i. m. Winkler (2009). "Expanding roles for metabolite-sensing regulatory RNAs." 12(2): 161-169.
- De Bruijn, I., et al. (2008). Massetolide A biosynthesis in *Pseudomonas fluorescens*, Am Soc Microbiol.

- De Bruijn, I., et al. (2007). "Genome-based discovery, structure prediction and functional analysis of cyclic lipopeptide antibiotics in *Pseudomonas* species." 63(2): 417-428.
- De Bruijn, I., et al. (2009). "Diversity and functional analysis of LuxR-type transcriptional regulators of cyclic lipopeptide biosynthesis in *Pseudomonas fluorescens*." 75(14): 4753-4761.
- de Souza, J. T., et al. (2003). "Effect of 2, 4-diacetylphloroglucinol on *Pythium*: cellular responses and variation in sensitivity among propagules and species." 93(8): 966-975.
- Dugar, G., et al. (2016). "The CsrA-FliW network controls polar localization of the dual-function flagellin mRNA in *Campylobacter jejuni*." 7(1): 1-18.
- Franze de Fernandez, M., et al. (1968). "Factor fraction required for the synthesis of bacteriophage Q β -RNA." 219(5154): 588-590.
- Gilchrist, F. J., et al. (2013). "Hydrogen cyanide concentrations in the breath of adult cystic fibrosis patients with and without *Pseudomonas aeruginosa* infection." 7(2): 026010.
- Giuliodori, A. M., et al. (2010). "The cspA mRNA is a thermosensor that modulates translation of the cold-shock protein CspA." 37(1): 21-33.
- Gray, E., et al. (2005). "Intracellular and extracellular PGPR: commonalities and distinctions in the plant-bacterium signaling processes." 37(3): 395-412.
- Gross, H. and J. E. J. N. p. r. Loper (2009). "Genomics of secondary metabolite production by *Pseudomonas* spp." 26(11): 1408-1446.
- Gross, H., et al. (2007). "The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters." 14(1): 53-63.
- Guttenberger, N., et al. (2017). "Recent developments in the isolation, biological function, biosynthesis, and synthesis of phenazine natural products." 25(22): 6149-6166.
- Haas, D. and G. J. N. r. m. Défago (2005). "Biological control of soil-borne pathogens by fluorescent pseudomonads." 3(4): 307-319.
- Hakkila, K., et al. (2002). "Reporter genes lucFF, luxCDABE, gfp, and dsred have different characteristics in whole-cell bacterial sensors." 301(2): 235-242.
- Hassan, K. A., et al. (2010). "Inactivation of the GacA response regulator in *Pseudomonas fluorescens* Pf-5 has far-reaching transcriptomic consequences." 12(4): 899-915.
- Hayat, R., et al. (2010). "Soil beneficial bacteria and their role in plant growth promotion: a review." 60(4): 579-598.
- Heeb, S., et al. (2002). "Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in *Pseudomonas fluorescens* CHA0." 184(4): 1046-1056.
- Hoegy, F., et al. (2014). Pyoverdine and pyochelin measurements. *Pseudomonas methods and protocols*, Springer: 293-301.
- Höfte, M. and N. J. R. i. M. Altier (2010). "Fluorescent pseudomonads as biocontrol agents for sustainable agricultural systems." 161(6): 464-471.
- Iavicoli, A., et al. (2003). "Induced systemic resistance in *Arabidopsis thaliana* in response to root inoculation with *Pseudomonas fluorescens* CHA0." 16(10): 851-858.
- Ishihama, Y., et al. (2008). "Protein abundance profiling of the *Escherichia coli* cytosol." 9(1): 1-17.
- Jeong, Y., et al. (2003). "Toxoflavin produced by *Burkholderia glumae* causing rice grain rot is responsible for inducing bacterial wilt in many field crops." 87(8): 890-895.

- Kay, E., et al. (2005). "Three small RNAs jointly ensure secondary metabolism and biocontrol in *Pseudomonas fluorescens* CHA0." 102(47): 17136-17141.
- Keel, C., et al. (1992). "Suppression of root diseases by *Pseudomonas fluorescens* CHA0: importance of the bacterial secondary metabolite 2, 4-diacetylphloroglucinol." 5(1): 4-13.
- Kennedy, R., et al. (2015). "Phenazine-1-carboxamide (PCN) from *Pseudomonas* sp. strain PUP6 selectively induced apoptosis in lung (A549) and breast (MDA MB-231) cancer cells by inhibition of antiapoptotic Bcl-2 family proteins." 20(6): 858-868.
- Kessler, B., et al. (1992). "A general system to integrate lacZ fusions into the chromosomes of gram-negative eubacteria: regulation of the P_m promoter of the TOL plasmid studied with all controlling elements in monocopy." 233(1): 293-301.
- Klinkert, B., et al. (2009). "Microbial thermosensors." 66(16): 2661-2676.
- Koch, B., et al. (2002). "Lipopeptide production in *Pseudomonas* sp. strain DSS73 is regulated by components of sugar beet seed exudate via the Gac two-component regulatory system." 68(9): 4509-4516.
- Koh, S., et al. (2011). "A novel light-dependent selection marker system in plants." 9(3): 348-358.
- Lapouge, K., et al. (2008). "Gac/Rsm signal transduction pathway of γ -proteobacteria: from RNA recognition to regulation of social behaviour." 67(2): 241-253.
- Lapouge, K., et al. (2007). "Mechanism of hcnA mRNA recognition in the Gac/Rsm signal transduction pathway of *Pseudomonas fluorescens*." 66(2): 341-356.
- Laville, J., et al. (1992). "Global control in *Pseudomonas fluorescens* mediating antibiotic synthesis and suppression of black root rot of tobacco." 89(5): 1562-1566.
- Loper, J. E., et al. (2012). "Comparative genomics of plant-associated *Pseudomonas* spp.: insights into diversity and inheritance of traits involved in multitrophic interactions." 8(7): e1002784.
- Loper, J. E., et al. (2016). "Rhizoxin analogs, orfamide A and chitinase production contribute to the toxicity of *Pseudomonas protegens* strain Pf-5 to *Drosophila melanogaster*." 18(10): 3509-3521.
- Loper, J. E., et al. (2008). "Isolation and identification of rhizoxin analogs from *Pseudomonas fluorescens* Pf-5 by using a genomic mining strategy." 74(10): 3085-3093.
- Ma, Z., et al. (2016). "Biosynthesis, chemical structure, and structure-activity relationship of orfamide lipopeptides produced by *Pseudomonas protegens* and related species." 7: 382.
- Maurhofer, M., et al. (1994). "Pyoluteorin production by *Pseudomonas fluorescens* strain CHA0 is involved in the suppression of *Pythium* damping-off of cress but not of cucumber." European Journal of Plant Pathology 100(3): 221-232.
- Maurhofer, M., et al. (1994). "Pyoluteorin production by *Pseudomonas fluorescens* strain CHA0 is involved in the suppression of *Pythium* damping-off of cress but not of cucumber." 100(3): 221-232.
- Maurhofer, M., et al. (1992). "Influence of enhanced antibiotic production in *Pseudomonas fluorescens* strain CHA0 on its disease suppressive capacity."
- Maurhofer, M., et al. (1998). "Salicylic acid biosynthetic genes expressed in *Pseudomonas fluorescens* strain P3 improve the induction of systemic resistance in tobacco against tobacco necrosis virus." 88(7): 678-684.
- Mavrodi, O. V., et al. (2001). "Genetic diversity of pHlD from 2, 4-diacetylphloroglucinol-producing fluorescent *Pseudomonas* spp." 91(1): 35-43.

- Mehnaz, S., et al. (2013). "Lahorenoic acids A–C, ortho-dialkyl-substituted aromatic acids from the biocontrol strain *Pseudomonas aurantiaca* PB-St2." 76(2): 135-141.
- Miethke, M., et al. (2007). "Siderophore-based iron acquisition and pathogen control." 71(3): 413-451.
- Miller, J. H. J. E. i. m. g. (1972). "Assay of β -galactosidase."
- Møller, T., et al. (2002). "Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction." 9(1): 23-30.
- Mondragón, V., et al. (2006). "pH-dependent activation of the BarA-UvrY two-component system in *Escherichia coli*." 188(23): 8303-8306.
- Nandi, M., et al. (2015). "Pyrrolnitrin and hydrogen cyanide production by *Pseudomonas chlororaphis* strain PA23 exhibits nematocidal and repellent activity against *Caenorhabditis elegans*." 10(4): e0123184.
- Nielsen, T., et al. (1999). "Viscosinamide, a new cyclic depsipeptide with surfactant and antifungal properties produced by *Pseudomonas fluorescens* DR54." 87(1): 80-90.
- Nielsen, T., et al. (2000). "Structure, production characteristics and fungal antagonism of tensin—a new antifungal cyclic lipopeptide from *Pseudomonas fluorescens* strain 96.578." 89(6): 992-1001.
- Papenfort, K., et al. (2010). "Regulatory RNA in bacterial pathogens." 8(1): 116-127.
- Péchy-Tarr, M., et al. (2005). "RpoN (σ 54) controls production of antifungal compounds and biocontrol activity in *Pseudomonas fluorescens* CHA0." 18(3): 260-272.
- Pedras, M. S. C., et al. (2003). "Structure, chemistry, and biological activity of pseudophomins A and B, new cyclic lipodepsipeptides isolated from the biocontrol bacterium *Pseudomonas fluorescens*." 62(7): 1105-1114.
- Perneel, M., et al. (2007). "Characterization of CMR5c and CMR12a, novel fluorescent *Pseudomonas* strains from the cocoyam rhizosphere with biocontrol activity." 103(4): 1007-1020.
- Pernestig, A.-K., et al. (2001). "Identification of UvrY as the cognate response regulator for the BarA sensor kinase in *Escherichia coli*." 276(1): 225-231.
- Philmus, B., et al. (2015). "Investigations into the biosynthesis, regulation, and self-resistance of toxoflavin in *Pseudomonas protegens* Pf-5." 16(12): 1782-1790.
- Pierson, L. S., et al. (2010). "Metabolism and function of phenazines in bacteria: impacts on the behavior of bacteria in the environment and biotechnological processes." 86(6): 1659-1670.
- Pliego, C., et al. (2011). "Screening for candidate bacterial biocontrol agents against soilborne fungal plant pathogens." 340(1): 505-520.
- Quail, J. W., et al. (2002). "Pseudophomins A and B, a class of cyclic lipodepsipeptides isolated from a *Pseudomonas* species." 58(5): o268-o271.
- Raaijmakers, J. M., et al. (2006). "Cyclic lipopeptide production by plant-associated *Pseudomonas* spp.: diversity, activity, biosynthesis, and regulation." 19(7): 699-710.
- Raaijmakers, J. M., et al. (2010). "Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics." 34(6): 1037-1062.
- Ramette, A., et al. (2011). "*Pseudomonas protegens* sp. nov., widespread plant-protecting bacteria producing the biocontrol compounds 2, 4-diacetylphloroglucinol and pyoluteorin." 34(3): 180-188.

- Reimann, C., et al. (2005). "Posttranscriptional repression of GacS/GacA-controlled genes by the RNA-binding protein RsmE acting together with RsmA in the biocontrol strain *Pseudomonas fluorescens* CHA0." 187(1): 276-285.
- Rokni-Zadeh, H., et al. (2012). "Genetic and functional characterization of cyclic lipopeptide white-line-inducing principle (WLIP) production by rice rhizosphere isolate *Pseudomonas putida* RW10S2." 78(14): 4826-4834.
- Romeo, T., et al. (1993). "Identification and molecular characterization of *csrA*, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties." 175(15): 4744-4755.
- Roongsawang, N., et al. (2003). "Cloning and characterization of the gene cluster encoding arthrofactin synthetase from *Pseudomonas* sp. MIS38." 10(9): 869-880.
- Roongsawang, N., et al. (2010). "Diversity of nonribosomal peptide synthetases involved in the biosynthesis of lipopeptide biosurfactants." 12(1): 141-172.
- Roth, A. and R. R. J. A. r. o. b. Breaker (2009). "The structural and functional diversity of metabolite-binding riboswitches." 78: 305.
- Schalk, I. J. and L. J. A. a. Guillon (2013). "Fate of ferrisiderophores after import across bacterial outer membranes: different iron release strategies are observed in the cytoplasm or periplasm depending on the siderophore pathways." 44(5): 1267-1277.
- Schubert, M., et al. (2007). "Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA." 14(9): 807-813.
- Shahid, I., et al. (2017). "Secondary metabolites production and plant growth promotion by *Pseudomonas chlororaphis* and *P. aurantiaca* strains isolated from cactus, cotton, and para grass." 27(3): 480-491.
- Shanmugaiah, V., et al. (2010). "Purification, crystal structure and antimicrobial activity of phenazine-1-carboxamide produced by a growth-promoting biocontrol bacterium, *Pseudomonas aeruginosa* MML2212." 108(2): 703-711.
- Sobrero, P. and C. J. C. r. i. m. Valverde (2012). "The bacterial protein Hfq: much more than a mere RNA-binding factor." 38(4): 276-299.
- Sobrero, P. M., et al. (2017). "A matter of hierarchy: activation of orfamide production by the post-transcriptional Gac-Rsm cascade of *Pseudomonas protegens* CHA0 through expression upregulation of the two dedicated transcriptional regulators." *Environ Microbiol Rep* 9(5): 599-611.
- Sobrero, P. M. and C. J. F. i. M. B. Valverde (2020). "Comparative genomics and evolutionary analysis of RNA-binding proteins of the CsrA family in the genus *Pseudomonas*." 7: 127.
- Song, C., et al. (2016). "Living on the edge: emergence of spontaneous *gac* mutations in *Pseudomonas protegens* during swarming motility." *Environ Microbiol* 18(10): 3453-3465.
- Song, C., et al. (2015). "The R sm regulon of plant growth-promoting *Pseudomonas fluorescens* SS 101: role of small RNA s in regulation of lipopeptide biosynthesis." 8(2): 296-310.
- Sørensen, D., et al. (2002). "Cyclic lipoundecapeptide lokisin from *Pseudomonas* sp. strain DSS41." 43(25): 4421-4423.
- Spence, C., et al. (2014). "Natural rice rhizospheric microbes suppress rice blast infections." 14(1): 1-17.
- Strieker, M., et al. (2010). "Nonribosomal peptide synthetases: structures and dynamics." 20(2): 234-240.

- Stutz, E., et al. (1986). "Naturally occurring fluorescent *Pseudomonads* involved in suppression." 76(2): 181-185.
- Sulochana, M., et al. (2014). "Siderophore as a potential plant growth-promoting agent produced by *Pseudomonas aeruginosa* JAS-25." 174(1): 297-308.
- Sun, X., et al. (2002). "Predicted structure and phyletic distribution of the RNA-binding protein Hfq." 30(17): 3662-3671.
- Takeuchi, K., et al. (2012). "ppGpp controlled by the Gac/Rsm regulatory pathway sustains biocontrol activity in *Pseudomonas fluorescens* CHA0." 25(11): 1440-1449.
- Tripathi, R. K. and D. J. J. o. b. Gottlieb (1969). "Mechanism of action of the antifungal antibiotic pyrrolnitrin." 100(1): 310-318.
- Valverde, C., et al. (2003). "RsmY, a small regulatory RNA, is required in concert with RsmZ for GacA-dependent expression of biocontrol traits in *Pseudomonas fluorescens* CHA0." 50(4): 1361-1379.
- Valverde, C. J. A. o. m. (2009). "Artificial sRNAs activating the Gac/Rsm signal transduction pathway in *Pseudomonas fluorescens*." 191(4): 349-359.
- Ventre, I., et al. (2006). "Multiple sensors control reciprocal expression of *Pseudomonas aeruginosa* regulatory RNA and virulence genes." 103(1): 171-176.
- Verhagen, B. W., et al. (2010). "*Pseudomonas* spp.-induced systemic resistance to *Botrytis cinerea* is associated with induction and priming of defence responses in grapevine." 61(1): 249-260.
- Vogel, J. and B. F. Luisi (2011). "Hfq and its constellation of RNA." Nat Rev Microbiol 9(8): 578-589.
- Wagner, E. G. H., et al. (2002). "Antisense RNAs in bacteria and their genetic elements." 46: 361-398.
- Weinberg, Z., et al. (2007). "Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline." 35(14): 4809-4819.
- Welbaum, G. E., et al. (2004). "Managing soil microorganisms to improve productivity of agroecosystems." 23(2): 175-193.
- Zuber, S., et al. (2003). "GacS sensor domains pertinent to the regulation of exoproduct formation and to the biocontrol potential of *Pseudomonas fluorescens* CHA0." 16(7): 634-644.

Capítulo 2

Papel de las orfamidas en la persistencia en el suelo y la competitividad para la colonización radicular en *Pseudomonas protegens* CHA0.

Lic. Andrés Muzlera

Laboratorio de Fisiología y Genética de
Bacterias Beneficiosas para Plantas 54

2.1 Resumen

Se ha visto que la producción de los lipopéptido orfamidas está relacionada con la motilidad en *Pseudomonas protegens* CHA0, ya que los mutantes impedidos en su biosíntesis, no son capaces de realizar *swarming in vitro*. Teniendo en cuenta esto y con el objetivo de profundizar la caracterización del rol de las orfamidas en el nicho biológico de esta bacteria, se realizó un ensayo para explorar su importancia tanto en su *fitness* en el suelo como en la colonización radicular.

En este capítulo, se describe un ensayo de competencia entre las cepas silvestre (CHA0; *wt*) y deficiente en la producción de orfamidas (Δofa). Se evaluó su capacidad para sobrevivir en suelo libre de plantas, y su capacidad de colonización radicular.

2.2 Introducción

2.2.1 Rol de los lipopéptidos en la naturaleza

Los lipopéptidos producidos por especies de *Pseudomonas* exhiben actividades líticas e inhibitoras del crecimiento contra una amplia gama de microorganismos (Haas and Défago 2005). En las cepas de *Pseudomonas* asociadas a plantas, los lipopéptidos también juegan un papel en la colonización de semillas (Nielsen et al. 2005) y raíces (Tran et al. 2007), en defensa contra microorganismos competidores y protozoos depredadores (Mazzola et al. 2009), y en la motilidad por *swarming* y la formación de biopelículas (*biofilms*) (Raaijmakers et al. 2010).

2.2.1.1 Lipopéptidos involucrados en la motilidad

La motilidad de las bacterias ha sido ampliamente estudiada, tanto el *swimming*, el *swarming* (Henrichsen 1972) y el *twitching* (Bradley et al. 1972). Durante el *swarming*, las células vegetativas de bacterias pueden diferenciarse en células híper-flageladas que generalmente son más alargadas (Harshey 2003). Los lipopéptidos tienen un rol fundamental en este tipo de motilidad, ya que mutantes deficientes en su producción no son capaces de realizar *swarming in vitro* en placas de agar semisólidas (Song et al. 2016). Además, la motilidad de estos mutantes se ve restaurada cuando se suplementan lipopéptidos purificados al medio (Andersen et al. 2003; De Bruijn et al. 2007), sin embargo, no es así en todos los casos; En *P. putida* IsoF, la putisolvina no funciona como un “bien público” (Carcamo-Oyarce et al 2015) y el *swarming* no se reestablece con la complementación del lipopéptido. Adicionalmente, el aporte de los lipopéptidos a la motilidad no ocurre necesariamente en el mismo sentido, las mutaciones en los genes que codifican para la biosíntesis del lipopéptido sesillina en *Pseudomonas sp.* CMR12a provocaron un aumento en la motilidad (D'aes et al. 2014). Hasta el momento, no se sabe con certeza si los lipopéptidos también contribuyen a la dispersión en hábitats naturales, pero existen reportes de su relevancia en la colonización radicular. La cepa *P. fluorescens* SS101 por ejemplo, demostró una colonización radicular más efectiva en semillas de tomate que mutantes

deficientes en la producción del lipopéptido massetolido A (Tran et al. 2007). De manera similar, el mutante deficiente en viscosina de la cepa de *P. fluorescens* 5064 patógena para plantas, no pudo colonizar la superficie de los floretes de brócoli en la misma medida que la cepa salvaje (Hildebrand et al. 1998). La anfisina producida por *Pseudomonas* sp. DSS73 también demostró ser esencial para la colonización de semillas de remolacha (Nielsen et al. 2005).

2.2.1.2 Lipopéptidos involucrados en la formación de biopelículas

En *Pseudomonas*, los lipopéptidos desempeñan un papel importante en la adhesión a superficies y la formación de biopelículas, aunque con diferentes resultados según el tipo de lipopéptido (Raaijmakers et al. 2010). Al igual que con la motilidad, en la mayoría de los casos, la formación de biopelículas se redujo en mutantes deficientes en la producción de lipopéptidos, pero en algunas cepas, la formación de biopelículas aumentó o no cambió en comparación con la del tipo salvaje. Por ejemplo, la cepa *Pseudomonas* MIS38 productora del lipopéptido antrofactina es capaz de formar biopelículas estables, mientras que mutantes no productores forman biopelículas en mayor cantidad pero esta es inestable (Roongsawang et al. 2003). Resultados similares fueron reportados para *Pseudomonas* productoras de putisolvina y cichofactina (Kuiper et al. 2004; Kruijt, Tran, and Raaijmakers 2009; Pauwelyn et al. 2013). Por otro lado, los mutantes deficientes en los lipopéptidos massetolido, viscosina, sesilina, WLIP (*white line-inducing principle*) y xantolisina de las cepas *P. fluorescens* SS101 y SBW25, *Pseudomonas* CMR12a, *P. putida* RW10S2, *P. putida* BW11M1, formaron significativamente menos biopelícula (De Bruijn et al. 2007; D'aes et al. 2014; Rokni-Zadeh et al. 2013; Li et al. 2013). En función de la evidencia previamente mencionada, en este capítulo se plantea estudiar el aporte de las orfamidas a la sobrevivencia de *P. protegens* CHA0 en suelo libre y en rizosfera.

2.3 Hipótesis de trabajo

Las orfamidas cumplen un papel fundamental en la motilidad de *Pseudomonas protegens* CHA0, por lo tanto, su presencia es relevante para la sobrevivencia de esta bacteria tanto en suelo libre como en rizosfera.

2.4 Objetivos específicos

- Evaluar el papel de las orfamidas en la sobrevivencia de *Pseudomonas protegens* CHA0 en suelo libre y rizosférico por medio de ensayos de competencia con mutantes deficientes en la producción de orfamidas.

2.5 Materiales y métodos

2.5.1 Cepas bacterianas y plásmidos

Las cepas bacterianas y plásmidos empleados en este capítulo se listan en la Tabla 1. *E. coli* se creció a 37°C y 200 rpm en caldo NYB (caldo nutritivo, 25 g/L; extracto de levadura, 5 g/L). Todas las cepas de *pseudomonas* se crecieron en NYB a 28°C y 200 rpm de forma rutinaria excepto que se detalle lo contrario. Las cepas se conservaron, en condiciones de esterilidad, en una solución acuosa de glicerol 20% final en ultrafreezer (-80°C). En los casos requeridos, los antibióticos fueron suministrados en las siguientes concentraciones: para *E. coli*, ampicilina (Ap) 100 µg/ml; kanamicina (Km) 50 µg/ml; gentamicina (Gm) 10 µg/ml; cloranfenicol 10 µg/ml; tetraciclina (Tc) 50 µg/ml. Para *pseudomonas*, kanamicina (Km) 50 µg/ml; gentamicina (Gm) 40 µg/ml; tetraciclina (Tc) 125 µg/ml.

Los cultivos de *pseudomonas* crecidos en placas multiwells destinados a cinéticas de crecimiento con medidas de fluorescencia fueron crecidos en GCM (Casaminoácidos 5 g/L, K₂HPO₄ 0,75 g/L, MgSO₄·7H₂O 0,75 g/L, glicerol 20 g/L).

Los cultivos de *pseudomonas* destinados a la selección de transconjugantes se crecieron en M9-citrato (MgSO₄ 2mM; CaCl₂ 0,1 mM; citrato de sodio 0,4 %; Na₂HPO₄ 7 g/L, KH₂PO₄ 3 g/L; NaCl 0.5g/L; NH₄Cl 1 g/L; 15 g/L de agar bacteriológico) (Green and Sambrook 2012)

Tabla 4. Cepas bacterianas, plásmidos y oligonucleótidos utilizados en este capítulo

Microorganismo	Genotipo característica o secuencia	Referencia
<i>Pseudomonas protegens</i>		
CHA0	Cepa Silvestre	(Stutz, Défago, and Kern 1986)
CHA19	$\Delta gacS$	(Zuber et al. 2003)
CHA89	$gacA::Km^R$	(Blumer et al. 1999)
CHA1009	$rsmA::\Omega Km$ $rsmE::\Omega Hg$;	(Reimann et al. 2005)
CHA1144	$\Delta rsmX$ $\Delta rsmY$ $\Delta rsmZ$	(Kay, Dubuis, and Haas 2005)
ARQ1	CHA0 silvestre con inserción Minitn7-gfp Km ^R en el sitio attTn7	(Jousset, Lara, Wall, Valverde, et al. 2006)
CHA5101	CHA0 silvestre con delección $\Delta ofaABC$. No produce lipopéptidos	(Flury et al. 2017)
CHA5101-m	CHA5101 con inserción Minitn7-gfp Gm ^R en el sitio attTn7	Este trabajo
<i>Escherichia coli</i>		
DH5 α	$recA1$ $endA1$ $hsdR17$ $deoR$ $thi-1$ $supE44$ $gyrA96$ $relA1\Delta(lacZYA-argF)$ U169 ($\Phi 80d lacZ\Delta M15$)	Stock de laboratorio
SM10/ λ pir	$thi-$ thr leu $tonA$ $lacY$ $supE$ $recA::RP4-2-Tc::Mu$, Km ^R , λ pir	(Miller and Mekalanos 1988)

HB101	Cepa <i>helper</i> ; F ⁻ λ – hsdS20(rB ⁻ mB ⁻) recA13 leuB6(Am) araC14 Δ (gpt-proA)62 lacY1 galK2(Oc) xyl-5 mtl-1 thiE1 rpsL20(Sm ^R) glnX44(AS)	(Boyer and Roulland-Dussoix 1969)
MT102	<i>leu pro thi</i> Sm ^R	(Casabadan and Cohen 1980)
Plásmidos		
pRK600	Vector <i>helper</i> para conjugación; ori(ColE1), RK2(mob+ tra+); Cm ^R	(Kessler et al. 1992)
pUX-BF13	Plásmido codificante para la transposasa Tn7	(Bao et al. 1991)
AKN104	miniTn7 PA1/04/03- <i>gfpAAV-a</i> Gm ^R	(Lambertsen, Sternberg, and Molin 2004)
pME7408	<i>prsmZ-gfp</i> ; <i>orfDR'</i> - <i>mCherry</i> ; Tc ^R , Km ^R	Este trabajo
pME6000	vector de clonado, pBBR1MCS derivative; Tc ^R	(Maurhofer et al. 1998)
pME6102	<i>orfUR'</i> - <i>gfp</i> ; <i>orfDR'</i> - <i>DsRed</i> . Tc ^R , Km ^R	Este trabajo
pME7402	<i>prsmZ-gfp</i> . Tc ^R	(Dubuis et al. 2006)
pBAMD1-6	Vector vehículo de mini-Tn5; ori(R6K); Ap ^R Gm ^R	(Martínez-García et al. 2014)
pR_mCherry	Vector para generar fusiones reporteras con <i>mCherry</i> (modificado del pR_EGFP)	(Torres-Quesada et al. 2013)
Oligonucleótidos		
21850FHindIII	TCAAAGCTTGCCACGGCTG	Este trabajo
21850RNheI	ACCGCTAGCGTTCGCAATACTGC	Este trabajo
Semillas		
Trigo	Variedad <i>Baguette 550</i>	

2.5.2 Generación de variantes marcadas

Se decidió evaluar el *fitness* de *P. protegens* CHA0 wt y Δ *ofa* en suelo libre no esterilizado, mediante el recuento de unidades formadoras de colonias (UFC) en medio selectivo. Para poder recuperar ambas cepas del suelo por separado fue necesario generar variantes marcadas de cada una de ellas con distintas resistencias a antibióticos. La variante de la cepa Δ *ofa* marcada fue generada para este trabajo utilizando el sistema mini-Tn7 descrito por (Koch, Jensen, and Nybroe 2001). Se realizó una conjugación tetraparental con la cepa *P. protegens* CHA5101 (Δ *ofa*), *E. coli* SM10 λ pir/pUX-BF13 (para transposición Tn7), *E. coli* HB101/pRK600 (para movilización de plásmidos mob+/tra-) y *E. coli* MT102/AKN104, poseedora de la construcción miniTn7(Gm)PA1/04/03-*gfpAAV-a* (Lambertsen, Sternberg, and Molin 2004). Se generaron cultivos líquidos saturados de estas cepas a 37 °C y de *P. protegens* CHA5101 (cepa receptora) a 35 °C para favorecer la recepción del ADN foráneo (Heeb, Blumer, and Haas 2002a). Todos los cultivos se incubaron bajo una agitación de 200 rpm. Se centrifugaron 1,5 ml de cada cultivo, se lavaron los *pellets* y se resuspendieron en 300 μ l de medio de cultivo NYB. Luego, se combinaron todos los *pellets* (volumen final de 1,2 ml) para centrifugarlos y finalmente resuspenderlos en 50 μ l de NYB. Esta suspensión que contiene todas las cepas implicadas en la conjugación tetraparental se sembró en el borde de una placa de Petri con medio de cultivo NA, y se incubó durante 4 hs a 37 °C. A continuación, se resuspendieron las células sembradas en la placa con 1 ml de NYB y se sembraron diferentes diluciones de esta suspensión (directa, 1/2 y

1/10) en placas de Petri con medio de cultivo mínimo M9-citrato suplementado con gentamicina 40 µg/ml, permitiendo únicamente el desarrollo de las pseudomonas que incorporaron el *cassette* Tn7 en su genoma (CHA5101-m Gm^R). Para la variante silvestre, nuestro laboratorio cuenta con la cepa ARQ1 (Jousset, Lara, Wall, and Valverde 2006), la cual fue generada con el mismo sistema mini-Tn7 a partir de la cepa receptora *P. protegens* CHA0, incorporando resistencia al antibiótico kanamicina y un gen de expresión constitutiva codificante para *gfp*.

2.5.3 Generación de vector pME7408 (*prsmZ-gfp; orfDR'-mCherry*)

Para la construcción del vector pME7408, se amplificó un fragmento de 569 pb con los oligonucleótidos 21850FHindII/21850RNheI conteniendo la región promotora del gen más 7 codones del ORF *orfDR* y se clonó en el vector pR_mCherry en los sitios *HindIII* y *NheI* generando una fusión traduccional en marco de lectura con *mCherry*. Se movilizó por conjugación triparental a *P. protegens* CHA0 y se evaluó la expresión de mCherry. Luego se subclonó la construcción al vector pBAMD1-6 en los sitios *HindIII* y *EcoRI*, los cuales se encuentran flaqueados por dos sitios *NotI*. Este subclonado se realizó para escindir la construcción *orfDR-mCherry* por digestión con *NotI*. Por último, se subclonó el fragmento al vector pME7402 en el sitio *NotI*, generando el vector pME7408 (*prsmZ-gfp; orfDR'-mCherry*; Tc¹²⁵, Km⁵⁰)

2.5.4 Persistencia en suelo natural

Para evidenciar si la producción de lipopéptido impacta en el *fitness* de CHA0, se decidió estudiar la supervivencia en suelo natural de la cepa silvestre (wt) y de una mutante isogénica deficiente en la producción de orfamidas (Δofa). Se llevó a cabo un ensayo en invernadero dividido en dos etapas. La primera etapa consistió en estudiar la contribución de las orfamidas a la supervivencia de ambas cepas en suelo libre de plantas. La segunda, en cambio, fue para evaluar el rol de las orfamidas para la competitividad en la colonización de trigo (Figura 30).

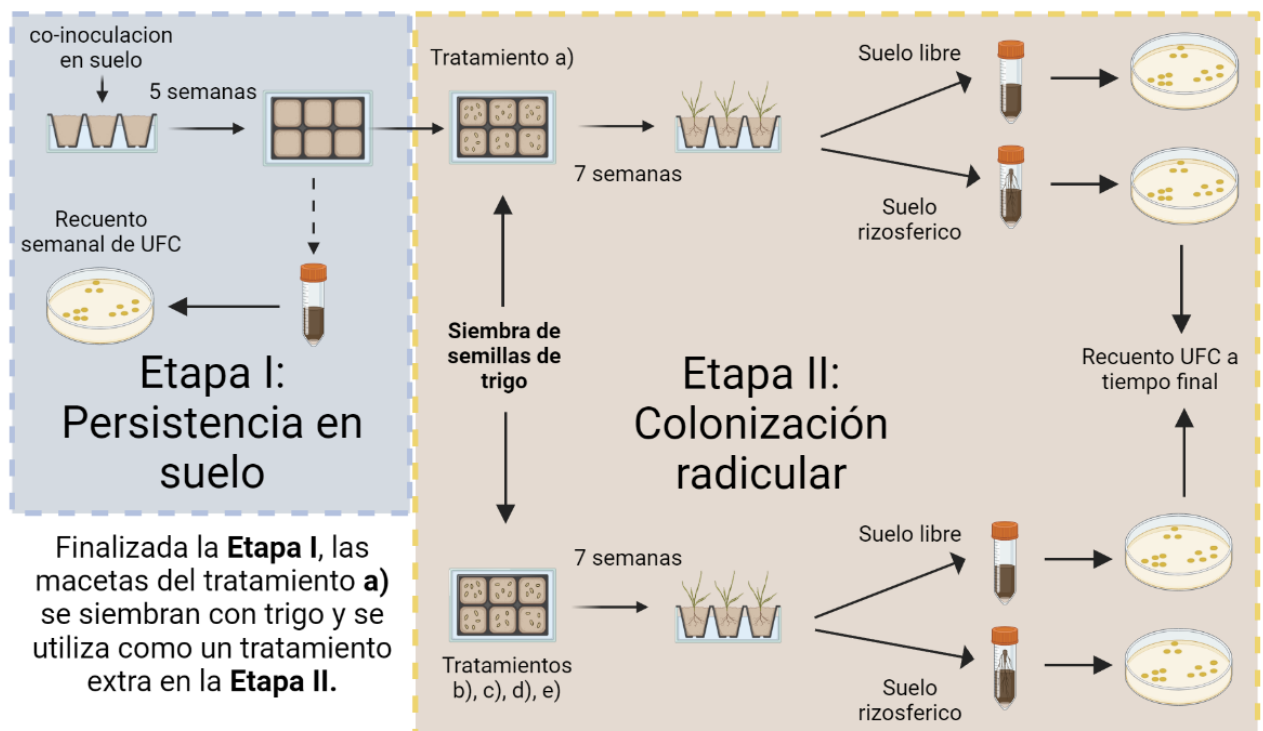


Figura 30. Esquema del ensayo realizado en invernadero

El experimento completo contó con los siguientes tratamientos:

- I. Primera etapa:
 - a) Co-inoculación wt y Δofa en suelo, y monitoreo de la carga bacteriana durante 5 semanas.
- II. Segunda etapa:
 - a) Co-inoculación wt y Δofa en suelo 5 semanas antes de la siembra de trigo. *
 - b) Co-inoculación wt y Δofa en suelo, inmediatamente antes de la siembra de trigo.
 - c) Co-inoculación wt y Δofa en semillas de trigo.
 - d) Inoculación en suelo con wt, y en semillas de trigo con mutante Δofa .
 - e) Inoculación en suelo con mutante Δofa , y en semilla de trigo con wt.

* Una vez finalizada la primera etapa, las macetas del tratamiento a) se utilizaron como un nuevo tratamiento en la segunda etapa del ensayo.

En todos los tratamientos se utilizaron macetas de plástico soplado de 3l recortadas hasta una altura de 8cm conteniendo 400g de sustrato compuesto por una mezcla de horizonte A de suelo natural y perlita en proporción 4:1 v/v. El suelo fue tamizado a través de una malla de 2mm de poro, para obtener un sustrato homogéneo y libre de restos vegetales. Este último proviene de un lote ubicado en la localidad de Gorina (34°54'39.1"S 58°01'56.6"O; partido de La Plata). El riego fue por aspersión superficial con agua destilada esterilizada, manteniendo la humedad a capacidad de campo por pesada. En el invernáculo, se mantuvo una temperatura de incubación

diurna máxima de 24 ± 2 °C y una humedad de entre 60 % y 80 %, con ciclos de luz:oscuridad de 16:8 horas (complementado con iluminación artificial para compensar las variaciones estacionales del fotoperiodo natural).

2.5.5 Inoculación

Se crecieron las cepas ARQ1 (wt) y CHA5101-m (Δofa) en 10 ml de NYB líquido conteniendo 40 µg/ml de gentamicina y 50 µg/ml de kanamicina, respectivamente, a 28 °C y 200 rpm hasta saturación (*overnight*). A partir de estos cultivos se prepararon suspensiones en solución salina que se emplearon para:

- Inocular sustrato (Etapas I y II). En una bandeja conteniendo una capa delgada de sustrato, se agregó agua estéril hasta obtener una humedad cercana a la saturación. En ese momento, se comenzó a rociar con suspensión de células mientras se mezcló constantemente de manera manual, hasta obtener una carga de 0.004 unidades de DO_{600} , lo que equivale aproximadamente a 1×10^6 UFC/g de suelo.
- Inocular semillas previamente esterilizadas superficialmente (Etapa II). Las semillas fueron lavadas con etanol 70% v/v por 1 minuto y se enjuagaron con agua destilada estéril. Luego se trataron con lavandina comercial diluida a 1.2% w/v final de cloro activo durante 10 min con agitación orbital suave. Finalmente se realizaron seis lavados con agua destilada estéril durante 10 min cada vez. Las semillas desinfectadas se guardaron o.n. a 4°C. Luego fueron incubadas con una suspensión de $DO_{600}=0.04$ durante 15 min en agitación orbital suave, previamente a su siembra.

En los tratamientos que correspondió de la Etapa II, luego de cargar cada maceta con aprox. 400g de sustrato, se colocaron 8 semillas por maceta y se cubrieron ligeramente con sustrato húmedo, permitiendo luego ralear las plantas sobrantes, para dejar 5 plantas por maceta.

2.5.6 Recolección de muestras

Para el caso del tratamiento a) de la Etapa I, se extrajo 1g de suelo subsuperficial con una espátula esterilizada, con una frecuencia semanal, hasta las 5 semanas luego de inocular. Al cabo de las 5 semanas de iniciado el tratamiento a), se sembraron semillas de trigo desinfectadas superficialmente en el centro de cada maceta, y se continuó la incubación y riego como tratamiento a) de la Etapa II junto al resto de los tratamientos b-e) (Figura 30).

La cosecha se llevó a cabo 7 semanas después de sembrar, para todos los tratamientos de la Etapa II (a-e), momento en el cual se tomaron muestras de “suelo libre de plantas” y “suelo rizosférico”. Se diferenció entre la fracción de suelo más lejana a la raíz y menos afectada por sus exudados (“suelo libre de plantas”) y la más cercana y firmemente adherida a la raíz (“suelo rizosférico”). Primero se extrajo 1g de suelo subsuperficial libre de raíces, con una espátula esterilizada, y se lo almacenó en un tubo estéril a 4°C hasta su procesamiento. Luego se descalzaron las macetas y se agitaron las plantas sobre una bandeja para desprender el suelo débilmente adherido a las raíces. Finalmente se cortaron trozos de raíces con suelo rizosférico adherido y se almacenaron en tubos estériles a 4°C hasta su procesamiento.

2.5.7 Recuentos en placa

Para poder realizar los recuentos de UFC en placa fue necesario recurrir a un medio selectivo para *Pseudomonas*, ya que se encontraron microorganismos naturales del suelo capaces de crecer en agar nutritivo suplementado con ambos antibióticos (información no mostrada). Se decidió utilizar el medio Gould's S1 (Gould et al. 1985) para este fin. Si bien no contamos con una versión comercial del medio selectivo S1, la formulación generada por nuestro laboratorio demostró ser específica para el género *Pseudomonas* en múltiples aislamientos de muestras de suelo o rizosfera de cultivos (Agaras, Wall, and Valverde 2012). Todos los recuentos se realizaron en placas de medio S1 con 40 µg/ml de gentamicina para la variante mutante Δofa (CHA5101-m) y 50 µg/ml de kanamicina para la variante silvestre (ARQ1). El protocolo llevado a cabo fue ligeramente distinto según se tratara de muestras de suelo libre o rizosférico:

Muestras de suelo libre: se pesó 1 gr de suelo libre y se lo suspendió en 9 ml de solución salina (NaCl 0.9%). Esta primera suspensión fue agitada en vortex a máxima potencia durante 1 min para romper los bloques de tierra restantes (procurando la fácil formación del vórtice en el proceso). Seguidamente, se sumergió en baño de sonicación durante 1 min (para esto se utilizó un equipo Testlab TB04 (Argentina) de 40 Khz y 160 W) y finalmente se centrifugó durante 10 min a 50×g. A partir del sobrenadante se obtuvieron diluciones seriadas, de las cuales se sembraron 20 µl por triplicado en placas de Petri (Naghili et al. 2013) con medio S1 suplementado con los antibióticos gentamicina para la cepa CHA5101-m o kanamicina para la cepa ARQ1. Para realizar el conteo, las placas fueron incubadas a 28 °C durante 48 h.

Muestras de suelo rizosférico con raíces: se pesaron entre 3-7g de raíces con suelo rizosférico, se agregaron 20 ml de solución salina, y se trataron igual que las muestras de suelo libre. Una vez finalizado el tratamiento, se retiraron las raíces lavadas sin suelo, se restó su peso y se calcularon los g de suelo rizosférico obtenidos por muestra, luego se ajustó la cantidad de UFC a g de suelo.

2.5.8 Microscopia de fluorescencia

Para la visualización de las cepas marcadas se utilizó un microscopio de fluorescencia Leica DMI6000B y el registro se realizó con una cámara Leica DFC295. Para ver la expresión de *gfp* se utilizó un filtro de excitación BP 470/40 y un filtro de emisión BP 525/50. Para ver la expresión de DsRed o mCherry se utilizó un filtro de excitación BP 515-560 y un filtro de emisión LP 590.

2.5.9 Análisis estadístico

Cada uno de los tratamientos en macetas fue llevado a cabo por quintuplicado (n = 5), considerando cada maceta como una réplica biológica. En cada una de ellas se desarrollaron 5 plantas de trigo, a partir de las cuales se realizó un promedio para el valor de cada réplica en todos los parámetros medidos.

El análisis estadístico de los datos se realizó con el software GraphPad Prism versión 8.0.2 para Windows, GraphPad Software, San Diego, California USA, www.graphpad.com

2.6 Resultados y discusión

2.6.1 *Fitness* de *P. protegens* silvestre (CHA0) en suelo natural.

Se realizó un ensayo de supervivencia con las cepas CHA0 y CHA5101 ($\Delta ofaABC$) en diferentes condiciones (Figura 30), con el objetivo de visualizar: I) el impacto de las orfamidas en la persistencia/supervivencia de la bacteria en suelo natural; II) el impacto de las orfamidas en la competitividad por la colonización de raíces. Complementariamente, se realizaron observaciones al microscopio de fluorescencia con distintas cepas reporteras (Tabla 1), para evidenciar de forma indirecta la expresión del *cluster* biosintético de orfamidas cuando la cepa silvestre CHA0 crece en un contexto rizosférico.

2.6.2 Primera etapa: Persistencia en suelo libre

El objetivo de la primera etapa de este ensayo fue el de evaluar la importancia de la producción de orfamidas para la supervivencia de CHA0 en suelo natural, libre de raíces. Para eso se co-inoculó la cepa silvestre y la mutante $\Delta ofaABC$ en igual proporción en suelo natural, se tomaron muestras de suelo semanalmente y se realizaron recuentos de UFC de cada cepa (Figura 31).

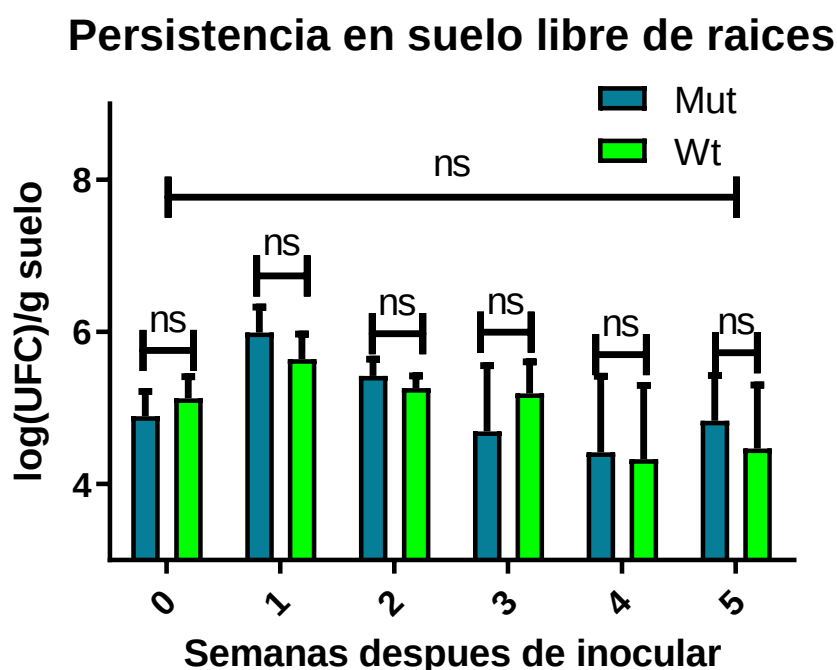


Figura 31. Recuento de UFC de CHA0 wt (verde) CHA5101 (azul). Los valores absolutos de UFC fueron relativizados por g de suelo.

La progresión de UFC en el tiempo (Figura 31) es coherente con lo observado por Mascher (Mascher et al. 2002), donde los recuentos de viables de CHA0 se mantuvieron relativamente constantes durante 4 semanas en suelo natural libre de plantas. En nuestro ensayo observamos una leve tendencia a la disminución en la abundancia de ambas cepas en función del tiempo, pero no resultó estadísticamente significativa (Figura 31). Debido a que el experimento fue monitoreado por medio de recuento de viables, no es posible saber si una porción del total de las células no pudo ser recuperada ya que se encontraban en un estado viable pero no cultivable. Esto podría explicar la ligera disminución que se observa a lo largo de 5 semanas en los recuentos de UFC. Independientemente de la aparición de variantes viables pero no cultivables, podemos afirmar que la ausencia de producción de orfamidas no parece afectar la abundancia de especies cultivables en las condiciones de este ensayo. Ambas cepas presentaron valores similares de UFC a lo largo de 5 semanas en suelo libre (Figura 31).

Uno de los posibles roles relevantes de las orfamidas en el *fitness* de CHA0 en suelo consiste en la necesidad de su expresión para que la bacteria pueda realizar *swarming*. La capacidad de desplazarse representa una ventaja sobre cepas que no puedan hacerlo, ya que podrán ocupar más y mejores nichos, obteniendo así más recursos del suelo. Sin embargo, hay que tener en cuenta que existen reportes de cepas deficientes en orfamidas que pueden realizar *swarming* en placa, siempre y cuando se encuentren en contacto con cepas silvestres (Song et al. 2016). Por lo tanto, si el impacto de las orfamidas sobre el *fitness* de la bacteria, en un estadio de vida libre, radica solo en la posibilidad de realizar *swarming*, es posible pensar que la cepa mutante podría estar “aprovechándose” de la cepa silvestre, y trasladarse junto con ella. En este escenario sería necesario analizar la supervivencia de cada cepa por separado, y no en una condición de competencia. Sin embargo, nuestra hipótesis hasta el momento es que en las condiciones de vida libre de la cepa silvestre CHA0 en nuestro ensayo, no existe una presión de selección que confiera alguna ventaja sobre las cepas capaces de sintetizar orfamidas por sobre las deficientes en el lipopéptido cíclico. El sustento de esta idea radica en dos observaciones realizadas en otras *Pseudomonas*: 1) El efecto de apropiación de la capacidad de *swarming* no afecta al 100% de las variantes mutantes, y si esto le otorgase una ventaja adaptativa, solo una pequeña porción de las cepas mutantes se vería beneficiada, por lo que el impacto en la supervivencia sería sutil. Además, dichos ensayos fueron realizados en placas de Petri, donde las concentraciones de células silvestres y mutantes son varios órdenes de magnitud superiores a las concentraciones que nosotros encontramos en suelo libre. 2) En 2005 se reportó un experimento de supervivencia en suelo libre con la cepa *Pseudomonas* sp. DSS73, la cual produce amfisina, un lipopéptido cíclico muy similar a la orfamida, tanto en estructura como en función y también necesario para el *swarming*. Los mutantes Δ amsY (que no producen el lipopéptido) no mostraron diferencias significativas en el número de recuentos en comparación con la cepa silvestre, a lo largo de 8 días en suelo natural, aunque evidenciaron menor competitividad para la colonización de semillas de remolacha desde el suelo (Nielsen et al. 2005).

2.6.3 Segunda etapa: Competencia por la colonización radicular

En la segunda etapa del ensayo, se evaluaron los recuentos de ambas cepas en diferentes escenarios de ubicación inicial de las células inoculadas (Figura 32). El primer tratamiento de co-inoculación en suelo corresponde a las macetas utilizadas en la primera etapa (Figura 30). Luego de 5 semanas de inoculación del suelo con una mezcla balanceada de las cepas wt y Δ ofa, se sembraron las semillas de trigo (tratamiento a; Figura 32) y se continuó tratándolas como al resto de los tratamientos (b-e; Figura 32).

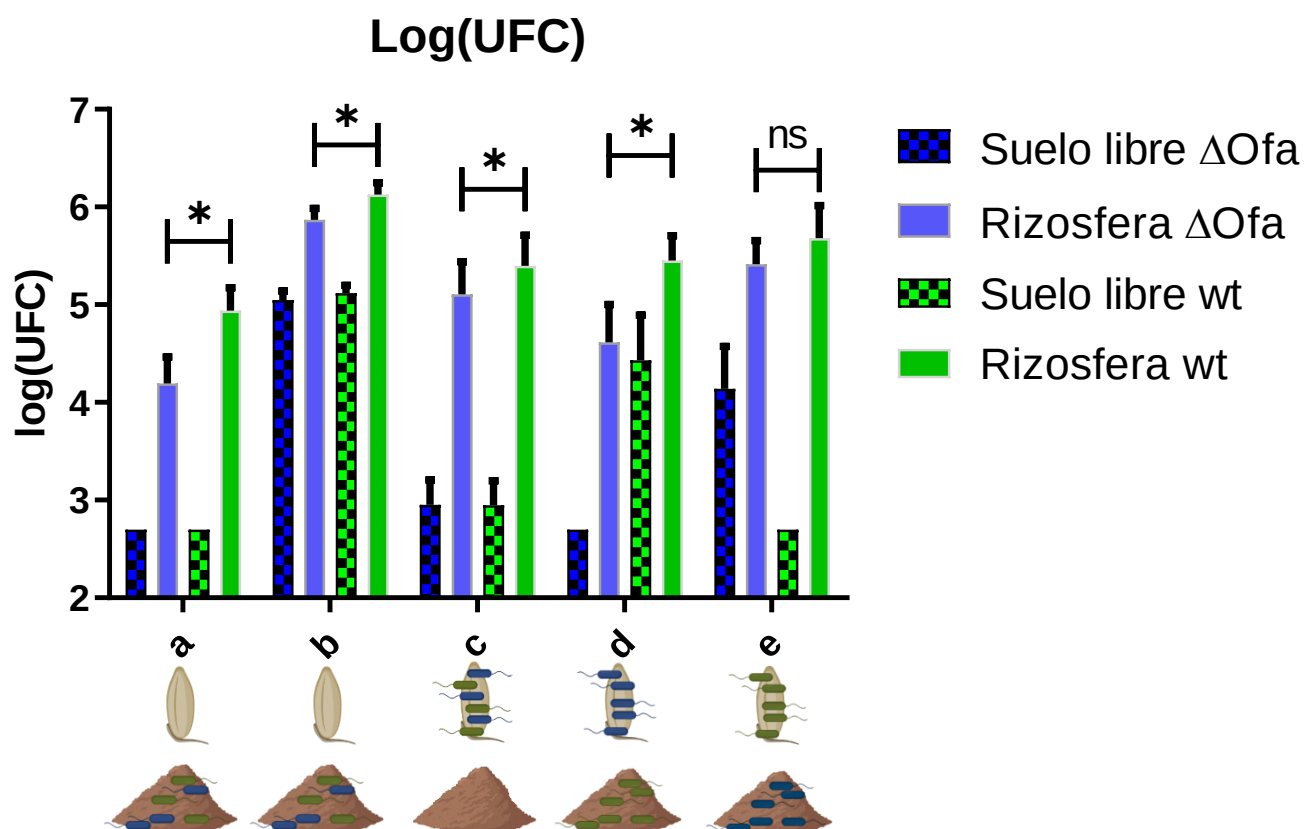


Figura 32. Recuentos de UFC/g suelo rizosférico en los distintos tratamientos. a) Co-inoculación en suelo proveniente de la primera etapa. b) Co-inoculación en suelo. c) Co-inoculación en semilla. d) Inoculación en suelo con cepa silvestre, y en semilla con mutante. e) Inoculación en suelo con cepa mutante, y en semilla con silvestre. Las barras azules corresponden a recuentos de cepas mutantes. Las barras verdes corresponden a recuentos de cepas silvestres. El patrón de relleno corresponde a suelo libre de raíces. El patrón liso corresponde a suelo rizosférico.

En todos los casos y en forma independiente de la localización inicial de las células inoculadas (suelo o semilla), se observó un mayor recuento de UFC de la cepa silvestre por sobre la mutante, aunque este efecto no resultó estadísticamente significativo para el tratamiento e) (cepa

silvestre inoculada en semilla, cepa deficiente en orfamidas inoculada en el suelo). Para comprender mejor la diferencia entre los recuentos de los tratamientos, es importante analizar el efecto del tipo de inoculación sobre los recuentos de UFC. La Figura 33 representa los recuentos de UFC, sin discriminar silvestre de mutante, solo diferenciando los dos tipos de inoculaciones utilizados: en suelo y en semilla. Al comparar los recuentos derivados de estas dos localizaciones iniciales de los inóculos, se observa que la inoculación en suelo resulta en un número significativamente mayor de UFC recuperadas de suelo rizosférico. De aquí en más, hay que tener en cuenta que la cepa que fue inoculada en suelo, contará con una “ventaja” de acceso a la rizosfera desde el suelo sobre las inoculadas en semilla, al comparar las UFC recuperadas de la colonización rizosférica.

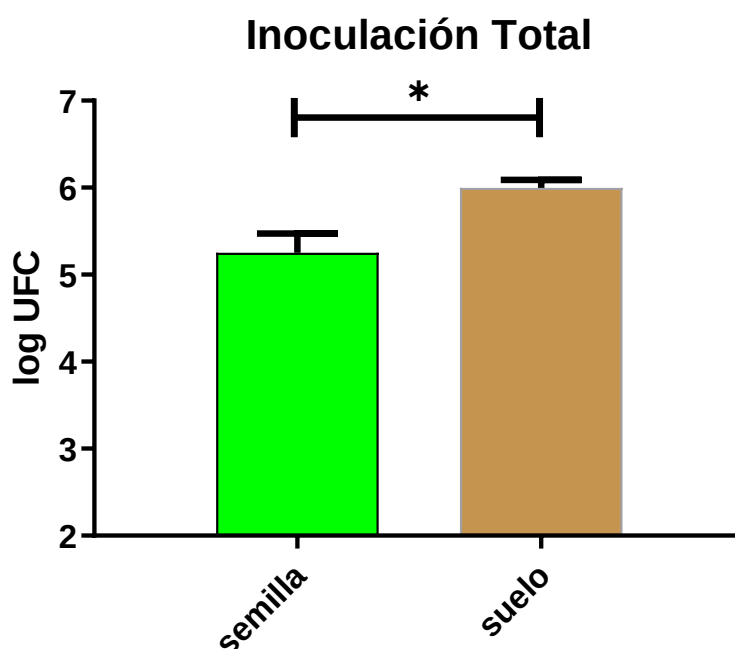


Figura 33. Recuentos de ambas cepas inoculadas en semilla (verde), recuentos totales de ambas cepas inoculadas en suelo (naranja).

Teniendo en cuenta esto, para analizar el impacto de las orfamidas en la colonización, podemos ver que en los tratamientos a), b) y c) los recuentos de la cepa silvestre recuperada de la rizosfera de trigo fueron significativamente superiores a los de la cepa mutante, a pesar que en estos tres tratamientos ambas cepas compartieron la misma ubicación de partida en el sistema experimental en la inoculación, y la única diferencia entre ellas es la capacidad de producir orfamidas (Figura 32). Si bien el rol de las orfamidas no está claramente definido, es evidente que su impacto en la colonización radicular en presencia de flora microbiana natural es significativo, siendo hasta 6 veces mayor la abundancia de la cepa silvestre sobre la mutante en el tratamiento a).

Para el análisis de los resultados de los tratamientos d) y e), debemos tener en cuenta que existen dos fenómenos distintos modulando la abundancia de cada cepa: la capacidad de producir orfamidas y el tipo de inoculación (ubicación inicial de las células en suelo o en semilla). Como se mencionó anteriormente las orfamidas impactan positivamente en el recuento de UFC en rizosfera (Figura 32), y la inoculación en suelo también lo hace en el mismo sentido (Figura 34). Esto explicaría que el tratamiento d) sea el que exhibió la mayor diferencia en la abundancia de la cepa silvestre sobre la mutante (entre 9 y 10 veces mayor; Figura 32). Ambos fenómenos operan en la misma dirección, generando una sinergia en la capacidad de colonización de la rizosfera (recuentos de UFC). Por el contrario, en el tratamiento e), la cepa silvestre fue inoculada en semilla, y la mutante Δ ofa en suelo. Es posible que el efecto positivo de la producción de surfactante en la capacidad de colonizar la rizosfera de las semillas por parte de la cepa silvestre localizada inicialmente en las semillas inoculadas se haya visto contrarrestado por la ventaja relativa de acceso a la rizosfera de la cepa Δ ofa presente en el suelo libre, dando como resultado una abundancia similar entre ambas cepas en este tratamiento en particular (e; Figura 32). Si bien la abundancia de la cepa silvestre fue mayor que la de la variante Δ ofa, la diferencia no resultó estadísticamente significativa (Figura 32).

Para ilustrar los resultados de la competencia por el acceso a la rizosfera en función de la capacidad de producción de orfamidas y de la ubicación inicial del inóculo, se graficaron las relaciones entre los recuentos de las cepas silvestre y Δ ofa (Figura 34). En esta representación, un valor igual a 1 equivaldría a la misma cantidad de UFC de células silvestre y Δ ofa. Como se advierte en la Figura 33, en todos los casos la abundancia de la cepa silvestre fue superior a la mutante al final del experimento, sin embargo, esta diferencia no fue homogénea en valor absoluto en todos los tratamientos, probablemente debido a las distintas metodologías de inoculación. Los tratamientos de co-inoculación en suelo y en semilla (b y c; respectivamente) presentaron relaciones de abundancia en rizosfera de aproximadamente el doble de células silvestres que Δ ofa (Figura 33). Sin embargo, hay que tener en cuenta que la carga del inóculo utilizado en el ensayo, es significativamente mayor a la carga de pseudomonas que naturalmente se estabiliza en el suelo utilizado en este ensayo (en el orden de 10^5 UFC/g suelo). En este sentido el tratamiento a), se asemeja más a lo que ocurre en un ambiente natural, donde la semilla es colonizada por las bacterias que habitan dicho suelo. En este tratamiento el suelo fue inoculado y luego transcurrió el tiempo suficiente para que se establezca una población residual de pseudomonas, la cual resultó de 2 órdenes de magnitud menor a la carga inicial introducida al inocular el suelo (Figura 30 y Figura 32). Al momento de la siembra, comienza la competencia por la colonización radicular, y en este caso, ambas cepas iniciaron a partir de una población de 10^5 UFC/g de suelo aproximadamente. Comparando con el tratamiento b), la diferencia es mucho mayor, a pesar de que ambos tratamientos son casi idénticos, excepto por que el a) parte de un inóculo menor, y entonces en estas condiciones de menor abundancia celular podría sacar mayor “provecho” de la ventaja otorgada por la orfamida. En conclusión, en un ambiente natural, el impacto de las orfamidas parece ser lo suficientemente relevante como para potenciar la capacidad de acceso a la rizosfera de la cepa silvestre CHAO aproximadamente 6 veces más que si no pudiese sintetizarla, y si bien en los tratamientos b) y c) la diferencia no

fue tan contundente, probablemente se deba a que ambas cepas se inocularon con un alto número de UFC, minimizando ligeramente las diferencias observadas en la competencia.

Relación silvestre/ ΔOfa

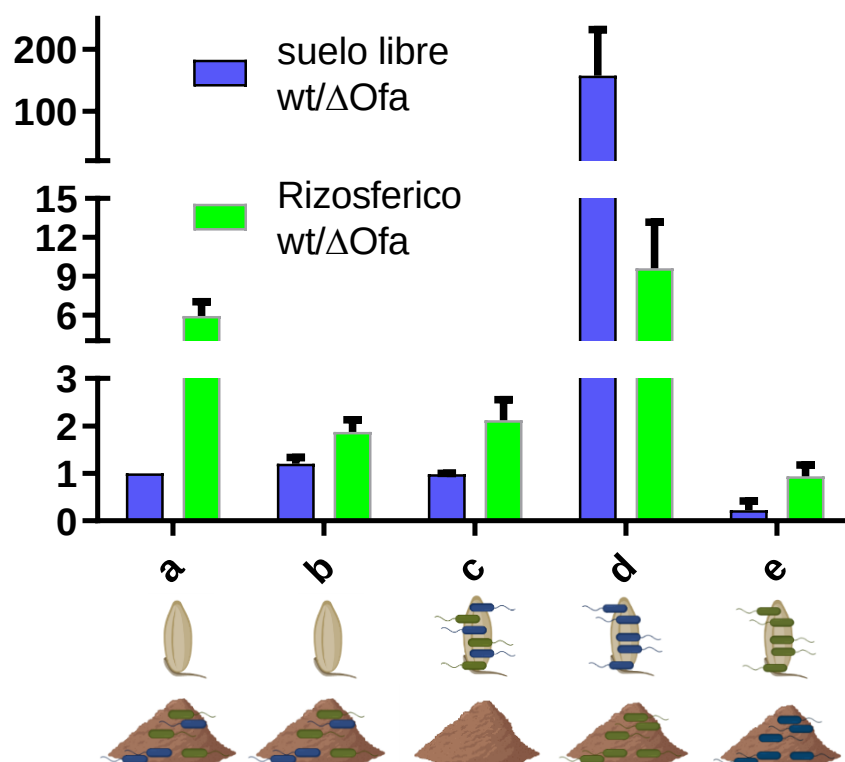
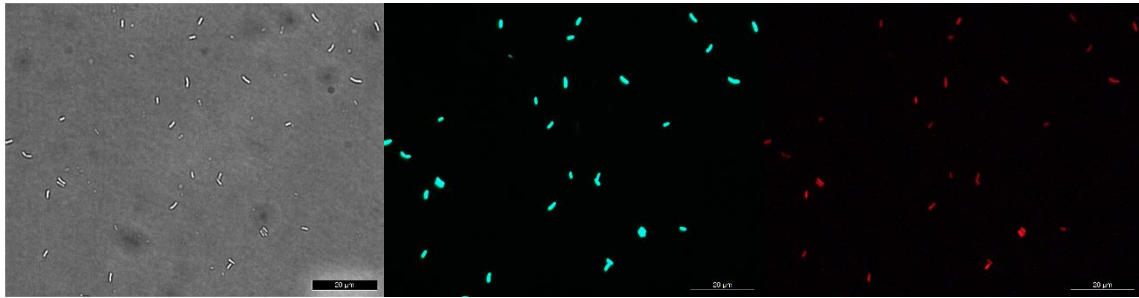


Figura 34. Relación entre recuentos de UFC de wt/ Δofa . a) Co-inoculación en suelo proveniente de la primera etapa. b) Co-inoculación en suelo. c) Co-inoculación en semilla. d) Inoculación en suelo con cepa silvestre, y en semilla con mutante. e) Inoculación en suelo con cepa mutante, y en semilla con silvestre.

2.6.4 Expresión de *ofa* en la rizosfera.

P. protegens CHA0 produce orfamidas cuando crece en cultivo puro en medios de cultivo complejos o definidos, desde dónde es posible aislar la actividad surfactante del sobrenadante y confirmar su presencia mediante bioensayos o determinación analítica (Capítulo I). Sin embargo, las condiciones fisiológicas de crecimiento en el suelo y en el ambiente rizosférico son sustantivamente diferentes, y podrían afectar la expresión de los genes biosintéticos de orfamida. Para investigar la expresión del *cluster* biosintético de orfamidas se recurrió a la observación de la expresión de cepas reporteras que indirectamente informarían sobre la capacidad biosintética de orfamidas en vida libre y sobre raíces de trigo.

CHA0/pME7408

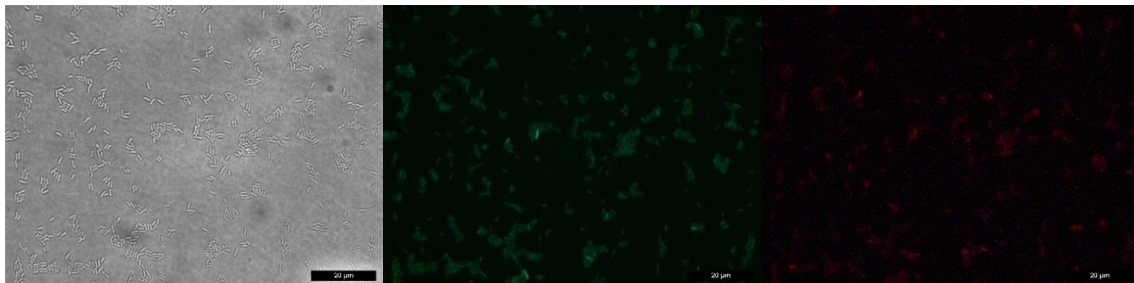


CHA89/pME7408



Figura 35. Microscopia de fluorescencia con un aumento de 1000x. Preparado de células provenientes de cultivo puro. El panel superior corresponde a la cepa *P. protegens* CHA0 silvestre con el plásmido reportero pME7408 (*prsmZ-gfp orfDR'-mCherry*). El panel inferior corresponde a la cepa mutante CHA89 (Δ *gacA*) con el plásmido reportero pME7408. De izquierda a derecha se puede ver el mismo campo con luz claro, el canal para visualizar GFP, y el canal para visualizar DsRed o mCherry.

CHA0/ pME6102



CHA89/ pME6102

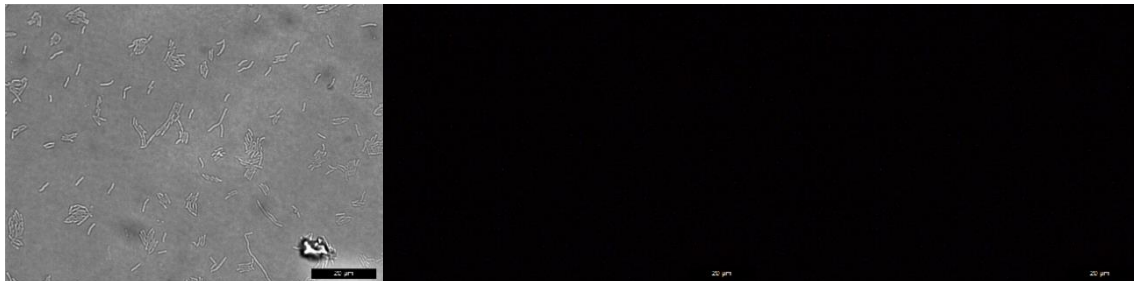


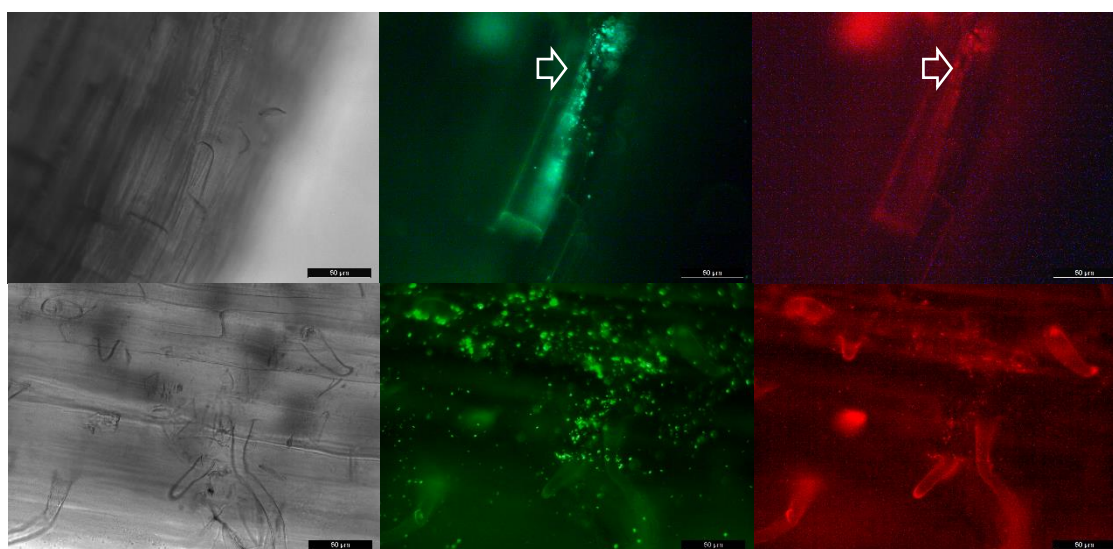
Figura 36. Microscopia de fluorescencia con un aumento de 1000x. Preparado de células provenientes de cultivo puro. El panel superior corresponde a la cepa *P. protegens* CHA0 silvestre con el plásmido reportero pME6102 (*orfUR'-gfp orfDR'-DsRed*). El panel inferior

corresponde a la cepa mutante CHA89 ($\Delta gacA$) con el plásmido reportero pME6102. De izquierda a derecha se puede ver el mismo campo con luz claro, filtro azul (para visualizar GFP), filtro verde (para visualizar DsRed o mCherry).

Para evidenciar indirectamente la expresión de los genes relacionados con la biosíntesis de la orfamida, se analizó la expresión de genes que codifican los reguladores transcripcionales OrfUR y OrfDR del cluster biosintético *ofaABC* (Capítulo I) mediante fusiones reporteras. Ambos reguladores deben ser expresados para activar la transcripción del cluster *ofaABC*, y además sabemos que su expresión está regulada positivamente por la cascada post-transcripcional Gac-Rsm (Capítulo I (Sobrero et al. 2017)). De esta forma, la detección de las proteínas reporteras fluorescentes evidenciaría la expresión del cluster biosintético para la producción de orfamida.

En primer lugar, verificamos la operatividad de la cascada Gac-Rsm mediante la expresión de una fusión reportera transcripcional del promotor del gen que codifica el riborregulador RsmZ y *gfp* (Dubuis et al. 2006), y la expresión (traduccional) del *orfDR* reportero en cultivo puro. En el panel superior de la Figura 35 se puede observar que la cepa silvestre expresa tanto *gfp* como *mCherry*, es decir la cascada Gac-Rsm está activa, y el gen *orfDR* (que codifica uno de los activadores del cluster de orfamida) se encuentra también activo (se transcribe y se traduce). En el panel inferior se puede ver que si la cascada Gac-Rsm no es funcional (ausencia de expresión de *rsmZ*), no se expresa el gen *orfDR*, y por lo tanto no se sintetizará orfamida. De forma complementaria, como se ilustra en la Figura 36, se confirmó que la expresión de los dos reguladores transcripcionales de *ofaABC*, *orfUR* y *orfDR*, son totalmente dependientes de la cascada Gac-Rsm. Cuando la cascada esta activa (panel superior) ambos se expresan, pero en un contexto $\Delta gacA$ no hay expresión detectable. Una vez confirmada la visualización de la expresión de *orfUR* y *orfDR* a través de microscopia de fluorescencia, se procedió a analizar las mismas cepas reporteras en raíces de trigo.

CHA0/pME7408



CHA89/pME7408

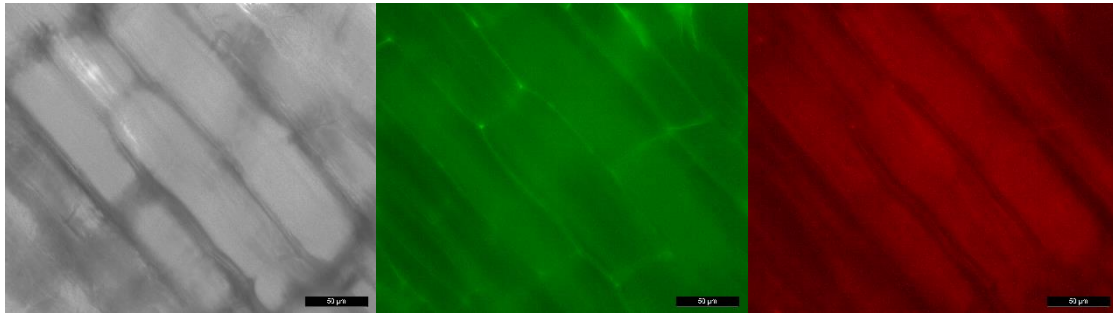
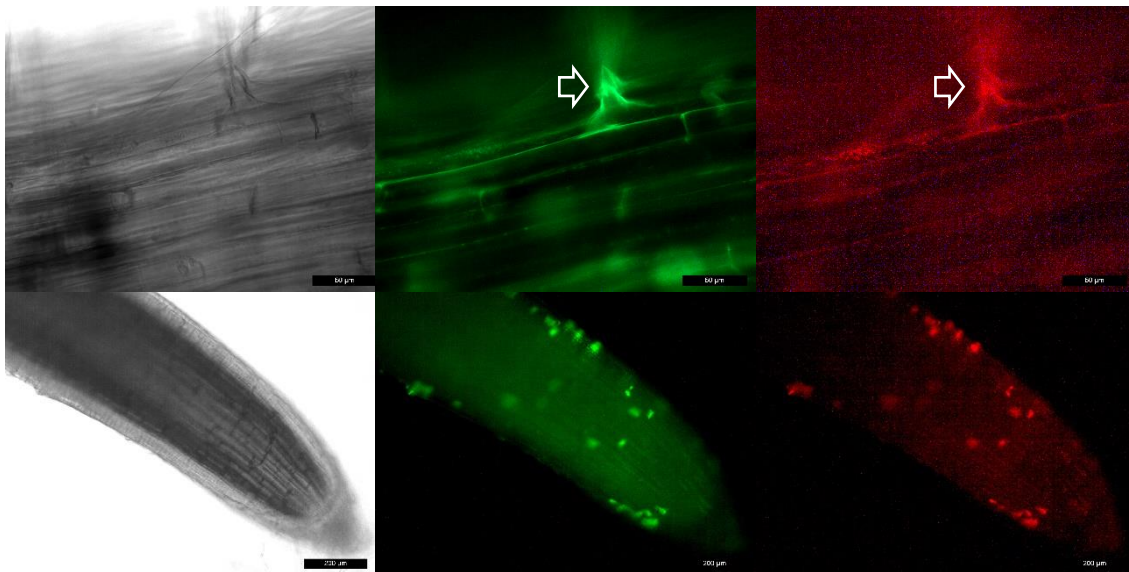


Figura 37. Preparado de raíces crecidas por 7 días en perlita estéril previamente inoculadas en semilla. El panel superior corresponde a la cepa *P. protegens* CHA0 silvestre con el plásmido reportero pME7408 (*prsmZ-gfp orfDR'-mCherry*). El panel inferior corresponde a la cepa mutante CHA89 ($\Delta gacA$) con el plásmido reportero pME7408. De izquierda a derecha se puede ver el mismo campo con luz claro, filtro azul (para visualizar GF), filtro verde (para visualizar DsRed o mCherry). En cada imagen se indica la escala de referencia de longitud (μm).

CHA0/ pME6102



CHA89/ pME6102

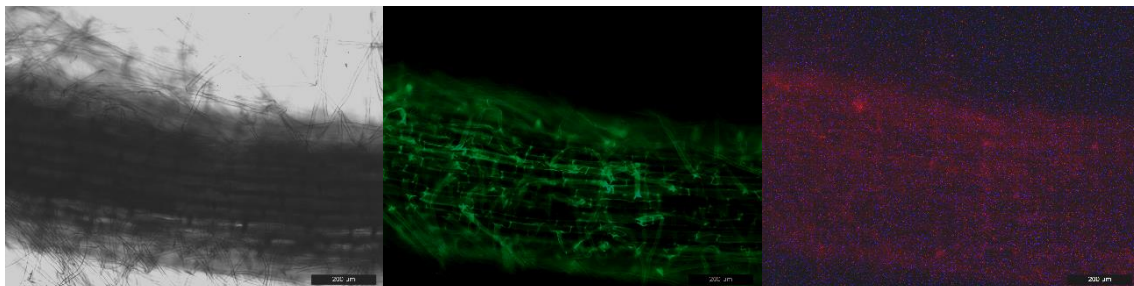


Figura 38. Preparado de raíces crecidas por 7 días en perlita estéril previamente inoculadas en semilla. Las tres primeras imágenes del panel superior corresponden a la cepa *P. protegens* CHA0 silvestre con el plásmido reportero pME6102 (*orfUR*'-'*gfp orfDR*'-'*DsRed*). El panel inferior corresponde a la cepa mutante CHA89 ($\Delta gacA$) con el plásmido reportero pME6102. De izquierda a derecha se puede ver el mismo campo con luz claro, filtro azul (para visualizar GFP), filtro verde (para visualizar DsRed o mCherry). En cada imagen se indica la escala de referencia de longitud (μm)

En la Figura 37, panel superior, se puede ver un conjunto de células adheridas a la raíz expresando tanto *gfp* como *mCherry*. Al igual que en la situación de cultivo puro, tanto la cascada Gac-Rsm como la expresión de uno de los reguladores del *cluster ofaABC* se encuentran activos. En el panel superior de la Figura 38 se observa que la expresión de ambos reguladores *orfUR* (canal verde) y *orfDR* (canal rojo), también está activa, por lo tanto, concluimos que las orfamidas sería sintetizada cuando la cepa silvestre CHA0 se encuentra colonizando las raíces de trigo. En el panel inferior de ambas figuras se muestran imágenes representativas de la expresión de las fusiones reporteras en el contexto de una cepa mutante *gacA* en la que la cascada Gac-Rsm no es funcional. En ningún caso fue posible detectar bacterias fluorescentes.

Teniendo en cuenta estas observaciones, podemos inferir que la cepa silvestre *P. protegens* CHA0 efectivamente sintetiza orfamidas en las condiciones del ensayo realizado en invernadero, y por lo tanto, las diferencias obtenidas en la colonización radicular (Figura 32, Figura 34, Figura 33) reflejarían la contribución del lipopéptido en al establecimiento de la bacteria en la rizosfera de trigo.

2.7 Conclusiones

La producción del surfactante orfamida no parece tener ningún efecto sobre la persistencia de la bacteria *P. protegens* CHA0 cuando esta se encuentra en vida libre en suelo, sin estar asociada a raíces. Por el contrario, cuando CHA0 tiene la posibilidad de acceder al nicho rizosférico, las orfamidas parecen tener un efecto positivo muy marcado en la competitividad por la colonización de raíces de trigo. Si bien las orfamidas son un metabolito especializado extracelular, el efecto positivo que observamos para la cepa productora no se trasladó a la variante no productora, de manera que las orfamidas no se comportarían como un “bien común” en las condiciones de este ensayo.

2.8 Bibliografía

- Agaras, B., et al. (2012). "Specific enumeration and analysis of the community structure of culturable pseudomonads in agricultural soils under no-till management in Argentina." **61**: 305-319.
- Andersen, J. B., et al. (2003). "Surface motility in *Pseudomonas* sp. DSS73 is required for efficient biological containment of the root-pathogenic microfungi *Rhizoctonia solani* and *Pythium ultimum*." **149**(1): 37-46.
- Bao, Y., et al. (1991). "An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria." **109**(1): 167-168.
- Blumer, C., et al. (1999). "Global GacA-steered control of cyanide and exoprotease production in *Pseudomonas fluorescens* involves specific ribosome binding sites." **96**(24): 14073-14078.
- Bradley, D. E. (1972). Evidence for the retraction of *Pseudomonas aeruginosa* RNA phage pili. *Biochemical and biophysical research communications*, **47**(1), 142-149.
- Boyer, H. W. and D. J. J. o. m. b. Roulland-Dussoix (1969). "A complementation analysis of the restriction and modification of DNA in *Escherichia coli*." **41**(3): 459-472.
- Casabadan, J. and S. J. I.-. Cohen (1980). "Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli* Journal of Molecular Biology **138**."
- D'aes, J., et al. (2014). "To settle or to move? The interplay between two classes of cyclic lipopeptides in the biocontrol strain *Pseudomonas* CMR 12a." **16**(7): 2282-2300.
- De Bruijn, I., et al. (2007). "Genome-based discovery, structure prediction and functional analysis of cyclic lipopeptide antibiotics in *Pseudomonas* species." **63**(2): 417-428.
- Dubuis, C., et al. (2006). "Thiamine-auxotrophic mutants of *Pseudomonas fluorescens* CHA0 are defective in cell-cell signaling and biocontrol factor expression." *Appl Environ Microbiol* **72**(4): 2606-2613.
- Flury, P., et al. (2017). "Antimicrobial and insecticidal: cyclic lipopeptides and hydrogen cyanide produced by plant-beneficial *Pseudomonas* strains CHA0, CMR12a, and PCL1391 contribute to insect killing." **8**: 100.
- Gould, W. D., et al. (1985). "New selective media for enumeration and recovery of fluorescent pseudomonads from various habitats." *Appl Environ Microbiol* **49**(1): 28-32.
- Green, M. R. and J. J. A. L. M. t. Sambrook (2012). "Molecular cloning."
- Haas, D. and G. J. N. r. m. Défago (2005). "Biological control of soil-borne pathogens by fluorescent pseudomonads." **3**(4): 307-319.
- Harshey, R. M. J. A. r. o. m. (2003). "Bacterial motility on a surface: many ways to a common goal." **57**: 249.
- Heeb, S., et al. (2002). "Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in *Pseudomonas fluorescens* CHA0." *J Bacteriol* **184**(4): 1046-1056.
- Henrichsen, J. J. B. r. (1972). "Bacterial surface translocation: a survey and a classification." **36**(4): 478-503.
- Hildebrand, P., et al. (1998). "Role of the biosurfactant viscosin in broccoli head rot caused by a pectolytic strain of *Pseudomonas fluorescens*." **20**(3): 296-303.

- Jousset, A., et al. (2006). "Secondary metabolites help biocontrol strain *Pseudomonas fluorescens* CHA0 to escape protozoan grazing." Appl Environ Microbiol **72**(11): 7083-7090.
- Jousset, A., et al. (2006). "Secondary metabolites help biocontrol strain *Pseudomonas fluorescens* CHA0 to escape protozoan grazing." **72**(11): 7083-7090.
- Kay, E., et al. (2005). "Three small RNAs jointly ensure secondary metabolism and biocontrol in *Pseudomonas fluorescens* CHA0." **102**(47): 17136-17141.
- Kessler, B., et al. (1992). "A general system to integrate lacZ fusions into the chromosomes of gram-negative eubacteria: regulation of the P_m promoter of the TOL plasmid studied with all controlling elements in monocopy." **233**(1): 293-301.
- Koch, B., et al. (2001). "A panel of Tn7-based vectors for insertion of the gfp marker gene or for delivery of cloned DNA into Gram-negative bacteria at a neutral chromosomal site." J Microbiol Methods **45**(3): 187-195.
- Kruijt, M., et al. (2009). "Functional, genetic and chemical characterization of biosurfactants produced by plant growth-promoting *Pseudomonas putida* 267." **107**(2): 546-556.
- Kuiper, I., et al. (2004). "Characterization of two *Pseudomonas putida* lipopeptide biosurfactants, putisolvin I and II, which inhibit biofilm formation and break down existing biofilms." **51**(1): 97-113.
- Lambertsen, L., et al. (2004). "Mini-Tn7 transposons for site-specific tagging of bacteria with fluorescent proteins." Environ Microbiol **6**(7): 726-732.
- Li, W., et al. (2013). "The antimicrobial compound xantholysin defines a new group of *Pseudomonas* cyclic lipopeptides." **8**(5): e62946.
- Martínez-García, E., et al. (2014). "New transposon tools tailored for metabolic engineering of Gram-negative microbial cell factories." **2**: 46.
- Mascher, F., et al. (2002). "Inactivation of the regulatory gene algU or gacA can affect the ability of biocontrol *Pseudomonas fluorescens* CHA0 to persist as culturable cells in nonsterile soil." Appl Environ Microbiol **68**(4): 2085-2088.
- Maurhofer, M., et al. (1998). "Salicylic acid biosynthetic genes expressed in *Pseudomonas fluorescens* strain P3 improve the induction of systemic resistance in tobacco against tobacco necrosis virus." **88**(7): 678-684.
- Mazzola, M., et al. (2009). "Protozoan-induced regulation of cyclic lipopeptide biosynthesis is an effective predation defense mechanism for *Pseudomonas fluorescens*." **75**(21): 6804-6811.
- Miller, V. L. and J. J. J. o. b. Mekalanos (1988). "A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires toxR." **170**(6): 2575-2583.
- Naghili, H., et al. (2013). "Validation of drop plate technique for bacterial enumeration by parametric and nonparametric tests." Vet Res Forum **4**(3): 179-183.
- Nielsen, T. H., et al. (2005). "Genes involved in cyclic lipopeptide production are important for seed and straw colonization by *Pseudomonas* sp. strain DSS73." Appl Environ Microbiol **71**(7): 4112-4116.
- Pauwelyn, E., et al. (2013). "New linear lipopeptides produced by *Pseudomonas cichorii* SF1-54 are involved in virulence, swarming motility, and biofilm formation." **26**(5): 585-598.
- Raaijmakers, J. M., et al. (2010). "Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics." **34**(6): 1037-1062.

- Reimann, C., et al. (2005). "Posttranscriptional repression of GacS/GacA-controlled genes by the RNA-binding protein RsmE acting together with RsmA in the biocontrol strain *Pseudomonas fluorescens* CHA0." **187**(1): 276-285.
- Rokni-Zadeh, H., et al. (2013). "Distinct lipopeptide production systems for WLIP (white line-inducing principle) in *Pseudomonas fluorescens* and *Pseudomonas putida*." **5**(1): 160-169.
- Roongsawang, N., et al. (2003). "Cloning and characterization of the gene cluster encoding arthofactin synthetase from *Pseudomonas* sp. MIS38." **10**(9): 869-880.
- Sobrero, P. M., et al. (2017). "A matter of hierarchy: activation of orfamide production by the post-transcriptional Gac-Rsm cascade of *Pseudomonas protegens* CHA0 through expression upregulation of the two dedicated transcriptional regulators." *Environ Microbiol Rep* **9**(5): 599-611.
- Song, C., et al. (2016). "Living on the edge: emergence of spontaneous gac mutations in *Pseudomonas protegens* during swarming motility." *Environ Microbiol* **18**(10): 3453-3465.
- Stutz, E., et al. (1986). "Naturally occurring fluorescent *Pseudomonads* involved in suppression." **76**(2): 181-185.
- Torres-Quesada, O., et al. (2013). "Independent activity of the homologous small regulatory RNAs AbcR1 and AbcR2 in the legume symbiont *Sinorhizobium meliloti*." **8**(7): e68147.
- Tran, H., et al. (2007). "Role of the cyclic lipopeptide massetolide A in biological control of *Phytophthora infestans* and in colonization of tomato plants by *Pseudomonas fluorescens*." **175**(4): 731-742.
- Zuber, S., et al. (2003). "GacS sensor domains pertinent to the regulation of exoproduct formation and to the biocontrol potential of *Pseudomonas fluorescens* CHA0." **16**(7): 634-644.

Capítulo 3

Aplicación de *molecular networking* para la exploración de metabolitos especializados extracelulares de aislamientos autóctonos de *Pseudomonas*.

Lic. Andrés Muzlera

Laboratorio de Fisiología y Genética de
Bacterias Beneficiosas para Plantas

3.1 Resumen

Nuestro laboratorio cuenta con una colección de aislamientos de *Pseudomonas* de gran interés debido a que presentan fenotipos biocontroladores frente a una variedad de hongos fitopatógenos (Agaras et al. 2015). Estos fenotipos en general, están asociados a la producción de metabolitos especializados y su naturaleza química suele ser muy diversa (Introducción Capítulo 1). Como los compuestos tipo lipopéptidos son de particular interés para esta Tesis, se seleccionaron los aislamientos que presentaran actividad surfactante en su sobrenadante de cultivo, para caracterizar posibles compuestos tipo lipopéptidos. El siguiente Capítulo describe la implementación de una metodología basada en *Molecular Networking* aplicada a la identificación de compuestos naturales, en sobrenadantes de cultivos de diferentes aislamientos de *Pseudomonas*.

3.2 Introducción

3.2.1 Colección de aislamientos autóctonos de la región agrícola núcleo Argentina. Proyecto BIOSPAS

En 2009 comenzó un proyecto público-privado denominado BIOSPAS, el cual tuvo como objetivo identificar indicadores biológicos de la calidad y productividad del suelo en campos agrícolas bajo manejo de labranza convencional en Argentina (Wall, et al. 2011) . Dentro de este consorcio, nuestro laboratorio se focalizó en caracterizar las poblaciones de *Pseudomonas* en dichos suelos. Durante el transcurso de este trabajo, varios aislamientos de *Pseudomonas* fueron conservados, y caracterizados en función de sus propiedades probióticas para el suelo y su potencial antagonismo *in vitro* sobre diversos hongos fitopatógenos (Agaras et al. 2015).

Este conjunto de aislamientos comprende una colección de bacterias pertenecientes al género *Pseudomonas* capaz de sintetizar una gran cantidad y variedad de compuestos naturales con actividades antibióticas y antifúngicas, entre otras (Agaras et al. 2015). La caracterización de estos metabolitos es un proceso costoso y de gran demanda de tiempo, es por esto que en su mayoría no están caracterizados. En la Tabla 5 se pueden observar los aislamientos que presentaron algún tipo de antagonismo *in vitro* frente a hongos fitopatógenos (Agaras et al. 2015). Los nombres de los aislamientos fueron designados como acrónimos del tipo de muestra (S, *bulk soil*; R, *rhizosphere*); el sitio de muestreo (B, Bengolea; M, Monte Buey; P, Pergamino; V, Viale) y el manejo agrícola (BP, buenas prácticas agrícolas o MP, malas prácticas agrícolas), seguido de su número de serie. Por ejemplo, el aislamiento SVBP6 corresponde al aislamiento #6 del suelo libre (S) de una parcela ubicada en Viale (V) que fue manejada bajo buenas prácticas agrícolas (BP).

Tabla 5. Colección de aislamientos de pseudomonas autóctonos de la región agrícola núcleo argentina que mostraron potencial antagonista frente a algún hongo fitopatógeno. (Agaras et al. 2015)

Aislamiento	<i>Pseudomonas</i> complex (aproximación por MLSA *)
RPAN5	<i>P. chlororaphis</i>
SPSA2	<i>P. chlororaphis</i>
SPAN5	<i>P. chlororaphis</i>
SVBP3	<i>P. chlororaphis</i>
SVBP8	<i>P. chlororaphis</i>
SMMP3	<i>P. chlororaphis</i>
SVBP6	<i>P. putida</i>
RPAN1	<i>P. chlororaphis</i>
SVAN4	<i>P. putida</i>
RBAN4	<i>P. asplenii</i>
SMAN5	<i>P. putida</i>
RBMP1	<i>P. fluorescens</i>
RMAN6	<i>P. fluorescens</i>
RBBP4	<i>P. fluorescens</i>
SBMP6	<i>P. putida</i>
SVMP4	<i>P. putida</i>
RMAN4	<i>P. fluorescens</i>
RPBP2	<i>P. asplenii</i>
RMAN5	<i>P. fluorescens</i>
RPAN5	<i>P. chlororaphis</i>
SPSA2	<i>P. chlororaphis</i>
SPAN5	<i>P. chlororaphis</i>
SVBP3	<i>P. chlororaphis</i>
SVBP8	<i>P. chlororaphis</i>
SMMP3	<i>P. chlororaphis</i>
SVBP6	<i>P. putida</i>
RPAN1	<i>P. chlororaphis</i>
SVAN4	<i>P. putida</i>
RBAN4	<i>P. asplenii</i>
SMAN5	<i>P. putida</i>
RBMP1	<i>P. fluorescens</i>
RMAN6	<i>P. fluorescens</i>
RBBP4	<i>P. fluorescens</i>
SBMP6	<i>P. putida</i>
SVMP4	<i>P. putida</i>
RMAN4	<i>P. fluorescens</i>
RPBP2	<i>P. asplenii</i>
RMAN5	<i>P. fluorescens</i>

Los aislamientos fueron asignados a un complejo dentro género *Pseudomonas* a partir de un análisis Multilocus de secuencia (MLSA) utilizando un concatenado de los genes ARNr 16s, *rpoB* y *oprF*. Para obtener una mayor resolución de las relaciones filogenéticas de especies dentro de un género o géneros dentro de una familia, el análisis de secuencias multilocus (del inglés

multilocus sequence analysis, MLSA) es actualmente un método ampliamente utilizado, por lo que a continuación, se describen algunas de sus principales características.

3.2.2 Análisis multilocus de secuencia

MLSA se basa en la tipificación de secuencias multilocus (MLST), que fue introducida por primera vez por Maiden et al. en 1998 como método de tipificación microbiana para estudios epidemiológicos y genéticos de poblaciones de bacterias patógenas (Maiden et al. 1998). MLST se define como una técnica de tipificación para la caracterización de cepas que muestran variaciones en múltiples genes de mantenimiento (Maiden 2006). Cada nueva variación en la secuencia de los fragmentos, incluso en un solo nucleótido, conduce a la asignación de un nuevo alelo único del gen respectivo. Los alelos se numeran arbitrariamente en el orden de aparición, pero los números de alelo obtenidos para los diferentes genes se combinan en un orden específico y el perfil de alelo obtenido se define como el tipo de secuencia (ST). Las cepas que comparten perfiles alélicos idénticos se refieren al mismo tipo de secuencia, mientras que las cepas que solo comparten algunos de los alelos están relacionadas y, por lo tanto, se refieren a complejos de secuencia (complejos CC) (Maiden et al. 1998).

Los análisis MLST en estudios epidemiológicos no pretenden calcular las relaciones filogenéticas, sino que utilizan la agrupación basada en "perfiles alélicos" para mostrar la relación y la aparición de tipos de secuencias y complejos de secuencias. Sin embargo, debe señalarse que los árboles resultantes no se basan en similitudes entre diferentes alelos. La información de secuencia que respalda los perfiles alélicos se puede emplear para cálculos filogenéticos y, por lo tanto, se puede usar para la identificación y clasificación de bacterias porque ofrece la oportunidad de proporcionar una visión más profunda de las relaciones filogenéticas de las especies procarióticas, lo que originalmente condujo a la definición de MLSA (Gevers et al. 2005).

En 2003, Zeigler ya estaba usando genes codificantes para predecir resultados basados en datos de secuenciación de genoma completo (Zeigler, et al 2003). Dos años más tarde, en 2005, Gevers et al. publicó un artículo de opinión sobre la reevaluación de las especies procarióticas, que se centró en el impacto de los enfoques basados en secuencias de nucleótidos multilocus para la taxonomía de procariotas (Gevers et al. 2005). Introdujeron el término MLSA como una nueva designación para el análisis de secuencias de múltiples genes que codifican proteínas para su aplicación taxonómica a la delimitación de especies dentro de un género. Como resultado, la integración de MLSA en la taxonomía procariótica debería conducir a un procedimiento de identificación de dos pasos. En primer lugar, se debe aplicar el análisis de la secuencia del gen 16S ARNr para asignar una nueva cepa al nivel de familia o incluso de género. En segundo lugar, en función de la asignación inicial, el aislamiento debe asignarse con mayor precisión dentro de un género utilizando el enfoque MLSA. Deben usarse (idealmente) genes que sean copias únicas y ubicuos, al menos en el taxón estudiado, y deben evitarse los genes que podrían conferir ventajas selectivas, como los genes de virulencia. Esto contrasta claramente con los estudios MLST, en los que se recomienda el uso de estos genes si se requiere un mayor poder de resolución intraespecie para estudios epidemiológicos (Maiden 2006).

Una evaluación genómica de MLSA realizada por Konstantinidis and Tiedje (2007) ilustró que una concatenación de ocho genes codificantes proporcionó una resolución filogenética robusta para las cepas que comparten entre el 70% y el 95% de identidad de nucleótidos promedio (ANI (Konstantinidis and Tiedje 2005)). Los autores afirmaron que MLSA constituye un compromiso razonable entre los métodos basados en un solo gen y los basados en el genoma completo para estudios de diversidad de especies, pero señalaron claramente que el índice ANI basado en el genoma completo se determinó entre varios parámetros probados como el más robusto. El potencial de los análisis por MLSA para la asignación taxonómica es un tema que será discutido en el próximo Capítulo.

3.2.3 Los productos naturales como objeto de estudio

La naturaleza nos ha proporcionado una fuente inigualable de pequeñas moléculas que han desempeñado y siguen desempeñando un papel destacado en diversas áreas de aplicación. Aproximadamente dos tercios de las terapias antibacterianas usadas clínicamente se derivan de productos naturales (Newman and Cragg 2012). Este éxito se puede atribuir en gran parte a propiedades únicas que han sido perfeccionadas por procesos evolutivos para proporcionar al organismo productor una ventaja selectiva. Atributos tales como la permeabilidad celular intrínseca, la gran diversidad química y la especificidad del objetivo (Stone and Williams 1992) a menudo están ausentes en las bibliotecas de compuestos químicos sintéticos. Sin embargo, tras los descubrimientos revolucionarios de la mayoría de las clases de fármacos antibacterianos conocidos hoy en día, las principales compañías farmacéuticas por ejemplo, han disminuido casi por completo su enfoque en los productos naturales en las últimas dos décadas (Li and Vederas 2009). La culpa es de las complejidades inherentes del descubrimiento de fármacos de productos naturales. Los esfuerzos se ven limitados por los desafíos del redescubrimiento de estructuras conocidas luego de una purificación y caracterización laboriosa, la baja disponibilidad de compuestos y los inevitables falsos positivos debido a la interferencia de otras sustancias en los extractos (Li and Vederas 2009).

3.2.4 Espectrometría de masas aplicadas al descubrimiento de productos naturales

La espectrometría de masas (MS) ha experimentado importantes avances instrumentales en los últimos quince años. El progreso en las técnicas de ionización, principalmente el desarrollo de fuentes de ionización ambiental, ha llevado al análisis de muestras en el ambiente libre (Monge et al. 2013; Ding and Duan 2015). Además, se ha logrado un progreso significativo en términos de resolución de masa y precisión (Nikolaev et al. 2011; Junot et al. 2014). Los espectrómetros de masas se han adaptado progresivamente con instrucciones de uso más sencillas para operar y mantener los instrumentos, lo que hace que esta herramienta analítica sea accesible para una comunidad más amplia de científicos. La facilidad de uso, la sensibilidad y la solidez combinadas con los costos razonables de la instrumentación han llevado a los investigadores de productos

naturales a adoptar la espectrometría de masas para el uso diario en el laboratorio (Bouslimani et al. 2014).

El cuarto premio Nobel de la espectrometría de masas llegó con el desarrollo de la ionización por electropulverización (Fenn et al. 1989) y la ionización por desorción láser asistida por matriz (MALDI) (Karas and Hillenkamp 1988; Hillenkamp and Karas 1990) en la década de 1980. Estas innovaciones permitieron que esta tecnología se utilizara para aplicaciones biológicas. Desde entonces, la comunidad de científicos de productos naturales ha adoptado, utilizado y avanzado ampliamente ambas técnicas. Para el estudio de compuestos naturales es particularmente relevante el resurgimiento de los métodos de ionización ambiental después del trabajo Graham Cooks sobre espectrometría de masas de ionización por electropulverización por desorción (DESI) en 2004, (Takats et al. 2004) y el trabajo de Robert B. Cody sobre análisis directo en tiempo real (DART) en 2005 (Cody, Laramée, and Durst 2005). Ambos métodos ofrecieron la capacidad de generar datos de espectrometría de masas a presión atmosférica sin necesidad de preparar muestras. El método DESI emplea electropulverización de gotas cargadas de solvente orgánico en la superficie de la muestra a través de un flujo de gas que resulta en la desorción de los materiales (Figura 39a), mientras que DART permite el análisis directo simplemente exponiendo la muestra a una corriente de gas excitado por plasma calentado (Figura 39b).

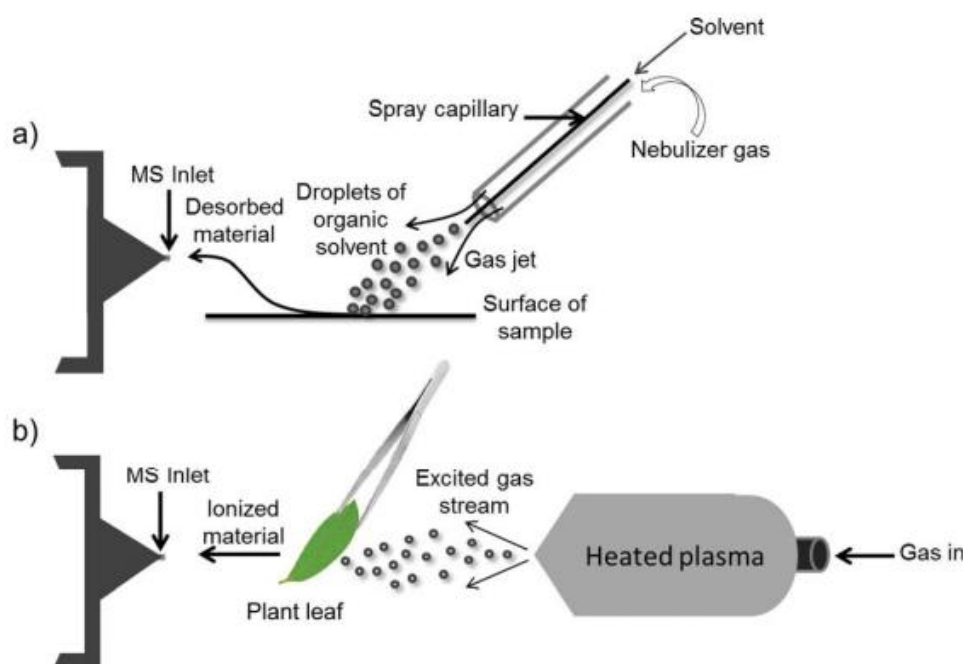


Figura 39. Dos fuentes de ionización atmosférica representativas. a) DESI b) Métodos de ionización DART. (Bouslimani et al. 2014)

Las técnicas analíticas básicas de espectrometría de masas, (Alexandrov and Kobarg 2011; Ye, Greer, and Li 2011) y la espectroscopia de RMN (Alexandrov et al. 2013; Seeley and Caprioli 2012; Watrous et al. 2013; Lane et al. 2009) se están utilizando para desreplicar (*dereplication*) y cuantificar metabolitos conocidos frente a nuevos productos naturales (Esquenazi, Dorrestein,

and Gerwick 2009; Kraft, Klitzing, and Lipids 2014). El concepto de desreplicar hace referencia al uso de análisis cromatográfico y espectroscópico para reconocer sustancias previamente caracterizadas presentes en un extracto. La desreplicación es un trabajo costoso (Debois et al. 2008; Vaidyanathan et al. 2008) porque los metabolitos secundarios tienen diversos arreglos atómicos, lo que resulta en variaciones en las propiedades químicas y físicas. Se debe utilizar una combinación de métodos analíticos que incluyan ultravioleta, espectrometría de masas en tándem y datos espectrales de RMN para garantizar la exactitud de la identificación de la estructura básica de los compuestos desreplicados en muestras crudas. El conjunto de datos UV está limitado en el análisis y desreplicación de metabolitos que contienen grupos cromóforos; mientras que, en la espectrometría de masas, existe el riesgo de perder compuestos poco ionizantes que solo son detectables en un modo. Por ejemplo, los compuestos fenólicos y de antraquinona se ionizan poco en el modo positivo, pero se ionizan muy bien en el modo negativo (Kraft, Klitzing, and Lipids 2014), mientras que se puede decir lo contrario de los alcaloides. Los metabolitos secundarios también se pueden encontrar en una amplia gama de concentraciones, por lo tanto, se requieren métodos analíticos confiables, robustos, selectivos y de alta resolución para identificar y cuantificar los múltiples grupos químicos de los productos naturales.

3.2.5 Desreplicación como metodología para la identificación rápida de compuestos naturales

La desreplicación es una parte integral del flujo de trabajo de descubrimiento de productos naturales (Lang et al. 2008). Hay un rápido progreso en la tecnología de la información (Esquenazi, Dorrestein, and Gerwick 2009), para el análisis rápido y eficiente de grandes conjuntos de datos de bibliotecas de productos naturales (Goodwin et al. 2012). Estos aprovechan enfoques innovadores para analizar datos de espectrometría de masas, por ejemplo, MZmatch (Buss and Butler 2010; Lang et al. 2008) y XCMS (Tu and Yan 2012), pueden ser utilizados para análisis de expresión diferencial y con bases de datos en línea (como ChemSpider, MarinLit y el Diccionario de productos naturales) o bases de datos internas (Yang et al. 2009). El análisis de expresión diferencial (Lang et al. 2008) implica una serie de procesos que incluyen el alineamiento no lineal del tiempo de retención de compuestos que son exclusivos de la muestra, seguida de la filtración combinada de picos auténticos en espectros de masas para la muestra sobre el ruido y los picos de fondo (como los que resultan de solventes o medios de cultivo). Finalmente, se detectan los picos y los metabolitos conocidos se comparan con los compuestos utilizando una base de datos como el Diccionario de Productos Naturales (<https://dnp.chemnetbase.com>).

Como se mencionó anteriormente, es posible analizar productos naturales de muestras complejas por espectrometría de masas. Sin embargo, un solo experimento puede recopilar miles de espectros de MS/MS en minutos (Bouslimani et al. 2014), y los proyectos individuales pueden adquirir millones de espectros. Estos conjuntos de datos son demasiado grandes para el análisis manual. Además, el software integral y la infraestructura computacional adecuada no se encuentran fácilmente disponibles. La información potencialmente útil en los conjuntos de

datos de MS/MS puede permanecer enterrada en documentos, cuadernos de laboratorio y bases de datos privadas, lo que dificulta la recuperación, extracción y el intercambio de datos y conocimientos. Aunque varias bases de datos de productos naturales (Dictionary of Natural Products, AntiBase (Laatsch 2011) y MarinLit (Blunt and Munro 2003)) ayudan en la desreplicación, estos recursos no están disponibles gratuitamente y no procesan datos de espectros de masa. Por el contrario, las bases de datos de espectros de masas, incluidas MassBank (Horai et al. 2010), Metlin (Smith et al. 2005), mzCloud (<https://www.mzcloud.org/>) y ReSpect (Sawada et al. 2012), albergan espectros de MS/MS, pero limitan los análisis de datos a varios espectros individuales o a una cantidad limitada de archivos de cromatografía líquida (LC)-MS. Existen otros recursos informáticos gratuitos en línea que aprovechan los espectros MS/MS de Metlin, como los proporcionados por mzCloud y XCMS Online. Sin embargo, ninguno de los dos permite la descarga gratuita de su biblioteca de referencia (Wang et al. 2016).

Teniendo en cuenta las problemáticas actuales, la plataforma GNPS (*Global Natural Products Social Molecular Networking*; <https://gnps.ucsd.edu>) proporciona una solución adecuada, accesible y colaborativa para en el análisis e identificación de compuestos naturales en muestras complejas. GNPS brinda la capacidad de analizar un conjunto de datos y compararlo con todos los datos disponibles públicamente. La infraestructura computacional se encuentra en el Centro de Espectrometría Computacional de Masas (CCMS; <http://proteomics.ucsd.edu/>) de la Universidad de California en San Diego (UCSD), por lo tanto, GNPS proporciona depósito y/o recuperación de conjuntos de datos públicos a través del entorno virtual interactivo de espectros de masas (*Mass Spectrometry interactive virtual environment*, MassIVE). La infraestructura de análisis de GNPS permite además la desreplicación en línea (Lang et al. 2008; Ito and Masubuchi 2014; Little et al. 2011), el análisis de redes moleculares automatizado (Watrous et al. 2012) y la curación de espectro MS/MS por colaboración colectiva. Cada conjunto de datos ingresado al repositorio de GNPS se vuelve a analizar automáticamente en el siguiente ciclo mensual de identificación continua. Cada una de estas decenas de millones de espectros en conjuntos de datos GNPS se compara con bibliotecas espectrales de referencia para anotar moléculas y descubrir análogos putativos.

3.2.6 *Molecular Networking*

Las redes moleculares son visualizaciones del espacio químico presente en los experimentos de espectrometría de masas. GNPS se puede utilizar para redes moleculares (Watrous et al. 2012), o para un enfoque de visualización y correlación espectral que puede detectar conjuntos de espectros de moléculas relacionadas (las llamadas redes espectrales (Bandeira et al. 2007)), incluso cuando los espectros en sí mismos no coinciden con ningún compuesto conocido. El alineamiento espectral (Guthals et al. 2012; Watrous et al. 2012) detecta espectros similares de moléculas estructuralmente relacionadas, asumiendo que estas moléculas se fragmentan de manera similar reflejada en sus patrones de MS/MS, de forma análoga a la detección de secuencias de proteínas o nucleótidos relacionadas por alineamiento de secuencias.

Dado que la fragmentación de MS de cada molécula individual da como resultado una huella dactilar de MS/MS única, se desarrollaron flujos de trabajo basados en redes para organizar grandes conjuntos de datos de espectros de masas en tándem basados en la similitud entre los patrones de fragmentación de iones precursores diferentes, pero relacionados (Watrous et al. 2012). Mediante el uso de una variación de redes espectrales diseñadas para aplicaciones proteómicas, los datos se simplifican inicialmente mediante la formación de espectros de consenso mediante los cuales se fusionan espectros idénticos que exhiben proporciones idénticas de masa a carga de iones precursores (m/z) y patrones de fragmentación (Frank et al. 2008). Los datos simplificados de MS/MS se utilizan luego para generar el análisis de redes moleculares (Figura 40). Las similitudes vectoriales se calculan para cada posible par de espectros con un mínimo de seis fragmentos de iones coincidentes (es decir, picos) con similitud determinada mediante el uso de un cálculo de coseno modificado que tiene en cuenta las intensidades relativas de los iones de fragmento, así como la diferencia del precursor m/z entre los espectros emparejados. Esto amplía el concepto de alineamiento espectral tal como se aplica en proteómica, con la excepción clave de que ciertos parámetros específicos de péptidos, como el uso de puntajes de probabilidad máxima, se generalizan para aplicar este enfoque a todas las clases de moléculas, incluidos lípidos, polisacáridos, péptidos, metabolitos y proteínas (Watrous et al. 2012). Esto es importante porque no se sabe de antemano qué clase de moléculas se ionizarán durante el análisis. Cuando esto se ha completado, los pares de espectros significativamente coincidentes se informan como una red molecular mediante el uso de scripts de MATLAB a través de los cuales cada espectro puede conectarse a sus coincidencias de más alta puntuación (normalmente se permiten un máximo de 10 conexiones por nodo). Los bordes entre espectros (conexiones) se retienen solo si el puntaje de similitud del vector, representado como un valor de coseno, de la coincidencia es mayor que el umbral definido por el usuario. Los valores de umbral de coseno generalmente se establecen entre 0,5 y 0,7, por lo que un valor de coseno de 1,0 indica espectros idénticos (Watrous et al. 2012) (Figura 40).

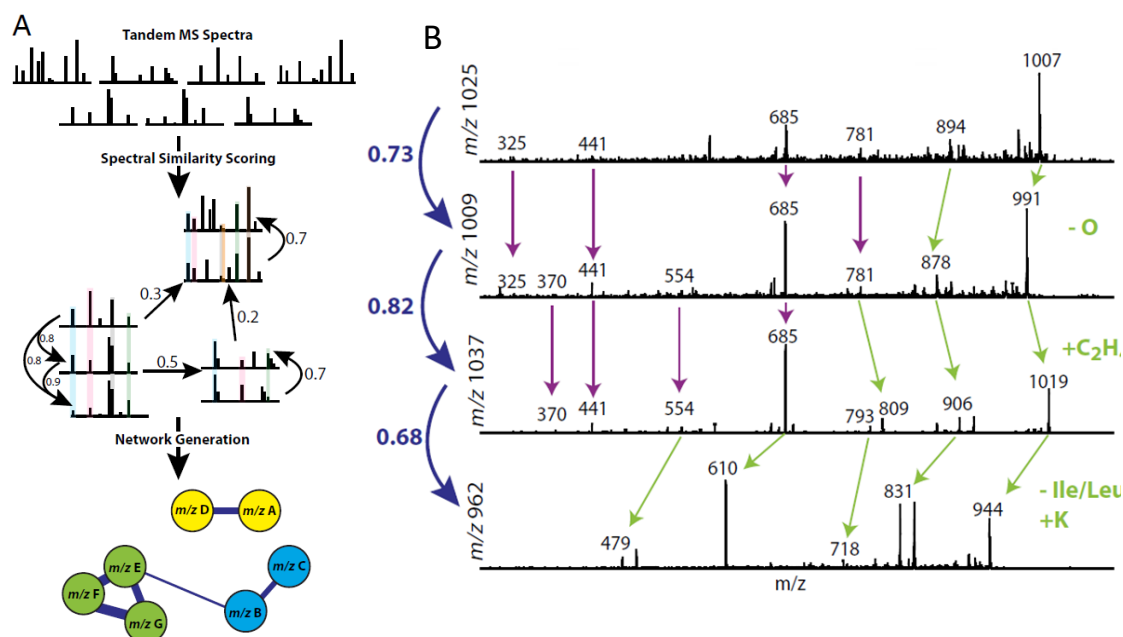


Figura 40. A) Representación esquemática de cómo se generan las redes moleculares. Los valores son representativos de las puntuaciones de coseno de 0 a 1, donde 1 indica espectros idénticos y 0 significa que no hay similitud alguna. B) Representación de cuatro espectros de fragmentación relacionados entre sí. Los valores de coseno (*score* de comparación) se encuentran en sobre las flechas azules a la izquierda de los espectros. A la derecha en color verde se encuentran las diferencias de masas entre los diferentes espectros. Modificado de (Watrous et al. 2012).

La visualización de redes moleculares en GNPS se lleva a cabo representando cada espectro como un nodo y los alineamientos entre espectros como bordes o líneas (conexiones) entre nodos. Los nodos se pueden complementar con metadatos, incluidas las coincidencias de desreplicación o la información proporcionada por el usuario, como la abundancia, el origen del producto, la actividad bioquímica o la hidrofobicidad, que se pueden reflejar en el tamaño o el color de un nodo. Es posible visualizar el mapa de moléculas relacionadas como una red molecular (Yang et al. 2013) en línea en GNPS o exportar para su análisis en softwares de terceros como el Cytoscape.

3.2.7 Feature base molecular networking

Dentro de la plataforma de GNPS es posible realizar un análisis alternativo denominado *Feature Base Molecular Networking* (FBMN). El FBMN aprovecha la capacidad ya bien establecida de los softwares de procesamiento de MS y mejora el *molecular networking* clásico al incorporar no solo información de MS¹, como patrones de isótopos y tiempo de retención, sino también la separación por movilidad de iones. Al confiar en la información espectral procesada, las redes moleculares obtenidas con FBMN pueden: 1) distinguir isómeros que producen espectros de MS² similares que se resuelven mediante separación cromatográfica o de movilidad iónica, que pueden haber permanecido ocultos en el *molecular networking* clásico; 2) Facilitar la anotación espectral; 3) Incorporar información cuantitativa relativa que permita un sólido análisis estadístico para análisis metabolómicos. Mientras que el *molecular networking* clásico usa el

conteo espectral o la suma de la intensidad de iones precursores para estimar la abundancia de iones, la FBMN usa las intensidades de iones precisas obtenidas por LC-MS. El método FBMN aporta abundancia de iones en todas las muestras utilizando el valor del área del pico cromatográfico o la altura del pico según lo determinado por el software de alineamiento y detección de características (*features*) de LC-MS (Wang et al. 2016). Al integrar ambos datos la precisión del cálculo de la abundancia de cada nodo aumenta significativamente, y es posible realizar comparaciones cuantitativas de alta fidelidad (Nothias et al. 2020). Durante el trabajo que se describe en este Capítulo, se realizó un FBNM con el solo objetivo de aumentar la precisión de la abundancia relativa de cada nodo. Para mayor detalle sobre la variedad de análisis disponibles referir a <https://gnps.ucsd.edu> (GNPS).

En el presente Capítulo nos propusimos aplicar un flujo de trabajo basado en espectrometría de masas en combinación con *Molecular Networking*, para la identificación de compuestos naturales producidos por aislamientos de pseudomonas con fenotipos atractivos desde el punto de vista del biocontrol.

3.3 Hipótesis de trabajo

La metodología propuesta por GNPS para la identificación de productos naturales por medio de *Molecular Networking*, es aplicable al análisis de precipitados ácidos de sobrenadantes de cultivos para la identificación de compuestos tipo lipopéptidos de aislamientos de *Pseudomonas*.

3.4 Objetivos específicos

- Identificar posibles compuestos tipo lipopéptidos producidos por aislamientos autóctonos de pseudomonas mediante espectrometría de masas en tándem acoplado al análisis por *Molecular Networking*.

3.5 Materiales y métodos

Las cepas bacterianas y oligonucleótidos empleados en este capítulo se listan en la Tabla 6. Las *Pseudomonas* se crecieron en NYB (caldo nutritivo, 25 g/L; extracto de levadura, 5 g/L) a 28°C y 200 rpm de forma rutinaria excepto que se detalle lo contrario. Las cepas se conservaron, en condiciones de esterilidad, en una solución acuosa de glicerol 20% final en ultrafreezer (-80°C).

Los cultivos de *Pseudomonas* destinados a la obtención de lipopéptidos fueron crecidos en medio PSFM (triptona 10 g/L, peptona 10 g/L, K₂HPO₄ 1,5 g/L, MgSO₄·7H₂O 1,5 g/L, glicerol 19 g/L). Los medios utilizados para evaluar la producción óptima de lipopéptidos fueron: King's B (peptona 20 g/L, K₂HPO₄ 1,5 g/L, MgSO₄·7H₂O 1,5 g/L, glicerol 7 g/L), DBM glucosa ((NH₄)₂SO₄ 1 g/L, Na₂HPO₄ 8.83 g/L, MgSO₄ 0.1 g/L, KH₂PO₄ 2 g/L, NaH₂PO₄ 1.15 g/L, glucosa 0.5 g/L), DBM manitol ((NH₄)₂SO₄ 1 g/L, Na₂HPO₄ 8.83 g/L, MgSO₄ 0.1 g/L, KH₂PO₄ 2 g/L, NaH₂PO₄ 1.15 g/L, manitol 0.5 g/L).

Tabla 6. Cepas bacterianas y oligonucleótidos utilizados en este capítulo

Aislamientos de <i>Pseudomonas</i>	Descripción	Referencia
SMMP3	Aislamiento autóctono de <i>Pseudomonas</i>	(Agaras et al. 2015)
RBAN4	Aislamiento autóctono de <i>Pseudomonas</i>	(Agaras et al. 2015)
RMAN4	Aislamiento autóctono de <i>Pseudomonas</i>	(Agaras et al. 2015)
SVMP4	Aislamiento autóctono de <i>Pseudomonas</i>	(Agaras et al. 2015)
SBMP6	Aislamiento autóctono de <i>Pseudomonas</i>	(Agaras et al. 2015)
SVBP6	Aislamiento autóctono de <i>Pseudomonas</i>	(Agaras et al. 2015)
Oligonucleótidos		
P0	GAGAGTTTGATCCTGGCTCAG Gen blanco: ARNr 16S	(Picard et al. 2000)
P6	CTACGGCTACCTGTTACGA Gen blanco: ARNr 16S	(Picard et al. 2000)
<i>rpoB</i> -F	TGGCCGAGAACCAGTTCCGCGT	(Tayeb et al. 2005)
<i>rpoB</i> -R	CGGCTTCGTCCAGCTTGTTTCAG	(Tayeb et al. 2005)
<i>gyrB</i> -F	AGCATYAARGTGCTGAARGG	(Agaras, Valverde, and Protocols 2018)
<i>gyrB</i> -R	GGTCATGATGATGATGTTGTG	(Agaras, Valverde et al. 2018)

3.5.1 PCR y secuenciación

Los genes 16s, *rpoB* y *gyrB* fueron amplificados por PCR utilizando los oligonucleótidos detallados en la Tabla 6. El perfil de ciclado fue idéntico para los tres genes a excepción de la temperatura de *annealing* y consistió en un paso de desnaturalización inicial de 5 min a 95°C, luego 35 ciclos de 30 s a 95°C, 30 s a 53°C para 16s; 50°C para *rpoB*; 57°C para *gyrB*, 1 min a 72°C y una extensión final de 10 a 72°C. Como molde de ADN, se utilizó 1 µL de un lisado celular térmico de cada cepa. Los lisados térmicos se obtuvieron tratando una colonia de 2–3 mm resuspendida en 100 µL de H₂O ultrapura (grado PCR) a 100 °C durante 10 min. Después de la amplificación, se corrieron 10 µL de cada reacción en geles de agarosa al 1% en Tris-borato-EDTA (TBE) 0,5x a 10 V/cm durante 50 min. El gel se tiñó con bromuro de etidio y se visualizó bajo luz

ultravioleta. Los productos de PCR se secuenciaron en Macrogen Inc. (Seúl, Corea) desde ambos extremos con los oligonucleótidos correspondientes (Tabla 6). La reconstrucción de la secuencia de cada uno de los amplicones y el control de la calidad de las secuencias se realizaron manualmente con la herramienta BlastN (Altschul et al. 1990).

3.5.2 Análisis multi locus de secuencia (MLSA)

Los alineamientos y los árboles filogenéticos se realizaron con el software MegaX (Kumar et al. 2018). Para los alineamientos múltiples se utilizó el método MUSCLE, y para la confección del árbol filogenético se infirió la historia evolutiva utilizando el método de Máxima Verosimilitud (*Maximun Likelihood*) y el modelo Tamura-Nei (Tamura, Nei, and evolution 1993). Los árboles iniciales para la búsqueda heurística se obtuvieron automáticamente aplicando los algoritmos Neighbor-Joining y BioNJ a una matriz de distancias por pares estimadas mediante el modelo Tamura-Nei y luego seleccionando la topología con un valor logarítmico de verosimilitud superior. Se seleccionó el método de Bootstrap para el test de filogenia con 500 repeticiones.

3.5.3 Antagonismo in vitro contra *Pythium ultimum*

Para evaluar el potencial antagonismo de las fracciones colectadas por HPLC, se las enfrentó con el aislamiento Pu67-1 del oomicete fitopatógeno *P. ultimum* (Maurhofer et al. 1994b). Se creció el oomicete en una placa de Petri con medio agar papa dextrosa (PDA, Laboratorios Britania) por 72h a 28°C, o hasta obtener un crecimiento que cubriera toda la placa. Luego, se extrajo una porción de agar de 0.5 cm², y se la depositó en el centro de un portaobjetos de vidrio recubierto en su superficie con PDA agarizado. En un extremo se depositó una gota de 20 µl de muestra disuelta en metanol, y en el otro extremo una gota de metanol como control. Los portaobjetos fueron guardados en cajas de plástico esterilizadas, selladas con Parafilm® y se incubaron a 28°C hasta que *P. ultimum* logró crecer sobre las gotas de metanol.

3.5.4 Molecular Networking

Los precipitados ácidos de sobrenadante fueron obtenidos al igual que para *P. protegens* CHAO siguiendo la metodología planteada en el capítulo 1. Luego fueron analizados por UHPLC acoplado a un espectrómetro de masas MS/MS según el protocolo propuesto por Petras et al. (2017) en el Centro de Innovación sobre Microbiomas, Universidad de San Diego, California, Estados Unidos.

Se creó una red molecular utilizando el flujo de trabajo según (<https://ccms-ucsd.github.io/GNPSDocumentation/>) en el sitio web de GNPS (<http://gnps.ucsd.edu>). La tolerancia de masa de iones precursores se fijó en 0,01 Da y una tolerancia de iones de fragmentos MS/MS de 0,01 Da. Luego se creó una red donde los bordes se filtraron para tener una puntuación de coseno superior a 0,7 y más de 6 picos coincidentes. Además, los bordes entre dos nodos se mantuvieron en la red sí y solo sí cada uno de los nodos aparecía en los 10 nodos más similares respectivos del otro. Finalmente, el tamaño máximo de una familia molecular se fijó en 100, y los bordes de puntuación más bajos se eliminaron de las familias

moleculares hasta que el tamaño de la familia molecular estuvo por debajo de este umbral. Luego se buscaron los espectros en la red contra las bibliotecas espectrales de GNPS. Los espectros de la biblioteca se filtraron de la misma manera que los datos de entrada. Todas las coincidencias mantenidas entre los espectros de la red y los espectros de la biblioteca debían tener una puntuación superior a 0,7 y al menos 6 picos coincidentes.

3.6 Resultados y discusión

Nuestro laboratorio contaba con una colección de aislamientos autóctonos de *pseudomonas* atractivos por sus fenotipos biocontroladores (Agaras et al. 2015), los cuales en ocasiones están asociados a la producción de lipopéptidos (Cawoy et al. 2015), por lo tanto, se decidió evaluar su presencia por el método del colapso de gota, como se detalla en la metodología del Capítulo 1 (Figura 41). Todos los cultivos fueron crecidos en el medio PSFM, ya que se había visto que *P. protegens* CHA0 producía significativamente mayor cantidad de lipopéptidos. Los aislamientos RBAN4, RMAN4, SVMP4, SBMP6 Y SMMP3 fueron positivos para el ensayo (Figura 41) y se decidió adicionarlos paralelamente al flujo de trabajo de identificación de lipopéptidos de CHA0 descrito en el capítulo 1. El aislamiento SVBP6 no fue positivo en el colapso de gota por lo que fue descartado para los posteriores análisis.

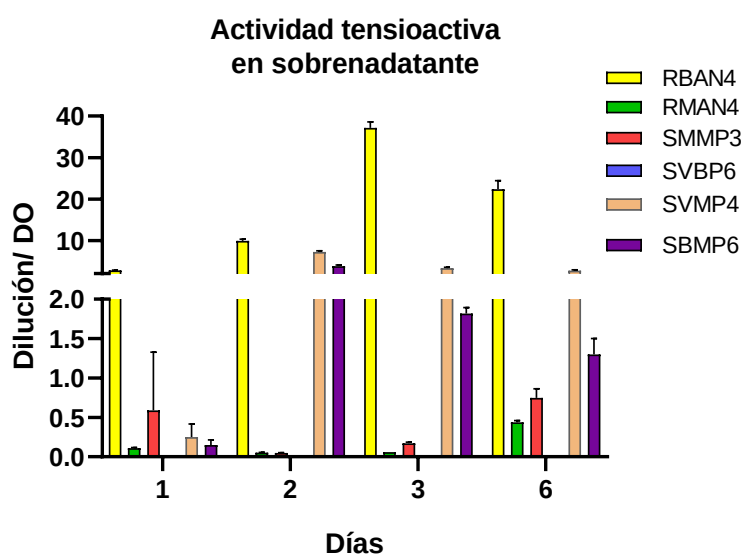


Figura 41. Actividad tensioactiva en sobrenadante de cultivo de diferentes aislamientos de *pseudomonas* autóctonos de la región agrícola núcleo argentina.

3.6.1 Extracción e identificación de compuestos surfactantes en aislamientos autóctonos de *pseudomonas*

Se crecieron cultivos de los aislamientos de *pseudomonas* que dieron positivo al ensayo del colapso de gota: RBAN4, RMAN4, SVMP4, SBMP6 Y SMMP3. Se realizó una precipitación ácida de los sobrenadantes de cultivo y se analizaron por HPLC del mismo modo que se realizó para CHA0 en el capítulo 1 de este trabajo.

3.6.1.1 RMAN4

El aislamiento RMAN4 no presentó señales relevantes en su cromatograma (Figura 42). Esta cepa es capaz de producir compuestos tensioactivos en su sobrenadante de cultivo (Figura 41), pero es posible que su naturaleza química sea diferente a la de un compuesto tipo lipopéptido. Dichos compuestos no son extraídos por el método de precipitación ácida, por lo tanto, el extracto analizado por HPLC presentó un cromatograma “vacío”.

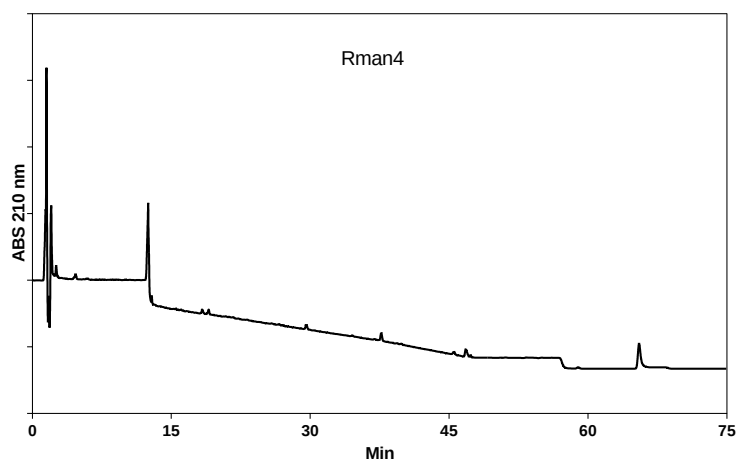


Figura 42. Perfil de extractos de lipopéptidos por HPLC del aislamiento RMAN4

3.6.1.2 RBAN4

El aislamiento RBAN4 exhibió un cromatograma idéntico al de CHA0. En la Figura 43 se puede ver que coinciden todos los picos en ambos cromatogramas. Este resultado generó la sospecha de que el aislamiento podría pertenecer a la especie *P. protegens*. Se aisló el pico mayoritario y se obtuvo el peso del ion molecular por espectrometría de masas. Este coincidió con la muestra correspondiente a la orfamida A purificada de CHA0. Teniendo en cuenta que los cromatogramas, tiempos de retención y las masas eran idénticas, asumimos que se trataba también de orfamida A. Por otro lado, se realizó un estudio genético para intentar identificar a RBAN4 a nivel de especie a partir de un análisis de secuencias multi locus (MLSA). Para esta cepa se eligieron los genes ADNr 16s, *rpoB* y *gyrB*. La elección de dichos genes se basó en bibliografía, aunque no se encontraron criterios claros o unificados sobre qué genes poseen el mayor poder de discriminación a nivel de especie para identificación de pseudomonas. Para el análisis genético del aislamiento RBAN4 se continuó con los genes previamente mencionados, pero el interrogante planteado a partir de esta búsqueda bibliográfica derivó en el trabajo que se detalla en el capítulo 4 de este trabajo.

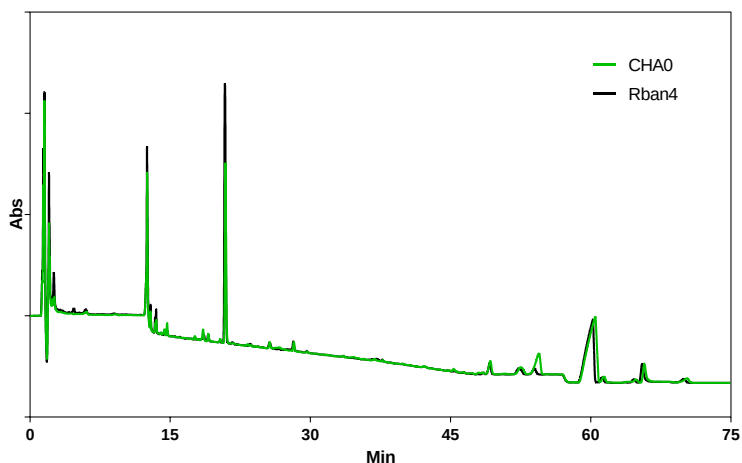


Figura 43. Perfil de extracto de lipopéptidos por HPLC de la cepa RBAN4 (negro). Se agregó el cromatograma de la cepa *P. protegens* CHA0 (verde) con fines comparativos.

Los genes fueron amplificados por PCR y secuenciados, luego se construyó un concatenado y se generó un árbol filogenético seleccionando algunas cepas representativas del género *Pseudomonas* (

Figura 44). Se puede apreciar que RBAN4 se agrupa dentro del grupo de cepas pertenecientes a la especie *P. protegens*, siendo CHA0 la cepa más cercana (que es además, la cepa tipo de la especie *P. protegens* (Ramette et al. 2011)). Teniendo en cuenta esto, y el perfil de producción de lipopéptidos podemos afirmar que el aislamiento se trata de una representante de *Pseudomonas protegens*.

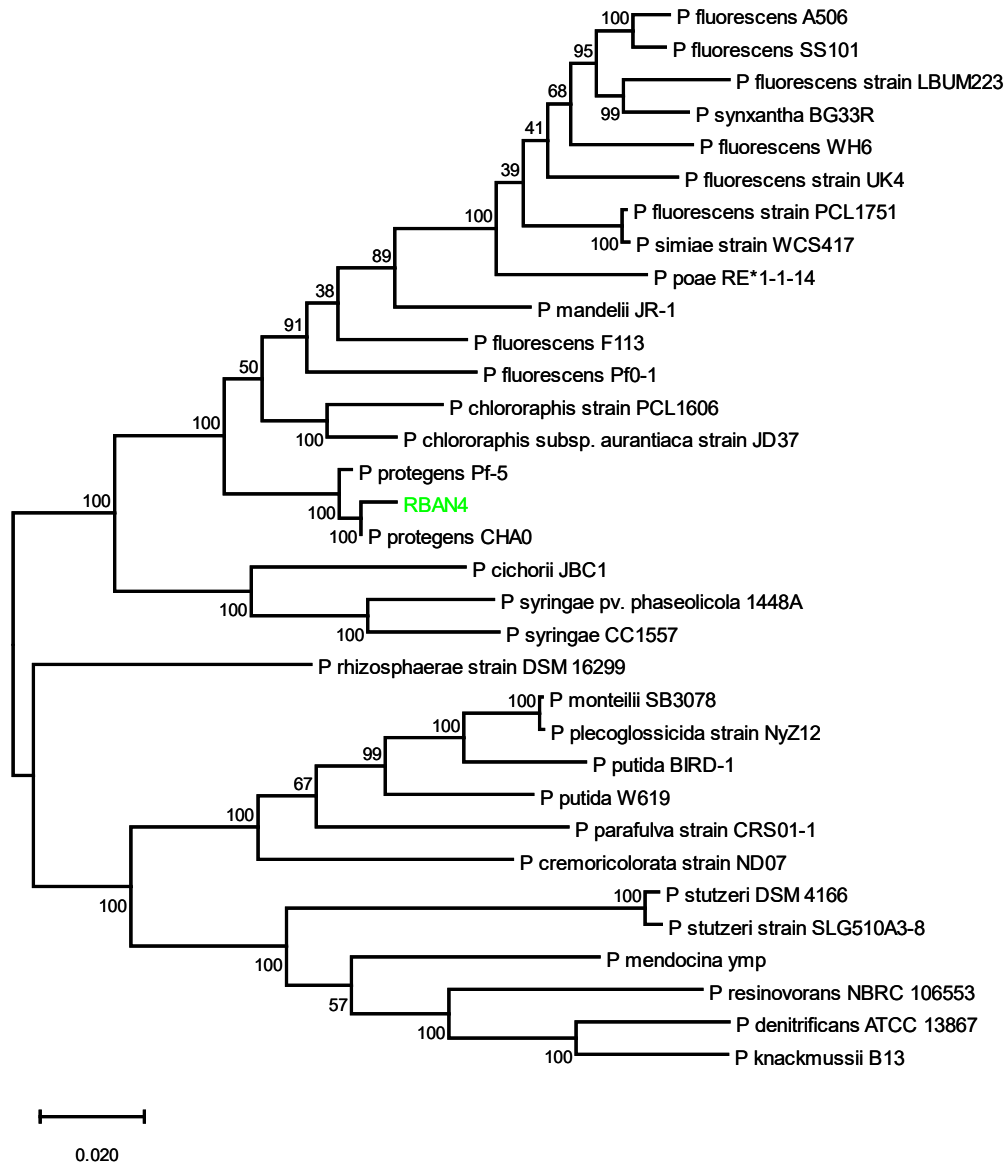


Figura 44. Árbol filogenético del concatenado 16s-*rpoB-gyrB*. Las distancias evolutivas fueron calculadas a partir del método *maximun likelihood*. El árbol está dibujado a escala, con las longitudes de las ramas medidas en número de sustituciones por sitio. Este análisis involucró 33 secuencias de nucleótidos. Hubo un total de 8065 posiciones en el conjunto de datos final.

3.6.1.3 SMMP3

El aislamiento SMMP3 fue estudiado y caracterizado dentro de nuestro laboratorio. A partir de distintos estudios realizados se determinó que pertenece a la especie *chlororaphis* subsp. *aurantiaca*, y que posee los genes necesarios para la producción de fenazinas, compuestos que se encuentran relacionados con el biocontrol (Chin-A-Woeng, Bloemberg, and Lugtenberg 2003). Además, en un ensayo preliminar no publicado hasta el momento, se observó que los

mutantes deficientes en la producción de fenazina pierden gran parte de su capacidad de antagonizar *P. ultimum* *in vitro*, sin embargo, el efecto no se pierde por completo, como sí lo hacen mutantes deficientes para activar la cascada Gac-Rsm. En la Figura 45 se observan los resultados del cálculo del halo de antagonismo de la cepa SMMP3 y dos mutantes sobre *P. ultimum*. Se puede observar que en SMMP3 Fen⁻, mutante impedido en la biosíntesis de fenazinas, el antagonismo se reduce significativamente pero no desaparece por completo. Por el contrario, la variante SMMP3 Gac, que es incapaz de activar todos sus genes regulados por la cascada Gac-Rsm perdió completamente la capacidad de antagonizar a *P. ultimum*.

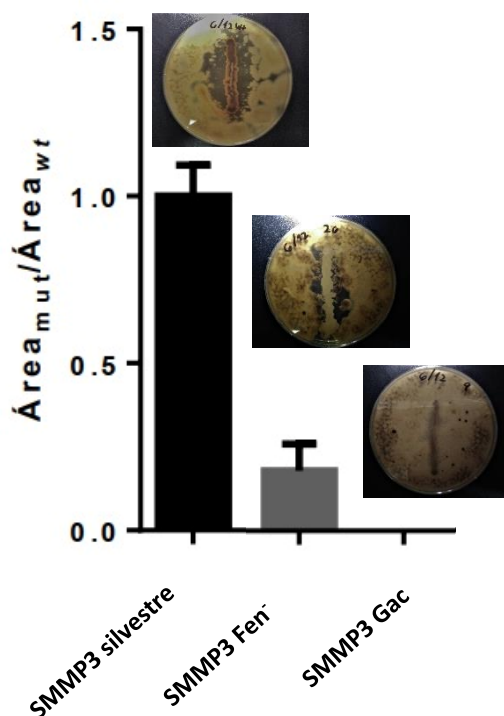


Figura 45. Antagonismo *in vitro* de la cepa SMMP3 contra *P. ultimum*. La cepa SMMP3 Fen⁻ proviene de una biblioteca de mutantes aleatorios (Tn5) y es incapaz de sintetizar fenazina. La cepa SMMP3 Gac es un mutante de la misma biblioteca y es incapaz de activar la cascada Gac-Rsm. Figura tomada del seminario de grado del Lic. Federico Muzio

Teniendo en cuenta toda la evidencia mencionada se decidió estudiar el perfil de compuestos tipo lipopéptidos, con dos objetivos principales: 1) Confirmar la presencia de fenazina, e identificarla dentro de su grupo químico. 2) Identificar otros posibles compuestos con actividad biocontroladora, responsable del antagonismo observado por cepas no productoras de fenazina. En la Figura 46, se puede observar el cromatograma de HPLC obtenido para el precipitado de sobrenadante de cultivo de la cepa SMMP3. Las flechas azules corresponden a las fracciones colectadas. Se utilizó una porción de cada fracción para evaluar su potencial antagonismo frente al oomicete *Pythium ultimum*. El ensayo consistió en enfrentar un taco de agar con *P. ultimum* en el centro de un porta objetos de vidrio, recubierto con PDA agarizado, contra una gota de cada fracción resuspendida en metanol (Figura 47). En el extremo contrario a la muestra, se depositó una gota de metanol, para controlar la posible inhibición del crecimiento por el solvente.

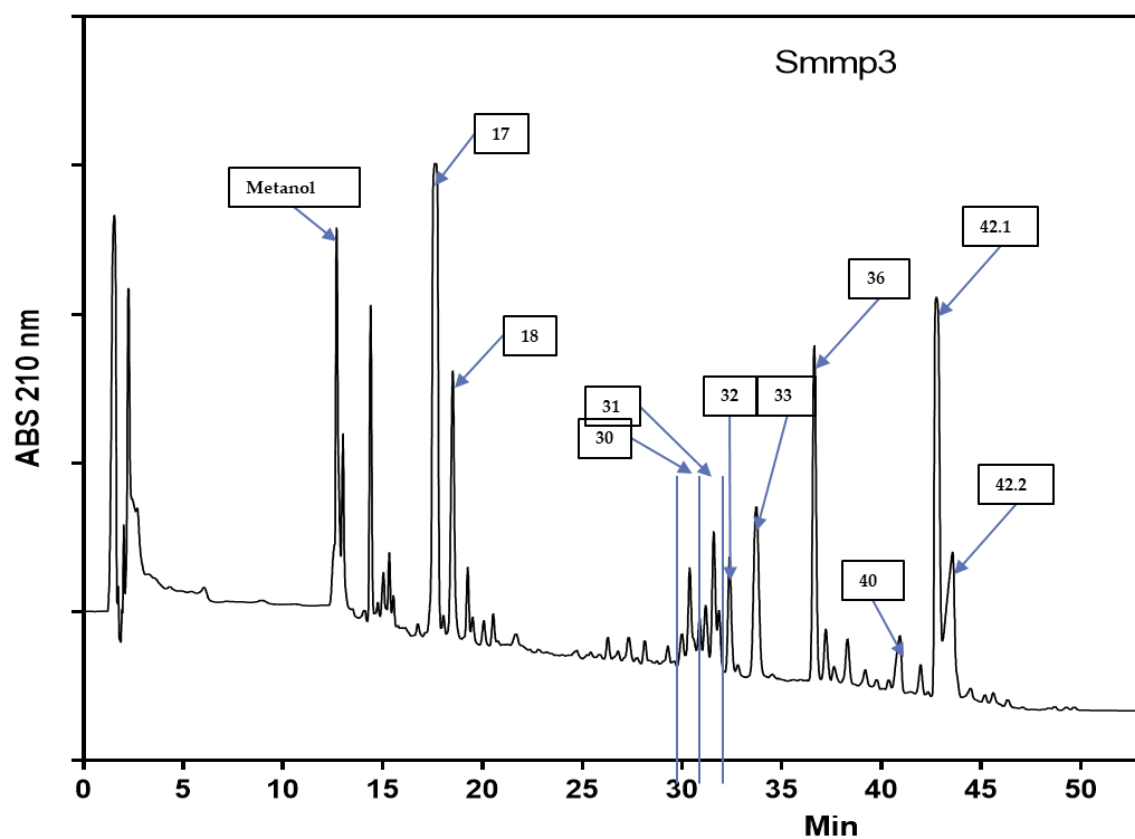


Figura 46. Perfil de extracto de lipopéptidos por HPLC de la cepa SMMP3. Cada flecha indica una fracción colectada, y el numero en el recuadro indica el minuto en el que se comenzó a colectar cada fracción. El perfil de elución fue el mismo que para los otros ensayos, pero se eliminaron los minutos finales porque no presentaron señales.

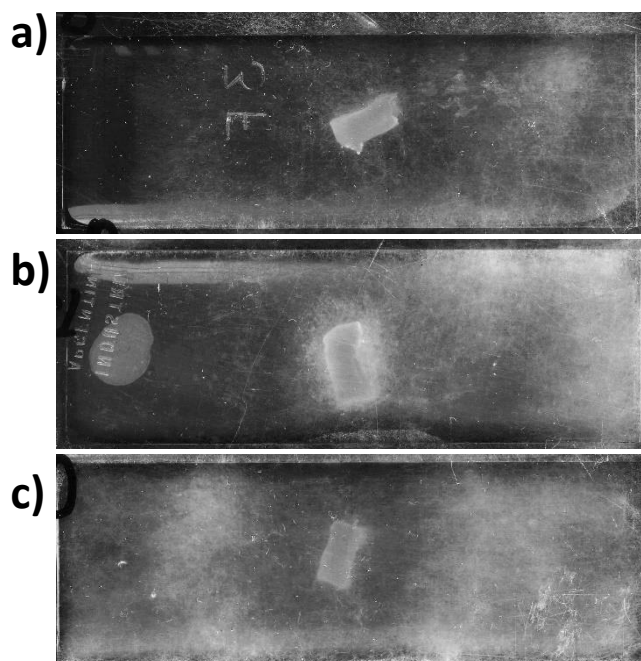


Figura 47. Antagonismo entre el oomicete *P. ultimum* y las fracciones colectadas por HPLC del medio de cultivo de la cepa *P. chlororaphis* subsp. *aurantiaca* SMMP3. Todos los portaobjetos recubiertos de PDA agarizado fueron sembrados con una muestra de 20 μ l en su extremo izquierdo y con 20 μ l de metanol en su extremo derecho. a) Fracción colectada al minuto 18. b) Precipitado previo a la separación por HPLC. c) El resto de las fracciones no fueron capaz de antagonizar al oomicete y se expone la fracción colectada al minuto 30 como representante del resto.

La única fracción que fue capaz de antagonizar el crecimiento de *P. ultimum* fue la colectada en el minuto 18, mientras que en el resto, el oomicete fue capaz de crecer sobre la muestra al mismo tiempo que sobre la gota de metanol. Una vez identificada la fracción con actividad inhibitoria, nos propusimos identificar el compuesto mediante espectrometría de masas MALDI-TOF, sin embargo, no obtuvimos resultados concluyentes por dos posibles problemas: 1) La masa de los compuestos no era suficiente. 2) La pureza de las muestras era muy baja. La identificación de compuestos naturales propone un desafío significativo, ya que no existe un camino o flujo de trabajo único para identificar un metabolito desconocido. Es por esto que nos propusimos modificar el enfoque con el cual veníamos trabajando, y analizar los extractos por espectrometría de masas en tándem, seguido del estudio enfocado en redes moleculares (*Molecular Networking*).

Los extractos crudos de precipitación acida de sobrenadante de cultivo fueron analizados por UHPLC acoplada a espectrometría de masas MS-MS, según el protocolo propuesto por la plataforma GNPS (*Global Natural Products Social molecular networking*) (Wang et al. 2016). Los datos obtenidos se componen de tres sets de datos complementarios: 1) Los cromatogramas obtenidos por UHPLC. 2) Los espectros de masas MS₁. 3) Los espectros de masas MS₂. La combinación de todos los datos nos da como resultado una red que integra todas las moléculas detectadas de todas las muestras, relacionadas entre sí según la similitud de sus espectros de masas MS₁ y MS₂. Además, podemos calcular la abundancia relativa de cada compuesto, según

el área bajo la curva del primer set de datos. Otra de las grandes ventajas de la plataforma GNPS, es que todos los espectros de masa obtenidos son comparados con diferentes bases de datos de compuestos naturales. En la Figura 48, se puede observar la red molecular completa generada a partir de la combinación de 4 muestras: Extracto de precipitado de sobrenadante de las cepas SMMP3, SVMP4, SBMP6 y blanco con metanol como control. En la parte inferior de la red se encuentran los nodos no relacionados, por el contrario, en la parte superior se encuentran las redes de mayor tamaño. Los nodos con aporte del blanco (rojos) no fueron analizados ya que se asume que la procedencia de sus compuestos no es específica de las muestras.

Una vez generada la red, se procedió a la búsqueda de fenazinas ya que estudios previos sugerían que la cepa SMMP3 era capaz de producir estos compuestos. Uno de los nodos provenientes de esta muestra presento alta similitud con la 1-carboxy-fenazina y este nodo era parte de una pequeña red de 5 nodos, la cual se analizó por separado en la Figura 49. Según la abundancia relativa de todos los nodos, SMMP3 produce principalmente 1-carboxy-fenazina, además se generaron predicciones de las posibles estructuras de los demás nodos en función de las diferencias de masas. La metodología empleada según la plataforma GNPS para *Molecular Networking* resultó de gran utilidad para el estudio del precipitado ácido de sobrenadante de cultivo de la cepa SMMP3. Se logró determinar que SMMP3 produce 1-carboxy-fenazina, y otros posibles derivados como la 1,6-dicarboxy-fenazina o la 2-hidroxí-1carboxi-fenazina, ambos reportados (Bilal et al. 2017), sin embargo para la confirmación de posibles derivados es necesario realizar estudios complementarios.

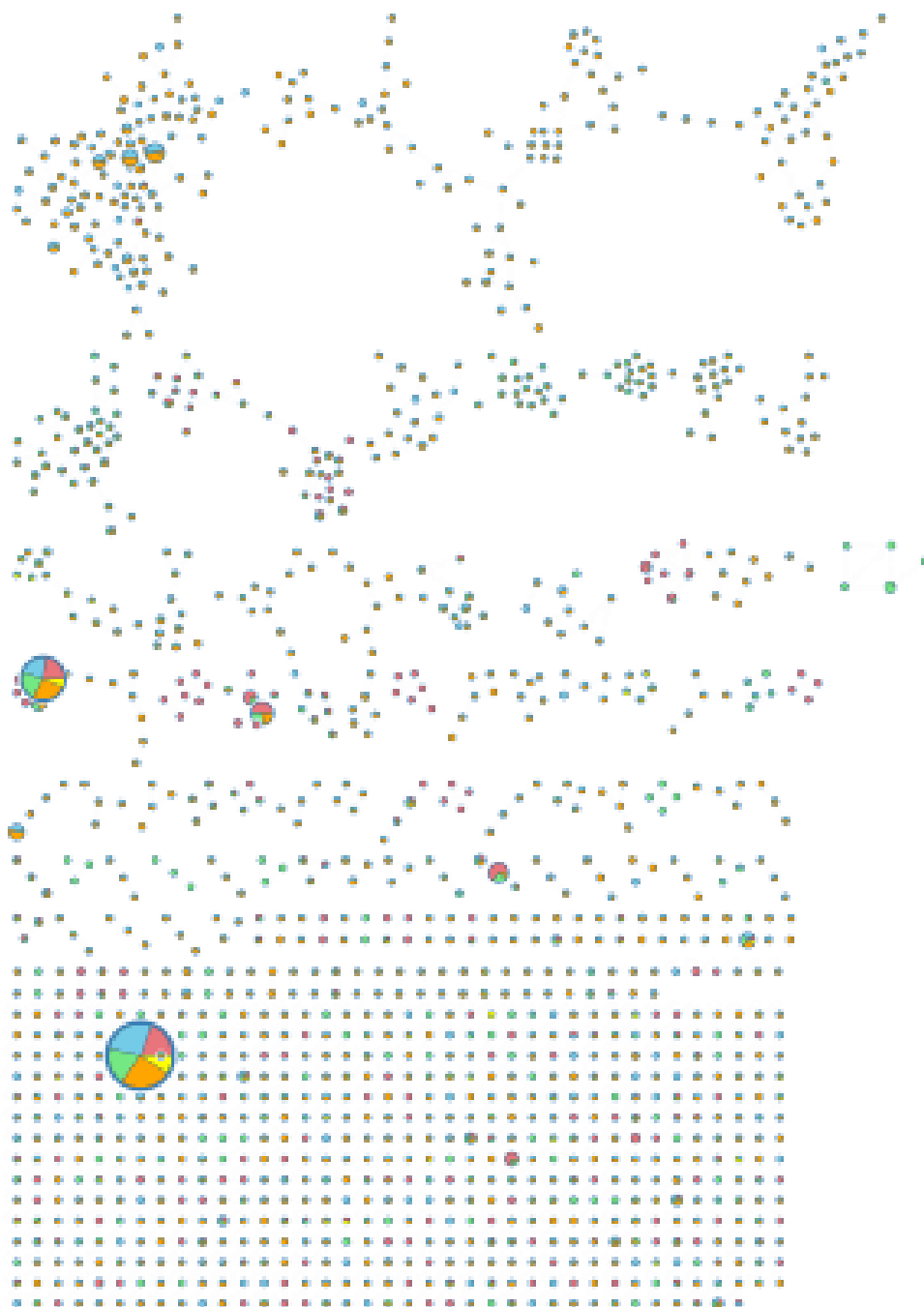


Figura 48. Red molecular completa de los precipitados ácidos del sobrenadante de las cepas SMMP3 (verde), SVMP4 (naranja), SBMP6 (azul) y un blanco conteniendo metanol puro (Rojo) generada con el software Cytoescape. El tamaño de los nodos (círculos de colores) se ajustó para representar la abundancia relativa de cada compuesto. Los colores internos de los nodos representan la procedencia de la muestra y el área de cada color corresponde a la abundancia relativa aportada por cada muestra.

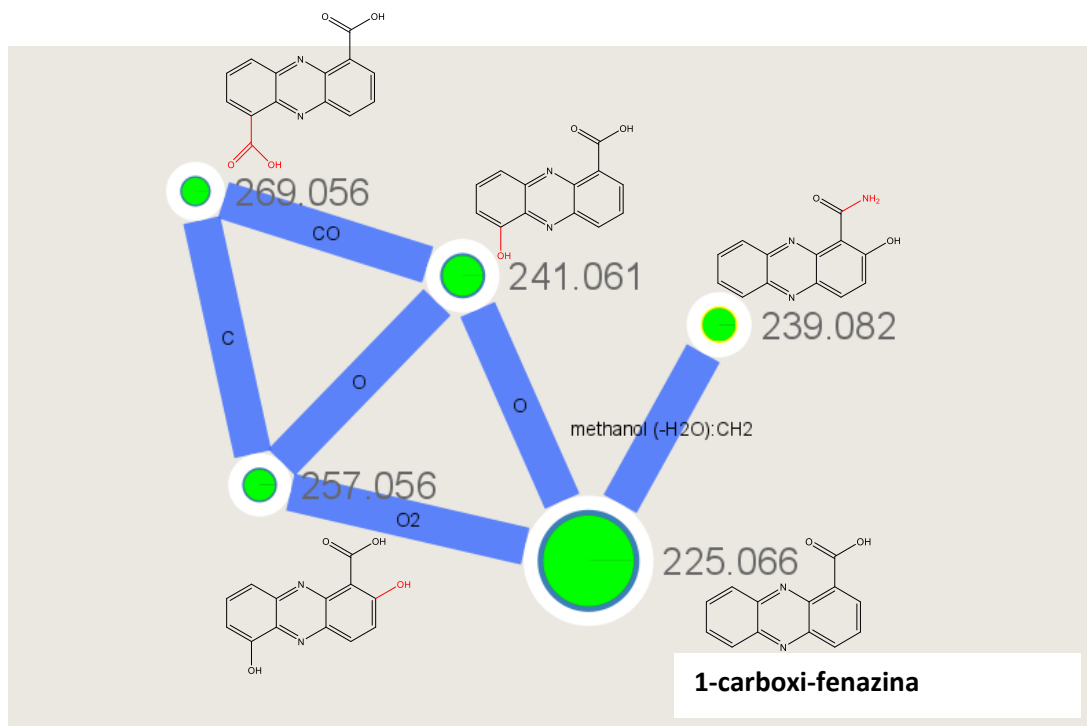


Figura 49. Red molecular de la 1-carboxi-fenazina. El tamaño de cada nodo representa la abundancia relativa. Cada nodo tiene a su derecha el valor de masa de su ion precursor. En las uniones entre los nodos se predice el átomo/molécula más probable para explicar la diferencia de masa. Se agregaron representaciones de las moléculas para predecir sus posibles estructuras químicas.

3.6.1.4 SVMP4 y SBMP6

Los aislamientos SVMP4 y SBMP6 presentaron cromatogramas idénticos entre sí, por lo tanto, se los analizó en conjunto. Se llevó a cabo el mismo flujo de trabajo que para la cepa SMMP3: en primer lugar, se colectaron las fracciones principales de su cromatograma (Figura 50), y luego se realizaron ensayos de antagonismo *in vitro* frente a *P. ultimum* con las fracciones de SVMP4 (Figura 51). El experimento no se repitió para la cepa SBMP6 ya que se consideró que se trata de los mismos compuestos, además en ninguna fracción pudimos detectar efecto antagonista contra *P. ultimum*. Igualmente procedimos a realizar análisis de masas basados en redes moleculares, con el objetivo de explorar el perfil de lipopéptidos purificado de ambas cepas.

A diferencia de la cepa SMMP3, no disponíamos de información sobre qué compuestos podrían producir las cepas SVMP4 y SBMP6, sin embargo, fue posible detectar pseudopironina B por una comparación de espectros con una de las bases de datos de GNPS. Este compuesto es un potente antibiótico (Singh et al. 2003), alguicida (Cho and biochemistry 2012) y tóxico para ciertos protozoarios (Giddens et al. 2008). Se elaboró una pequeña red conteniendo el nodo correspondiente a la pseudopironina B, para visualizar su entorno químico (Figura 52). Ambas cepas parecen producir mayoritariamente pseudopironina B, aunque también existe un nodo que coincide con la masa de la pseudopironina A, la cual probablemente se encuentre presente en el medio. Es habitual encontrar ambos compuestos juntos, principalmente porque la

pseudopironina A es un intermediario de la ruta de síntesis de la pseudopironina B. Estas moléculas se sintetizan a partir de la adición de unidades de dos carbonos (Bauer et al. 2015), que es exactamente la diferencia de masas entre una y otra (C_2H_4). Es posible que el resto de los nodos sean intermediarios también de su biosíntesis. Existe un nodo relacionado de ácido ricinoleico el cual proviene de la muestra de la cepa SMMP3 el cual fue agrupado por similitud en los espectros de masas, pero no por un vínculo biológico.

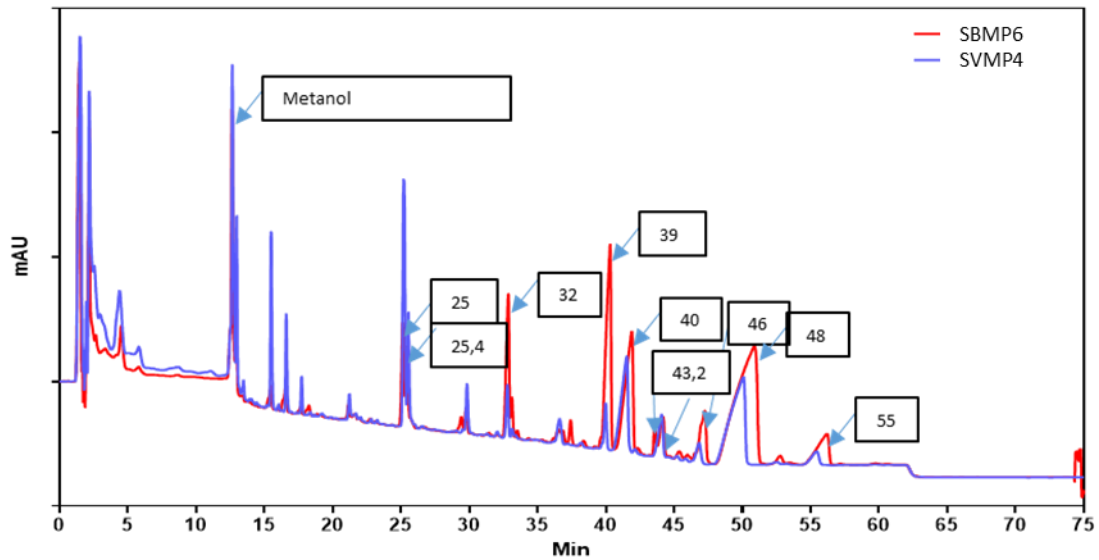


Figura 50. Perfil de extracto de lipopéptidos por HPLC de las cepas SVMP4 y SBMP6. Cada flecha indica una fracción colectada, y el numero en el recuadro indica el minuto en el que se comenzó a colectar cada fracción.

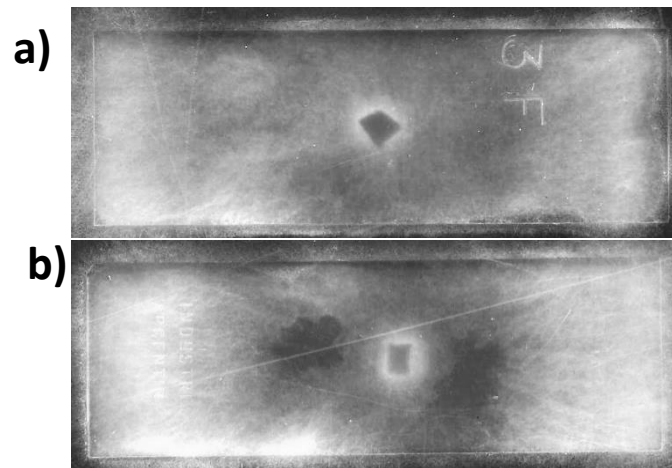


Figura 51. Antagonismo entre *P. ultimum* y las fracciones colectadas por HPLC del medio de cultivo de la cepa SVMP4. Todos los portaobjetos recubiertos de PDA agarizado fueron sembrados con una muestra de 20µl en su extremo izquierdo y con 20µl de metanol en su extremo derecho. a) Precipitado previo a la separación por HPLC. b) El resto de las fracciones no fueron capaz de antagonizar al oomicete y se expone la fracción colectada al minuto 48 como representante del resto.

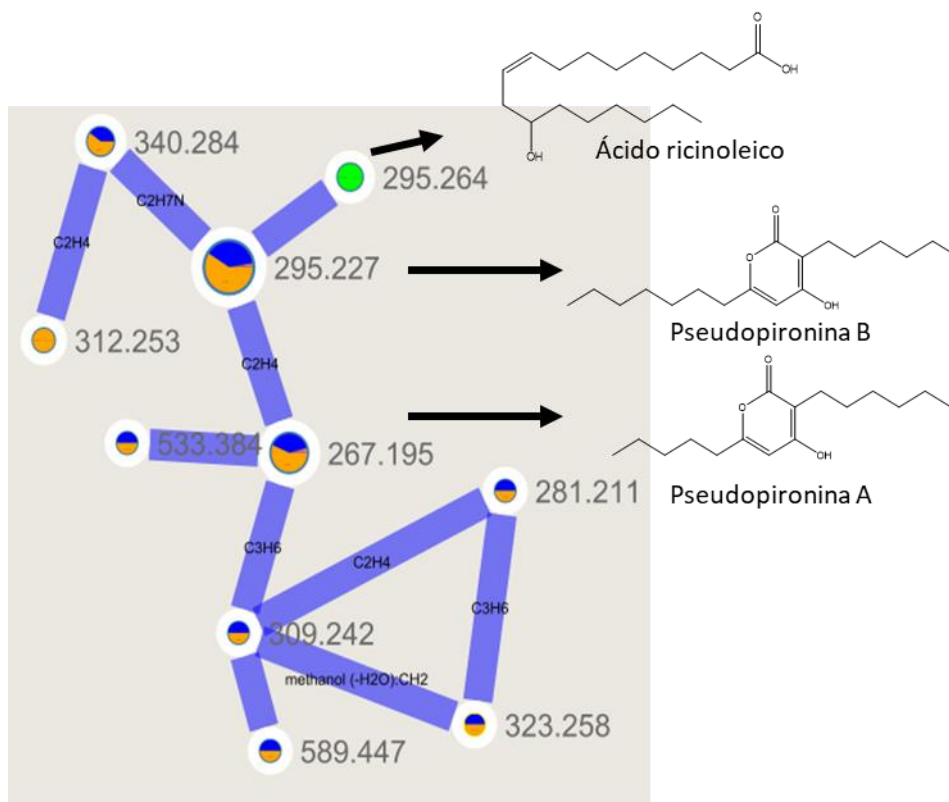


Figura 52 Red molecular de la pseudopironina B. El tamaño de cada nodo representa la abundancia relativa. Cada nodo tiene a su derecha el valor de masa de su ion precursor. En las uniones entre los nodos se predice el átomo/molécula más probable para explicar la diferencia de masa. Se agregaron representaciones de las moléculas para predecir sus posibles estructuras químicas. El color de cada nodo representa la abundancia relativa aportada por cada muestra: SVMP4 en azul, SBMP6 en naranja y SMMP3 en verde.

3.7 Discusión

En este capítulo se reporta la caracterización de productos metabólicos especializados de 6 aislamientos autóctonos de *Pseudomonas* de una colección de nuestro laboratorio, proveniente de suelo agrícola o rizosfera de plantas de cultivos de la zona núcleo del país (Agaras, Wall, and Valverde 2012). Del aislamiento RMAN4 no pudimos obtener demasiada información, excepto que la naturaleza química del compuesto responsable de su actividad surfactante en el medio de cultivo, no pertenece a la familia de los lipopéptidos. Se logró posicionar al aislamiento RBAN4 dentro del grupo de *Pseudomonas protegens*, y se confirmó la producción de orfamida. Los aislamientos SVMP4 y SBMP6 mostraron similitudes significativas en su perfil de producción de lipopéptidos, por lo que probablemente pertenezcan a la misma especie. Precisamente, la afiliación taxonómica de estas dos cepas dentro del género *Pseudomonas*, se discutirá en el siguiente Capítulo. Por otro lado, la espectrometría de masas en tándem en combinación con el análisis propuesto por la plataforma GNPS, resultó de gran utilidad para la identificación de varios compuestos presentes en sobrenadantes de cultivo. Fue posible detectar pseudopironina

A y B en el sobrenadante de SVMP4 y SBMP6, el cual probablemente sea el responsable tanto de la actividad surfactante en su sobrenadante, como de algunos de sus rasgos biocontroladores (Agaras et al. 2015). De la misma manera, se logró corroborar la presencia de varias fenazinas en el sobrenadante de SMMP3, la 1-carboxy-fenazina entre ellas. Esta metodología presenta varias ventajas para la identificación de compuestos naturales, principalmente la de evitar pasos de purificación exhaustiva sin perder profundidad en el análisis. Desafortunadamente para lograr una rápida identificación dependemos de que exista un espectro de masas similar en una base de datos, sin embargo, el mismo uso de la plataforma sostenido en tiempo debería resultar en el aumento de espectros disponibles hasta alcanzar un sistema comparable al BLAST para ADN.

3.8 Conclusión

La identificación de productos naturales es una tarea compleja que requiere la combinación de diversas disciplinas, y en general, el diseño un flujo de trabajo *ad hoc* para cada caso de estudio. Particularmente para nuestro laboratorio, la identificación de metabolitos bacterianos es sumamente importante, ya sea que estudiemos interacciones involucradas en la regulación, interacciones entre microorganismos, con plantas, etc. La identificación molecular de los compuestos involucrados en dichas interacciones es uno de los pilares en los cuales debemos apoyar nuestros resultados, sin embargo, la elevada dificultad o la falta de experiencia en ciertas disciplinas ha llevado a retrasar investigaciones o incluso tener que modificar el curso de las mismas. En este sentido, los análisis basados en *Molecular Networking* no solo proporcionan una alternativa a la identificación de compuestos naturales, sino que además permite realizar estudios de metabolómicas completos utilizando las mismas muestras. En este capítulo se analizaron los precipitados ácidos de diferentes aislamientos de *Pseudomonas* en busca de lipopéptidos, se construyeron diversas redes moleculares y se lograron identificar algunos metabolitos secundarios. Ante estos resultados, creemos que esta metodología podría presentarse como una solución a numerosas investigaciones dentro de nuestro grupo de trabajo.

3.9 Bibliografía

- Agaras, B., et al. (2012). "Specific enumeration and analysis of the community structure of culturable pseudomonads in agricultural soils under no-till management in Argentina." 61: 305-319.
- Agaras, B. C., et al. (2015). "Quantification of the potential biocontrol and direct plant growth promotion abilities based on multiple biological traits distinguish different groups of *Pseudomonas* spp. isolates." 90: 173-186.
- Agaras, B. C., et al. (2018). "A novel oligonucleotide pair for genotyping members of the *Pseudomonas* genus by single-round PCR amplification of the *gyrB* gene." 1(3): 24.
- Alexandrov, T., et al. (2013). "Analysis and interpretation of imaging mass spectrometry data by clustering mass-to-charge images according to their spatial similarity." 85(23): 11189-11195.
- Alexandrov, T. and J. H. J. B. Kobarg (2011). "Efficient spatial segmentation of large imaging mass spectrometry datasets with spatially aware clustering." 27(13): i230-i238.
- Altschul, S. F., et al. (1990). "Basic local alignment search tool." 215(3): 403-410.
- Bandeira, N., et al. (2007). "Protein identification by spectral networks analysis." 104(15): 6140-6145.
- Bauer, J. S., et al. (2015). "Biosynthetic origin of the antibiotic pseudopyronines A and B in *Pseudomonas putida* BW11M1." 16(17): 2491-2497.
- Bilal, M., et al. (2017). "Engineering *Pseudomonas* for phenazine biosynthesis, regulation, and biotechnological applications: a review." 33(10): 1-11.
- Blunt, J. and M. J. C. Munro, New Zealand: Marine Chemistry Group, Department of Chemistry, University of Canterbury (2003). "Marinlit: a database of the literature on marine natural products."
- Bouslimani, A., et al. (2014). "Mass spectrometry of natural products: current, emerging and future technologies." 31(6): 718-729.
- Buss, A. D. and M. S. Butler (2010). *Natural product chemistry for drug discovery*, Royal Society of Chemistry.
- Cawoy, H., et al. (2015). "Lipopeptides as main ingredients for inhibition of fungal phytopathogens by *Bacillus subtilis*/*amyloliquefaciens*." 8(2): 281-295.
- Chin-A-Woeng, T. F., et al. (2003). "Phenazines and their role in biocontrol by *Pseudomonas* bacteria." 157(3): 503-523.
- Cho, J. Y. J. B., biotechnology, and biochemistry (2012). "Algicidal activity of marine *Alteromonas* sp. KNS-16 and isolation of active compounds." 76(8): 1452-1458.
- Cody, R. B., et al. (2005). "Versatile new ion source for the analysis of materials in open air under ambient conditions." 77(8): 2297-2302.
- Debois, D., et al. (2008). "In situ localisation and quantification of surfactins in a *Bacillus subtilis* swarming community by imaging mass spectrometry." 8(18): 3682-3691.
- Ding, X. and Y. J. M. s. r. Duan (2015). "Plasma-based ambient mass spectrometry techniques: The current status and future prospective." 34(4): 449-473.

- Esquenazi, E., et al. (2009). "Probing marine natural product defenses with DESI-imaging mass spectrometry." 106(18): 7269-7270.
- Fenn, J. B., et al. (1989). "Electrospray ionization for mass spectrometry of large biomolecules." 246(4926): 64-71.
- Frank, A. M., et al. (2008). "Clustering millions of tandem mass spectra." 7(01): 113-122.
- Gevers, D., et al. (2005). "Re-evaluating prokaryotic species." 3(9): 733-739.
- Giddens, A. C., et al. (2008). "Natural product inhibitors of fatty acid biosynthesis: synthesis of the marine microbial metabolites pseudopyronines A and B and evaluation of their anti-infective activities." 64(7): 1242-1249.
- Goodwin, C. R., et al. (2012). "Structural mass spectrometry: rapid methods for separation and analysis of peptide natural products." 75(1): 48-53.
- Guthals, A., et al. (2012). "The spectral networks paradigm in high throughput mass spectrometry." 8(10): 2535-2544.
- Hillenkamp, F. and M. Karas (1990). [12] Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. *Methods in enzymology*, Elsevier. 193: 280-295.
- Horai, H., et al. (2010). "MassBank: a public repository for sharing mass spectral data for life sciences." 45(7): 703-714.
- Ito, T. and M. J. T. J. o. a. Masubuchi (2014). "Dereplication of microbial extracts and related analytical technologies." 67(5): 353-360.
- Junot, C., et al. (2014). "High resolution mass spectrometry based techniques at the crossroads of metabolic pathways." 33(6): 471-500.
- Karas, M. and F. J. A. c. Hillenkamp (1988). "Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons." 60(20): 2299-2301.
- Konstantinidis, K. T. and J. M. J. C. o. i. m. Tiedje (2007). "Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead." 10(5): 504-509.
- Konstantinidis, K. T. and J. M. J. P. o. t. N. A. o. S. Tiedje (2005). "Genomic insights that advance the species definition for prokaryotes." 102(7): 2567-2572.
- Kraft, M. L., et al. (2014). "Imaging lipids with secondary ion mass spectrometry." 1841(8): 1108-1119.
- Kumar, S., et al. (2018). "MEGA X: molecular evolutionary genetics analysis across computing platforms." 35(6): 1547.
- Laatsch, H. (2011). *AntiBase 2014: the natural compound identifier*, Wiley-Vch Weinheim, Germany.
- Lane, A. L., et al. (2009). "Desorption electrospray ionization mass spectrometry reveals surface-mediated antifungal chemical defense of a tropical seaweed." 106(18): 7314-7319.
- Lang, G., et al. (2008). "Evolving trends in the dereplication of natural product extracts: new methodology for rapid, small-scale investigation of natural product extracts." 71(9): 1595-1599.
- Li, J. W.-H. and J. C. J. S. Vederas (2009). "Drug discovery and natural products: end of an era or an endless frontier?" 325(5937): 161-165.
- Little, J. L., et al. (2011). "Identification of "known unknowns" utilizing accurate mass data and ChemSpider." 23(1): 179-185.

- Maiden, M. C., et al. (1998). "Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms." 95(6): 3140-3145.
- Maiden, M. C. J. A. R. M. (2006). "Multilocus sequence typing of bacteria." 60: 561-588.
- Monge, M. E., et al. (2013). "Mass spectrometry: recent advances in direct open air surface sampling/ionization." 113(4): 2269-2308.
- Newman, D. J. and G. M. J. J. o. n. p. Cragg (2012). "Natural products as sources of new drugs over the 30 years from 1981 to 2010." 75(3): 311-335.
- Nikolaev, E. N., et al. (2011). "Initial experimental characterization of a new ultra-high resolution FTICR cell with dynamic harmonization." 22(7): 1125-1133.
- Nothias, L.-F., et al. (2020). "Feature-based molecular networking in the GNPS analysis environment." 17(9): 905-908.
- Petras, D., et al. (2017). "High-resolution liquid chromatography tandem mass spectrometry enables large scale molecular characterization of dissolved organic matter." 4: 405.
- Picard, C., et al. (2000). "Frequency and biodiversity of 2, 4-diacetylphloroglucinol-producing bacteria isolated from the maize rhizosphere at different stages of plant growth." 66(3): 948-955.
- Ramette, A., et al. (2011). "Pseudomonas protegens sp. nov., widespread plant-protecting bacteria producing the biocontrol compounds 2, 4-diacetylphloroglucinol and pyoluteorin." 34(3): 180-188.
- Sawada, Y., et al. (2012). "RIKEN tandem mass spectral database (ReSpec) for phytochemicals: a plant-specific MS/MS-based data resource and database." 82: 38-45.
- Seeley, E. H. and R. M. J. A. c. Caprioli (2012). "3D imaging by mass spectrometry: a new frontier." 84(5): 2105-2110.
- Singh, M. P., et al. (2003). "Novel α -Pyrones Produced by a Marine Pseudomonas sp. F92S91 Taxonomy and Biological Activities." 56(12): 1033-1044.
- Smith, C. A., et al. (2005). "METLIN: a metabolite mass spectral database." 27(6): 747-751.
- Stone, M. and D. J. M. m. Williams (1992). "On the evolution of functional secondary metabolites (natural products)." 6(1): 29-34.
- Takats, Z., et al. (2004). "Mass spectrometry sampling under ambient conditions with desorption electrospray ionization." 306(5695): 471-473.
- Tamura, K., et al. (1993). "Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees." 10(3): 512-526.
- Tayeb, L. A., et al. (2005). "Molecular phylogeny of the genus Pseudomonas based on rpoB sequences and application for the identification of isolates." 156(5-6): 763-773.
- Tu, Y. and B. Yan (2012). High-throughput fractionation of natural products for drug discovery. High-Throughput Phenotyping in Plants, Springer: 117-126.
- Vaidyanathan, S., et al. (2008). "Subsurface biomolecular imaging of Streptomyces coelicolor using secondary ion mass spectrometry." 80(6): 1942-1951.
- Wall, L. G. J. H. o. m. m. e. l. m. and c. approaches (2011). "The BIOSPAS consortium: soil biology and agricultural production." 299-306.
- Wang, M., et al. (2016). "Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking." 34(8): 828-837.

- Watrous, J., et al. (2012). "Mass spectral molecular networking of living microbial colonies." 109(26): E1743-E1752.
- Watrous, J. D., et al. (2013). "Microbial metabolic exchange in 3D." 7(4): 770-780.
- Yang, J. Y., et al. (2013). "Molecular networking as a dereplication strategy." 76(9): 1686-1699.
- Yang, Y.-L., et al. (2009). "Translating metabolic exchange with imaging mass spectrometry." 5(12): 885-887.
- Ye, H., et al. (2011). "From pixel to voxel: a deeper view of biological tissue by 3D mass spectral imaging." 3(3): 313-332.
- Zeigler, D. R. J. I. j. o. s. and e. microbiology (2003). "Gene sequences useful for predicting relatedness of whole genomes in bacteria." 53(6): 1893-1900.

Capítulo 4

Diseño de un método económico para la tipificación molecular de aislamientos del género *Pseudomonas* con poder de resolución comparable a la filogenómica.

Lic. Andrés Muzlera

Laboratorio de Fisiología y Genética de Bacterias Beneficiosas para Plantas 106

4.1 Resumen

A lo largo de los años, las especies bacterianas se han definido en base a su morfología y sus actividades metabólicas. Luego, mediante hibridación DNA-DNA y por comparación de la secuencia de genes altamente conservados como el ARNr 16S, pero estas técnicas fueron sustituyéndose por métodos de caracterización genotípica utilizando una serie de genes marcadores que se encuentran conservados en diferentes especies. En esto se basan los Análisis Multilocus de Secuencia (del inglés *Multi Locus Sequence Analysis*, MLSA), aunque más recientemente se han desarrollado métodos basados en el estudio del genoma completo para obtener filogenias más precisas que las que se obtienen mediante MLSA, tales como la identidad media de nucleótidos (del inglés *average nucleotide identity*, ANI). Si bien los costos de secuenciación disminuyeron con el transcurso de los años, obtener la secuencia de un genoma bacteriano completo no es lo suficientemente sencillo ni económico para ser considerada una técnica de rutina para la identificación de múltiples aislamientos ambientales o clínicos. Es por esto que el MLSA es una metodología ampliamente utilizada para asignación preliminar o permanente de especie de un aislamiento nuevo, ya que es sencilla, rápida y económica. En general, la elección de los genes utilizados como estándar se ha basado solamente en 3 factores: que sean codificantes, que estén conservados y que se puedan diseñar oligonucleótidos para su amplificación. Sin embargo, la calidad informativa de los genes no suele ser un parámetro a tener en cuenta, lo cual puede resultar en una asignación taxonómica incorrecta (Öhrman et al. 2021; Tong et al. 2018; Colston et al. 2014).

El presente Capítulo se centra en la identificación sistemática y no sesgada de marcadores filogenéticos para el género *Pseudomonas* que, en base a su calidad de contenido informativo, permitan diseñar un MLSA con capacidad de discriminación de especies equivalente al de filogenómica basada en ANI. Este trabajo se realizó en colaboración con el Dr. Matías Garavaglia (becario postdoctoral en el LFGBBP hasta 2019, y actualmente docente-investigador en la Universidad Nacional de Hurlingham).

4.2 Introducción

4.2.1 El género *Pseudomonas*

El género *Pseudomonas* se introdujo a fines del siglo XIX (Palleroni 1984) y desde entonces ha experimentado una notable expansión hasta convertirse en el género Gram-negativo más grande en términos de número de especies tipo válidamente actualizadas, con 299 representantes en la LPNS (*List of Prokaryotic Names with Standing in Nomenclature*) (Parte et al. 2020) (<https://lpsn.dsmz.de/genus/pseudomonas>). A nivel celular y fisiológico, los miembros del género *Pseudomonas* se caracterizan por las siguientes características claves: son bacilos rectos o ligeramente curvados, de 0,5 a 1,0 μm de ancho $\times$ 1,5 a 5,0 μm de largo; móvil por uno o varios flagelos polares; quimioorganotrófico; de crecimiento aeróbico con un tipo de

metabolismo estrictamente respiratorio con oxígeno como aceptor final de electrones, o en algunos casos con crecimiento anaeróbico utilizando nitrato como aceptor alternativo de electrones; por lo general no requieren factores de crecimiento orgánicos; oxidasa positivo o negativo; catalasa positivo; cuando crece en alcanos o gluconato pueden acumular gránulos de polihidroxialcanoatos de longitudes de monómero superiores a cuatro carbonos; la mayoría de las especies no crecen en condiciones ácidas (pH 4,5 o inferior) (Palleroni 1984).

Quizás una de las propiedades más interesantes del género es su facilidad de aislamiento en medios de cultivo estándar, a partir de una gran variedad de nichos ecológicos, lo que en última instancia está determinado por su amplia versatilidad metabólica y su muy baja demanda nutricional (Silby et al. 2011; Ramos 2011). Se han aislado o detectado representantes del género *Pseudomonas* en estudios metagenómicos de hábitats naturales como agua dulce, suelo, animales, insectos, plantas y hongos, en condiciones relativamente suaves o extremas (Silby et al. 2011; Ramos 2011; Deakin et al. 2018; Saxena et al. 2017). El género comprende especies de relevancia clínica para humanos, animales y plantas, lo que genera un interés para la medicina, la veterinaria y la agronomía, pero también especies con características probióticas para animales y plantas, o con capacidades metabólicas con fines de biorremediación (Silby et al. 2011; Ramos 2011). Dada la amplia relevancia de los miembros del género para la salud humana, animal y vegetal, así como para aplicaciones biotecnológicas, es fundamental contar con un enfoque simple, rápido y confiable para su asignación taxonómica.

4.2.2 Metodologías aplicadas para la taxonomía en el género *Pseudomonas* en la historia

El género *Pseudomonas* fue descrito a finales del siglo XIX (Palleroni 1984) cuando las descripciones de los géneros se basaban en las morfologías macro y microscópicas, práctica universalmente aceptada por los taxónomos microbianos de esa época (Cohn 1872). Ya en el siglo XX las características fisiológicas se propusieron como criterio básico para la taxonomía bacteriana (Orla-Jensen 1909). En el Manual de Bergey publicado en 1923, se añadieron varias características fenotípicas, además de la morfología, tinción de Gram, tipo de flagelos y metabolismo con respecto al oxígeno, en un intento de diferenciar las especies del género *Pseudomonas* (Bergey 1923). Cuando las técnicas basadas en el ADN comenzaron a hacer posible el uso de enfoques genéticos en la taxonomía bacteriana (Schildkraut, Marmur, and Doty 1961), la composición de bases del ADN (G+C) y la hibridación de ADN-ADN fueron las primeras técnicas aplicadas a la taxonomía de *Pseudomonas* (Johnson and Ordal 1968; Colwell and Mandel 1964; Colwell, Citarella, and Ryman 1965). Así, el contenido de G+C de todas las especies de *Pseudomonas* fue incluido en el Manual de Bergey de 1974 (Doudoroff, Palleroni, and Wilkins Co. 1974). Más tarde, las pseudomonas se dividieron en cinco subgrupos de ARNr en función de las mediciones de hibridación ARN-ADN y esta clasificación se informó en el Manual de Bacteriología Sistemática de Bergey publicado en 1984 (Palleroni 1984).

Los cambios más profundos en la taxonomía bacteriana ocurrieron en la década de 80, cuando Woese y colaboradores propusieron el análisis de las secuencias del gen del ARN ribosómico 16S

para la clasificación de bacterias ubicando el género *Pseudomonas* en Gamma Proteobacteria (Woese et al. 1984). No obstante, lo más relevante de estos cambios que afectan al género *Pseudomonas* se inició a partir del año 2000 con un primer trabajo de recopilación de secuencias del gen 16S ARNr de 128 especies de *Pseudomonas* realizado por Anzai et al. (2000), quienes demostraron que muchas especies no encajaban dentro del grupo de *Pseudomonas*, en sentido estricto, que contenía los miembros del grupo I de ARNr de (Palleroni 1984). Los miembros de los grupos de ARNr restantes se dividieron en más de 25 géneros pertenecientes a las clases Alpha, Beta y Gamma proteobacteria (Peix et al. 2009; García-Valdés and Lalucat 2016). Estos cambios quedaron registrados en la edición del Manual de Bacteriología Sistemática de Bergey del año 2005, que pasó de formato impreso a formato en línea en 2015 donde cada género constituye un capítulo independiente (Whitman et al. 2015). Este nuevo formato permitirá actualizar con mayor frecuencia la información, lo que actualmente es fundamental teniendo en cuenta que el número de géneros y especies bacterianas es cada vez mayor.

4.2.3 Metodologías aplicadas para la taxonomía en el género *Pseudomonas* en la actualidad

La genómica modificó profundamente la forma en que se desarrolla la taxonomía bacteriana. El conocimiento de la secuencia del genoma completo de las cepas bacterianas y el análisis de sus secuencias de genes ha adquirido un valor superior a los rasgos fenotípicos para la clasificación e identificación bacteriana. Las comparaciones de genomas completos permiten un marco taxonómico más sólido conocido como taxogenómica. Las comparaciones *in silico* de genoma completo mediante el uso de la identidad de nucleótidos promedio (ANI) o cálculos de distancia de genoma a genoma (GGDC) son los nuevos estándares de oro (*gold standards*) para la asignación de especies, que sustituyen a los procedimientos experimentales de hibridación de ADN-ADN (Konstantinidis and Tiedje 2005; Meier-Kolthoff et al. 2013; Richter and Rosselló-Móra 2009). La secuencia del gen 16S ARNr fue inicialmente utilizada para definir relaciones filogenéticas entre bacterias (Anzai et al. 2000), pero hoy en día se reconoce que carece del poder discriminatorio suficiente para diferenciar especies en muchos géneros (por ejemplo, *Aeromonas*, *Bacillus*, *Pseudomonas*, *Streptococcus*, entre otros) (Janda and Abbott 2007). Por lo tanto, también se han propuesto otros genes de mantenimiento (*housekeeping*) como genes alternativos para estudios filogenéticos (Mulet, Lalucat, and García-Valdés 2010). Estos genes deben ser universales o al menos estar presentes en todas las especies del grupo a estudiar y pueden ser analizados individualmente, o sus secuencias pueden ser concatenadas y analizadas desde una perspectiva filogenética por MLSA. Para realizar estos análisis se han recomendado un conjunto de siete genes (16S rDNA, *rpoD*, *rpoB*, *gyrB*, *atpD*, *carA* y *recA*) (Stackebrandt et al. 2002), basándose solamente en los criterios previamente mencionados. Más recientemente, con el advenimiento de la secuenciación masiva y la accesibilidad de secuencias de genomas bacterianos completos, la filogenómica ha contribuido sustancialmente a una taxonomía bacteriana moderna. La filogenómica se centra en el estudio de las relaciones evolutivas entre las bacterias mediante el alineamiento múltiple de secuencias homólogas en el genoma central del grupo bacteriano a estudiar y mediante la inferencia de un árbol filogenético

correspondiente. Recientemente, Parks et al. (2018) propusieron una taxonomía bacteriana estandarizada (taxonomía GTDB; Genome Taxonomy Database) basada en la filogenia de genomas mediante el análisis de la secuencia de aminoácidos de 120 proteínas codificadas por 120 genes universales. En esta propuesta, los autores también incluyeron genomas ensamblados a partir de metagenomas (MAG) que aumentan sustancialmente la diversidad de especies bacterianas cultivadas hasta ahora. Los MAG podrían predecir las especies bacterianas existentes dentro de mezclas biológicas complejas sin la necesidad de su cultivo.

Las herramientas basadas en genomas completos han proporcionado criterios sólidos para el posicionamiento taxonómico de las especies de *Pseudomonas* (Parks et al. 2018; Lalucat et al. 2020). Sin embargo, aunque el costo de la secuenciación del genoma se redujo significativamente, sigue siendo una limitación para la catalogación rápida y precisa de colecciones de aislamientos. El único estudio que abordó el poder discriminativo para marcadores genéticos (Mulet, Lalucat, and García-Valdés 2010) útiles para MLSA, apoyó la selección de *rpoD*, *rpoB* y *gyrB* dentro de un conjunto de un total de 7 genes conservados, aunque las razones para considerar solo estos seis genes codificadores de proteínas, y no otros, no fueron explícitas (Mulet, Lalucat, and García-Valdés 2010). En este contexto, destacamos la importancia de contar con un método rápido, robusto y económico para definir la filiación taxonómica de nuevos aislamientos.

4.3 Hipótesis de trabajo

Las regiones comúnmente utilizadas para análisis por MLSA en *Pseudomonas* no fueron elegidas según su potencial informativo para realizar asignaciones taxonómicas. Por lo tanto, existen mejores regiones para amplificar, generar concatenados y utilizarlos para la filiación taxonómica de nuevos aislamientos con mayor precisión.

4.4 Objetivos específicos

- Desarrollar un flujo de trabajo no sesgado para la identificación de marcadores genéticos para realizar asignaciones taxonómicas por MLSA con la mayor precisión posible.
- Determinar el mejor concatenado de tres regiones para la filiación taxonómica de nuevos aislamientos del género *Pseudomonas*.

4.5 Materiales y métodos

4.5.1 Cepas bacterianas

Las cepas bacterianas utilizadas en este Capítulo se detallan en la Tabla 7. Las *Pseudomonas* fueron crecidas en agar nutritivo NA (agar nutritivo, 25 g/L; extracto de levadura, 5 g/L) o en caldo nutritivo NYB (caldo nutritivo, 25 g/L; extracto de levadura, 5 g/L), a 28°C, excepto *P. aeruginosa* que fue crecida a 37°C.

Tabla 7. Microorganismos utilizados

Cepa	Numero de acceso en <i>GenBank</i>	Referencia
<i>P. protegens</i> CHA0	NC_021237.1	(Stutz, Défago, and Kern 1986)
<i>P. aeruginosa</i> PAO1	NC_002516.2	(Stover et al. 2000)
<i>P. putida</i> KT2440	NC_002947.4	(Belda et al. 2016)
<i>P. sp.</i> BP01		(Sosa et al. 2020)
<i>P. sp.</i> SBMP6		(Agaras et al. 2015)
<i>P. sp.</i> SVMP4		(Agaras et al. 2015)
<i>P. sp.</i> BF61		(Martini et al. 2018)

4.5.2 Herramientas y softwares bioinformáticos

Para la comparación de genomas completos (obtenidos de www.pseudomonas.com; abril 2020) se utilizó la herramienta FastANI (Jain et al. 2018). Los alineamientos múltiples se generaron con MACSE v2 (Kumar et al. 2018). Los análisis de componentes principales se realizaron con el comando “prcomp” del paquete R “stats” (versión 3.6.2). Todos los concatenados se construyeron con la herramienta SEDA (Lopez-Fernandez et al. 2020). Los gráficos de correlación se generaron con el comando “chart.Correlation” del paquete R “PerformanceAnalytics” (versión 2.0.4). El método de correlación seleccionado fue “spearman” porque los datos no mostraban una distribución normal. Los árboles filogenéticos se generaron con MEGA X (Kumar et al. 2018). Los valores de Bootstrapping se generaron con el comando “boot.phylo” del paquete R “ape” (versión 5.4-1). Las pruebas t de Student se realizaron con el comando “t_test” del paquete R “rstatix” (versión 0.7.0), con el método de ajuste de Bonferroni.

4.5.3 Diseño de oligonucleótidos para PCR

Se utilizó la herramienta en línea PrimerDesign-M (Yoon and Leitner 2015) para diseñar oligonucleótidos degenerados capaces de amplificar la región objetivo en cada genoma de *Pseudomonas*. Se utilizó un alineamiento múltiple de secuencias para definir los oligonucleótidos y la longitud de amplificación se fijó en 850 pb aproximadamente. Cada posible par de oligonucleótidos se probó *in silico* con la herramienta ipcress (Slater and Birney 2005) para determinar el porcentaje de genomas en los que podrían amplificar la región correspondiente. También se analizaron posibles amplificaciones inespecíficas. Los

oligonucleótidos óptimos (Tabla 9) fueron aquellos que produjeron amplicones de al menos el 95% de los genomas, permitiendo el apareamiento incorrecto de solo un par de bases, y sin amplicones inespecíficos.

4.5.4 Amplificaciones por PCR

Las condiciones de las PCR se optimizaron para obtener amplicones únicos para cada región interna de los genes *core* seleccionados para el nuevo MLSA. Las mezclas de reacción consistieron en 20 µL que contenían buffer 1× (TransGen Biotech, China, con MgCl₂ 2,0 mM), desoxirribonucleótidos (dNTP) 0,2 mM (PB-L, Argentina), 0,8 µM de cada oligonucleótido (Tecnolab S.A, Argentina), dimetilsulfóxido al 5% (v/v) (DMSO, Sigma Aldrich, St Louis, Missouri, EE. UU.) y 1 U de ADN polimerasa EasyTaq (TransGen Biotech, China). El ciclo de PCR consistió en un paso inicial de desnaturalización de 5 min a 95 °C, seguido de 35 ciclos de 30 s a 95 °C, 30 s a 57 °C y 30 s a 72 °C, y un paso final de extensión a 72 °C durante 5 min. Como molde de ADN, se utilizó 1 µL de un lisado celular térmico de cada cepa. Los lisados térmicos se obtuvieron tratando una colonia de 2–3 mm resuspendida en 100 µL de H₂O ultrapura (grado PCR) a 100 °C durante 10 min. Después de la amplificación, se corrieron 10 µL de cada reacción en geles de agarosa al 1% en Tris-borato-EDTA (TBE) 0,5× a 10 V/cm durante 50 min. El gel se tiñó con bromuro de etidio y se visualizó bajo luz ultravioleta. Los productos de PCR se secuenciaron en MacroGen Inc. (Seúl, Corea) desde ambos extremos con los oligonucleótidos correspondientes (Tabla 9). La reconstrucción de la secuencia de cada uno de los fragmentos internos de los genes *core* y el control de la calidad de las secuencias se realizaron manualmente con la herramienta BlastN (Altschul et al. 1990).

4.6 Resultados y discusión

4.6.1 Curado de base de datos y obtención de genes *core* (CG)

Se descargaron un total de 526 genomas completos de la base de datos de *pseudomonas* (www.pseudomonas.com; abril 2020). 72 fueron removidos ya que no estaban asignados a ninguna especie. *Pseudomonas mesoacidophila* ATCC 31433 fue removida porque recientemente se reclasificó dentro del complejo *Burkholderia cepacia* (Loveridge et al. 2017). Para el resto de los 453 genomas se definieron dos bases de datos locales, la primera enumeraba cada genoma anotado en una línea separada (fna-DB) y la segunda enumeraba cada gen anotado por genoma (ffn-DB). A continuación, se desarrolló una serie de scripts en Bash para definir el conjunto de trabajo de genes *core* (CG) para el género *Pseudomonas* (Figura 53). Cada gen de ffn-DB se utilizó como interrogación (*query*) en una búsqueda por BlastN contra fna-DB, con un umbral de $E = 1 \times 10^{-15}$. Aquellos genes con solo un homólogo en cada uno de los 453 genomas de *pseudomonas* se mantuvieron y sus secuencias se recuperaron del ffn-DB para generar un alineamiento de los 453 homólogos por gen. En este punto nos dimos cuenta de que algunos genes no estaban anotados en el ffn-DB, pero su secuencia estaba en el fna-DB. Para superar este problema, se llevó a cabo una búsqueda BlastN utilizando todos los genes por genoma como *query* contra la fna-DB. El genoma con el mayor número posible de aciertos (*hits*) anotados

(453) fue el de *P. chlororaphis* JD37. Este genoma sirvió como fuente de secuencias de genes para anotar los ORF faltantes en algunos de los genomas recuperados y, más tarde, para definir el conjunto de trabajo de CG, es decir, genes que tienen homólogos en todos los 453 genomas recuperados. Todas las secuencias de genes anotadas de ffn-DB se recuperaron por homología con el genoma de la cepa JD37, y se construyeron archivos individuales con todas las secuencias homólogas de cada gen. La inspección de estos archivos permitió la identificación de genes faltantes (no anotados) y sus genomas correspondientes. Para recuperar la secuencia de los genes no anotados, realizamos una búsqueda BlastN contra fna-DB y recuperamos las secuencias alineadas más 500 pb río arriba y río abajo. Se utilizó la herramienta en línea ORFinder de NCBI para determinar el ORF de esas secuencias y se verificó su consistencia alineándolas como secuencias codificantes con el resto de sus homólogos. Al completar la anotación de ORF faltantes, llegamos a una lista final de 52 secuencias de genes que estaban presentes en los 453 genomas recuperados, que definieron nuestro conjunto de CG de *Pseudomonas* (Tabla 8).

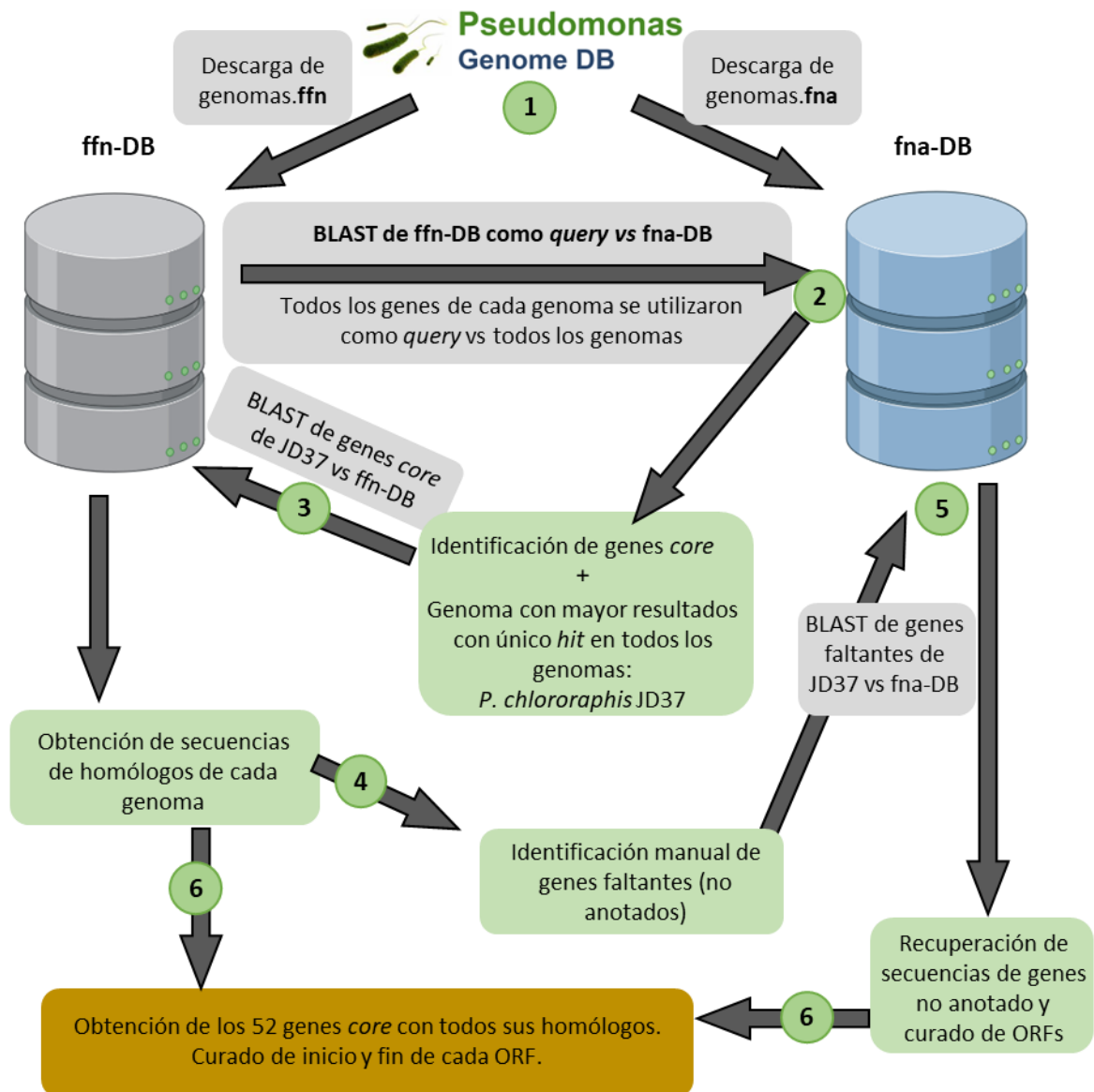


Figura 53. Esquema del flujo de trabajo empleado para la obtención de los 52 genes *core* del género *Pseudomonas*. 1) Se descargaron 453 genomas completos de *pseudomonas* en formato .ffn (fasta múltiple con los genes anotados) y en formato .fna (genoma completo) y se generó una base de datos con cada uno. 2) Se realizó un blast sobre fna-DB utilizando los genes de cada genoma en ffn-DB como *query* y se identificaron los 52 genes *core*. 3) Se recuperaron las secuencias de todos los genes *core* de cada genoma por homología realizando un blast contra ffn-DB. 4) Se confeccionaron fastas múltiples y se identificaron los genes faltantes (no anotados). 5) Se realizó un blast sobre fna-DB utilizando los genes de JD37 como *query* sobre los genomas con genes no anotados. 6) Determinación manual del correcto inicio y fin de cada ORF. Confección de 52 fastas múltiples con sus 453 genes *core* homólogos.

Tabla 8 Genes core del género *Pseudomonas*

Gen	Función	Anotación en el genoma de <i>P. aeruginosa</i> PAO1	coordenadas	Largo del gen
ahcY	S-Adenosyl-L-homocysteine hydrolase	PA0432	482706..484115 (-)	1410
alaT	Aminotransferase	PA2828	3181137..3182348 (+)	1212
argG	Argininosuccinate synthase	PA3525	3944662..3945879 (-)	1218
asd	Aspartate-semialdehyde dehydrogenase	PA3117	3499486..3500598 (-)	1113
atpA	ATP synthase subunit alpha	PA5556	6250596..6252140 (-)	1545
atpD	ATP synthase subunit beta	PA5554	6248278..6249654 (-)	1377
bioA	Adenosylmethionine-8-amino-7-oxononanoate transaminase	PA0420	466833..468236 (-)	1404
bioB	Biotin synthase	PA0500	559644..560702 (+)	1059
carB	Carbamoyl phosphate synthase large subunit	PA4756	5339864..5343085 (-)	3222
cysN	ATP sulfurylase GTP-binding subunit/APS kinase	PA4442	4974694..4976595 (-)	1902
dxs	1-Deoxyxylulose-5-phosphate synthase	PA4044	4526547..4528430 (+)	1884
gap	Glyceraldehyde-3-phosphate dehydrogenase	PA3195	610214..611275 (+)	1062
gatB	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B	PA4484	5015535..5016980 (+)	1446
glnA	Glutamine synthetase	PA5119	5766484..5767893 (+)	1410
glnS	Glutamyl-tRNA synthetase	PA1794	1943067..1944737 (+)	1671
gltA	Type II citrate synthase	PA1580	1719109..1720395 (-)	1287
gltB	Glutamate synthase subunit alpha	PA5036	5667696..5672141 (-)	4446
gltD	Glutamate synthase subunit beta	PA5035	5666234..5667667 (-)	1434
gltX	Glutamyl-tRNA synthetase	PA3134	3515416..3516900 (-)	1485
groEL	GroEL 60 kDa molecular chaperone	PA4385	4915481..4917124 (-)	1644
gyrB	DNA gyrase subunit B	PA0004	4275..6695 (+)	2421
ilvB	Acetolactate synthase	PA4696	5274008..5275732 (-)	1725
metG	Methionyl-tRNA synthetase	PA3482	3895324..3897357 (+)	2034
metH	B12-dependent methionine synthase	PA1843	1999889..2003593 (-)	3705
metK	S-Adenosylmethionine synthase	PA0546	603706..604896 (-)	1191
parE	DNA Topoisomerase IV subunit B	PA4967	5576028..5577917 (-)	1890
pepN	Aminopeptidase N	PA3083	3460380..3463037 (-)	2658
pnp	Polynucleotide phosphorylase/polyadenylase	PA4740	5323374..5325479 (-)	2106
prkA	PrkA family serine protein kinase	PA0588	646730..648652 (-)	1923
prlC	Oligopeptidase A	PA0067	78710..80755 (+)	2046
purB	Adenylosuccinate lyase	PA2629	2972699..2974069 (+)	1371
pyrB	Aspartate carbamoyltransferase	PA0402	444687..445691 (-)	1005
rep	ATP-dependent DNA helicase Rep	PA5296	5962716..5964725 (+)	2010
rplV	50S Ribosomal protein L22	PA4258	4764243..4764575 (-)	333
rpoA	DNA-directed RNA polymerase alpha	PA4238	4754423..4755424 (-)	1002
rpoB	DNA-directed RNA polymerase subunit beta	PA4270	4776544..4780617 (-)	4074
rpoD	RNA polymerase sigma factor	PA0576	634371..636224 (-)	1854
rpoH	RNA polymerase sigma factor	PA0376	420683..421537 (+)	855
rpsA	30S Ribosomal protein S1	PA3162	3548110..3549789 (-)	1680
rpsD	Ribosomal protein S4	PA4239	4755447..4756067 (-)	621
rpsH	Ribosomal protein S8	PA4249	4760480..4760872 (-)	393
sdhA	Succinate dehydrogenase	PA1583	1721496..1723268 (+)	1773
speA	Arginine decarboxylase	PA4839	5431381..5433291 (-)	1911
spoVR	SpoVR family protein	PA0586	643808..645361 (-)	1554
sucA	Oxoglutarate dehydrogenase E1 component	PA1585	1724244..1727075 (+)	2832

<i>thiC</i>	Phosphomethylpyrimidine synthase	PA4973	5581762..5583645 (-)	1884
<i>trpE</i>	Anthranilate synthase component I	PA0609	670980..672458 (+)	1479
<i>typA</i>	Translational GTPase	PA5117	5762659..5764611 (-)	1818
<i>uvrA</i>	Excinuclease ABC subunit A	PA4234	4748756..4751593 (+)	2838
<i>valS</i>	Valyl-tRNA	PA3834	4291356..4294208 (+)	2853
<i>ygiR</i>	Uncharacterized radical SAM protein	PA4928	5528659..5530902 (-)	2244
<i>yhgF</i>	RNA-binding transcriptional accessory protein	PA5201	5853585..5855924 (+)	2340

4.6.2 Evaluación del potencial resolutivo de cada amplicón

Esta sección resume el flujo de trabajo (Figura 53) realizado por el Dr. Matías Garavaglia para identificar regiones internas de tres CG, a partir del conjunto presentado en la Tabla 8, para generar un concatenado que permita realizar una asignación filogenética por MLSA con poder de resolución comparable a la filogenómica basada en ANI.

Los 453 genomas fueron comparados entre ellos con FastANI (tamaño de k-mero = 16, largo del fragmento = 3,000 (Jain et al. 2018)) para seleccionar representantes únicos de cada especie. Dentro de cada especie (basado en un corte ANI <95% (Zhao et al. 2021)), se seleccionaron genomas representativos considerando aquellos con el rango ANI más amplio entre ellos y limitando su número a un máximo del 10% del total de genomas seleccionados. Se obtuvieron un total de 126 genomas completos para análisis posteriores.

El siguiente paso fue identificar qué secuencia interna de un CG de un tamaño apropiado para PCR y secuenciación bidireccional sería mejor para fines taxonómicos, en comparación con los estándares actuales. Primero, se obtuvieron alineamientos para todos los 52 CG, utilizando MUSCLE en MEGA X (Kumar et al. 2018). Los alineamientos se modificaron manualmente para eliminar espacios al inicio y al final de cada secuencia teniendo en cuenta los tripletes codificantes. En esta etapa, se generaron dos alineamientos de referencia (estándares) para fines comparativos posteriores: 1) la referencia "MLSA" se obtuvo alineando un concatenado de las secuencias internas de los genes marcadores *rpoD*, *rpoB* y *gyrB* según lo definido por Mulet, Lalucat, and García-Valdés (2010), para cada uno de los 126 genomas de trabajo; 2) la referencia "MLSA52" se obtuvo alineando la concatenación de los 52 CG para cada uno de los 126 genomas.

Con el objetivo de construir matrices de distancia y árboles filogenéticos para análisis posteriores, se seleccionó el mejor modelo de sustitución para el alineamiento MLSA52, y se aplicó al MLSA de referencia, a cada CG individual, y a cada alineamiento múltiple de fragmentos individuales de 850 pb derivadas de todos los CG. Los árboles filogenéticos se construyeron con el método de máxima verosimilitud, a excepción de MLSA52 que se generó con unión de vecinos (*Neighbor-Joining*) debido al alto costo computacional asociado con el tamaño de la muestra. Se aplicó la prueba de congruencia entre matrices de distancia (CADM (Campbell, Legendre, and Lapointe 2011)) para comparar cada matriz de distancia con la proveniente de FastANI o la proveniente de MLSA52. Por otro lado, se aplicaron una serie de métodos para comparar la topografía y coherencia de los nodos de los árboles filogenéticos generados (*Robinson-Foulds*, *Path distance*, *Cophenetic correlation*, *Quartet distances*, *Match Split test*). Los resultados

numéricos de cada una de estas pruebas se normalizaron en un rango de 0 a 1 para facilitar más comparaciones.

Para buscar correlaciones intra o entre métodos, se aplicó la prueba de correlación de Pearson y un análisis de componentes principales (PCA), y los valores obtenidos se promediaron para los árboles ANI y MLSA52. Se seleccionaron aquellas pruebas que mostraron resultados congruentes entre sí, luego se realizó un PCA con todos los valores de las pruebas congruentes y se calculó la distancia euclidiana entre las coordenadas de cada muestra con árboles ANI o MLSA52. Tras la normalización de los valores en un rango de 0 a 1, se generó una lista con las regiones individuales de 850 pb con las puntuaciones más altas, que resultaron ser: *rpoD*⁵²⁶⁻¹³⁷⁵, *dxs*¹²⁶⁻⁹⁷⁵, *pepN*¹³⁵¹⁻²²⁰⁰, *gltB*⁹⁷⁶⁻¹⁸²⁵, *spoVR*⁵²⁶⁻¹³⁷⁵, *uvrA*¹⁸²⁶⁻²⁶⁷⁵, *ilvB*⁸⁵¹⁻¹⁷⁰⁰, *bioA*²⁶⁻⁸⁷⁵, *gltX*⁴⁷⁶⁻¹³²⁵ and *sucA*⁴⁷⁶⁻¹³²⁵.

Por último, se repitió el flujo de trabajo propuesto para elegir las mejores regiones internas de los diferentes CG, pero utilizando todos los posibles concatenados de a tres de las diez mejores regiones. De esta manera se obtuvo que el concatenado de los genes *rpoD*⁵²⁶⁻¹³⁷⁵, *gltX*⁴⁷⁶⁻¹³²⁵ y *pepN*¹³⁵¹⁻²²⁰⁰ proporciona la mejor resolución para la asignación filogenética en *Pseudomonas* a través de MLSA.

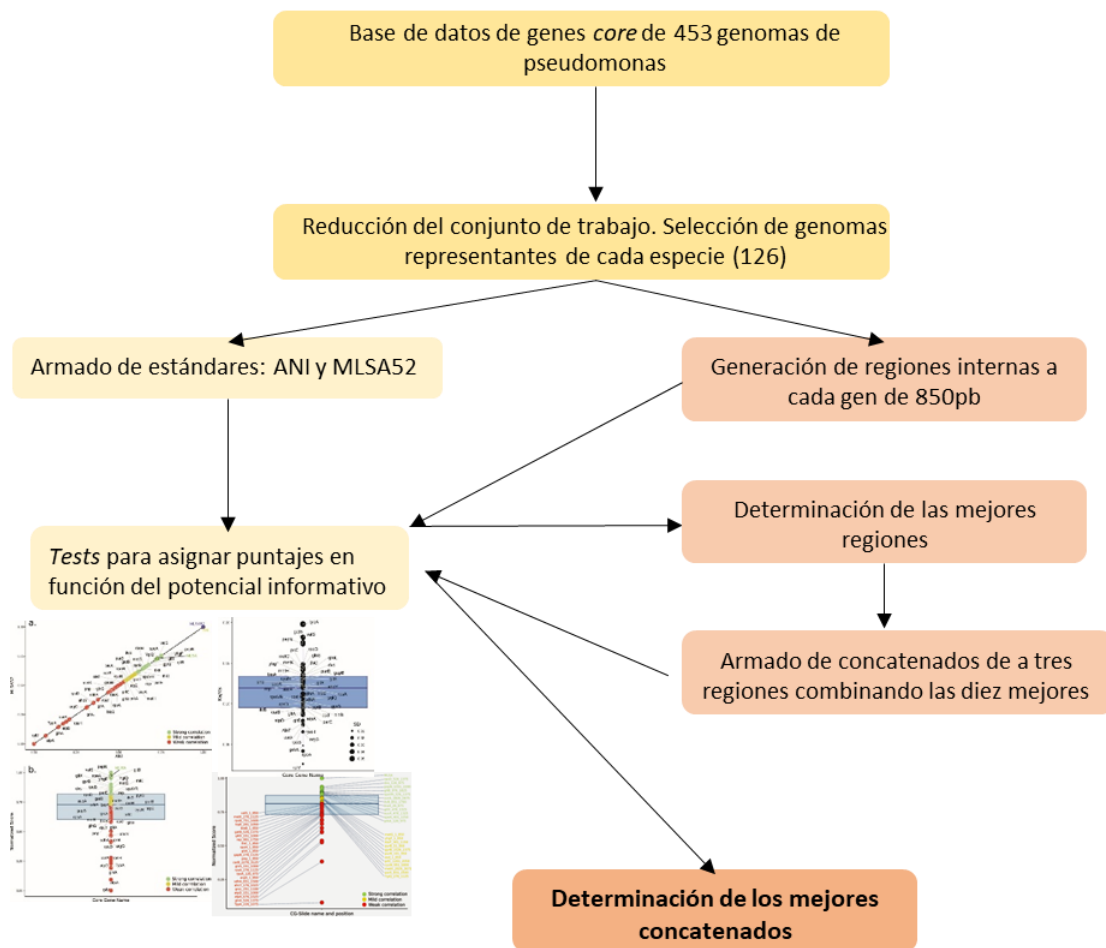


Figura 54. Esquema del flujo de trabajo empleado para la obtención de los mejores concatenados para utilizar en MLSA

Además de generar un árbol robusto con una topología similar a los derivados de los valores ANI o MLSA52 (nuestros estándares de referencia externos), es importante que el concatenado propuesto logre una excelente discriminación a nivel de especie. Para probar esto, se transformó la matriz derivada de la similitud ANI asignando valores de 0 a aquellas especies con valores ANI = 100 % (matriz diagonal), 0.5 a aquellos valores > 94 % y 1 a aquellos valores ≤ 94 %. Elegimos el 94 % como valor de corte del ANI (en lugar del 95 %) porque los valores entre el 94 y el 95 % requieren un procedimiento más exhaustivo para la discriminación de especies (Richter and Rosselló-Móra 2009), y también porque los diagramas de densidad de los valores del ANI revelaron dos picos en el 93,5 % y 95,5%, y luego un valle hasta el último pico de valores cercanos al 100% de ANI (Figura 55a). A continuación, generamos matrices equivalentes con las distancias euclidianas calculadas para cada concatenado y/o la referencia MLSA. Como no se sabía de antemano qué valores de distancia TN93+G+I se correlacionaban con el umbral ANI de 93,5-95,5% (Figura 55a), se evaluó un rango de valores de distancia entre 0,010 y 0,099 (Figura 55b).

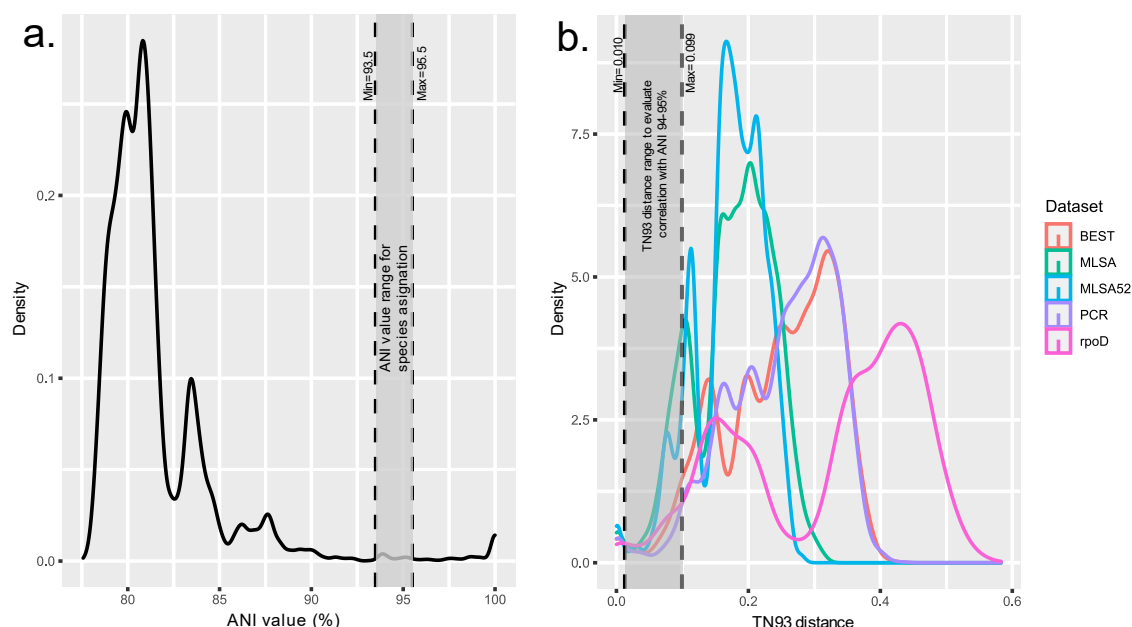


Figura 55. Comparación del rendimiento filogenético del concatenado *rpoD-pepN-gltX* y ANI. a) Diagrama de densidad de valores ANI para el conjunto de datos de genomas utilizados. Los valores entre líneas discontinuas indican los valores para los cuales la asignación de especies es dudosa (93,5% a 95,5%). b) Diagrama de densidad de los valores de distancia TN93 para cada conjunto de datos. La región entre líneas discontinuas corresponde a una posible zona de concordancia con ANI >94%. "BEST" se refiere al concatenado *pepN-rpoD-sucA* y "PCR" al concatenado *rpoD-pepN-gltX*. Modificado de Garavaglia, Muzlera et. al. (2022)

Luego se realizó un análisis de correlación de matrices de Mantel (CADM) para encontrar el valor de distancia TN93+G+I con una mejor representación de la matriz ANI transformada y para encontrar el valor de umbral TN93 correspondiente. Con este enfoque, encontramos que el valor de correlación de W Kendall más alto (0.998) fue el del concatenado *rpoD-pepN-gltX* con un valor de umbral TN93 de 0.054 (Figura 56)

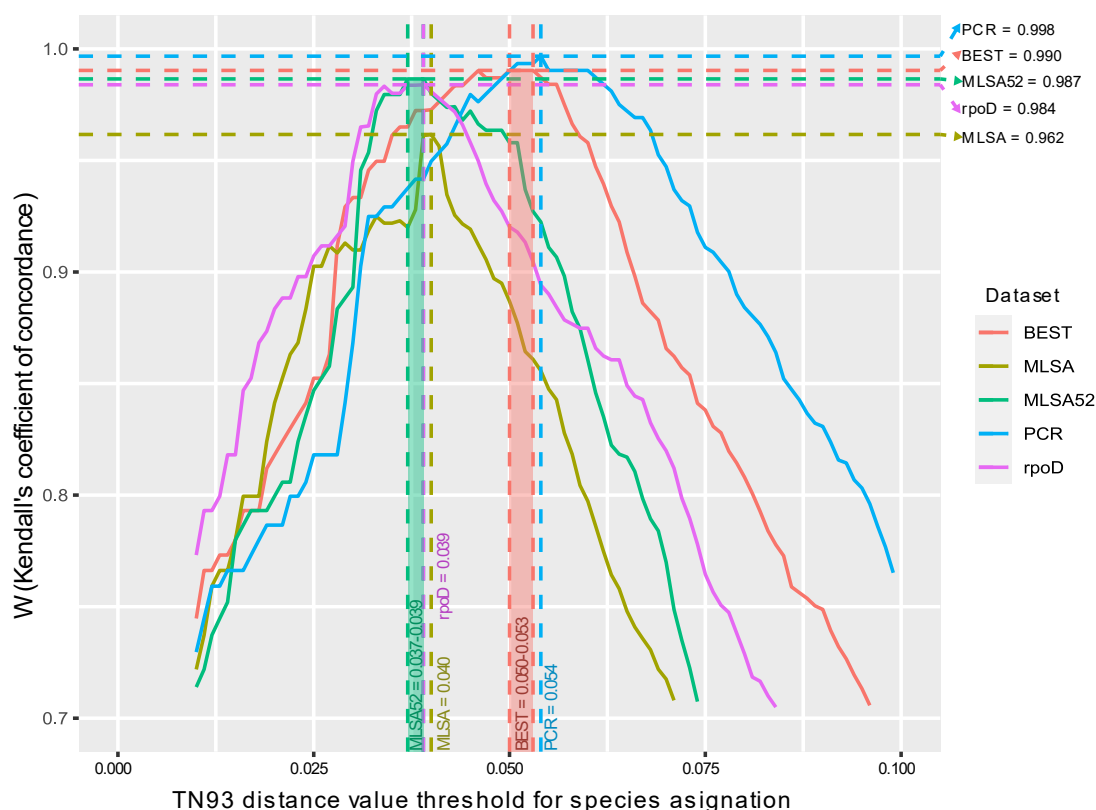


Figura 56. Valores de concordancia de Kendall entre las matrices de distancia de ANI y todos los valores posibles de TN93 para cada conjunto de datos. Las líneas discontinuas resaltan los valores de concordancia más altos (w) entre ANI y cada conjunto de datos. "BEST" se refiere al concatenado *pepN-rpoD-sucA* y "PCR" al concatenado *rpoD-pepN-gltX*. Modificado de Garavaglia, Muzlera et. al. (2022)

4.6.3 Diseño y validación de oligonucleótidos

Se diseñaron pares de oligonucleótidos con el objetivo de amplificar las regiones definidas previamente, generando productos de 850pb aproximadamente. Partiendo de alineamientos múltiples y utilizando la herramienta online PrimerDesign-M, se seleccionaron entre 2 y 5 pares de oligonucleótidos posibles para amplificar cada región. Para definir el mejor par, se realizaron simulaciones de PCRs empleando la herramienta ipcress, y seleccionamos el par que pueda ser capaz de amplificar la región objetivo en el mayor número de genomas posibles, sin generar amplicones inespecíficos. Para amplificar la región interna del gen *rpoD*, se seleccionó el par diseñado por Mulet et al. (2009) ya que la región definida se solapaba con dicho amplicón en un 80% aproximadamente (nuevo amplicón : *rpoD*⁵²⁶⁻¹³⁷⁵; propuesto por Mulet : *rpoD*³⁵⁵⁻¹⁰⁹⁷). Se analizó el concatenado generado con el fragmento de *rpoD* propuesto por Mulet y se observó que no se modificó su potencial para MLSA, por lo que se tomó la decisión de no rediseñarlos. Una vez seleccionados los tres pares de oligonucleótidos a utilizar (Tabla 9), se procedió a evaluar las condiciones de PCR, a su secuenciación y a realizar MLSA. Se seleccionaron tres cepas de referencia con genomas completos disponibles (*P. protegens* CHA0, *P. putida* KT2440 y *P. aeruginosa* PAO1), tres aislamientos sin asignación taxonómica del cepario de nuestro laboratorio (SVMP4, SBMP6, (Agaras, Wall, and Valverde 2012) y BP01 (Sosa et al. 2020), y un

cuarto aislamiento BF61 provisto por la Dra. M.F. del Papa (Martini et al. 2015)). Las reacciones de secuenciación por Sanger fueron exitosas y produjeron lecturas completas de muy buena calidad para las siete cepas. Las secuencias crudas se curaron manualmente para eliminar los extremos 5' y 3' mal resueltos, ambas cadenas se alinearon para resolver posibles incongruencias, y finalmente obtuvimos secuencias internas de CG útiles de 713-716 nt (*rpoD*) y de 785-815 nt (*pepN* y *gltX*) para el posterior posicionamiento filogenético. Las secuencias internas *rpoD*, *pepN* y *gltX* para las cepas de *Pseudomonas* no caracterizadas se han depositado en GenBank con los códigos de acceso OL451240, MT154529.1, MT154528.1 para BF61; OL451241, MT154526.1, MT154525.1 para BP01; OL451242, MT154523.1, MT154522.1 para SBMP6; OL451243, MT154520.1, MT154519.1 para SVMP4.

Tabla 9. Oligonucleótidos diseñados para MLSA

Gen Core	Oligonucleótido	Secuencia	Posición dentro del gen	Tamaño del amplicón (bp)	Referencia
<i>rpoD</i>	PsEG30F	ATYGAAATCGCCAARCG	355-372	742	(Mulet et al. 2009)
	PsEG790R	CGGTTGATKTCCTTGA	1082-1097		(Mulet et al. 2009)
<i>pepN</i>	PepN-FW	GGYTTCAAYCGYTGGGANGC	1711-1730	844	Este trabajo
	PepN-rv	CRTAYTTRCGCCAGCGRGTC	2535-2554		Este trabajo
<i>gltX</i>	GltX-FW	GGCGGYSARTTCATYCTGCG	103-122	863	Este trabajo
	GltX-rv	TTSAGCCASGASAGYTTCTC	946-965		Este trabajo

Se recuperaron las regiones de los genes correspondientes a los 126 genomas utilizados en nuestro flujo de trabajo, se combinaron con los concatenados provenientes de la secuenciación de las siete cepas analizadas, y se generó un alineamiento múltiple que se utilizó para construir un árbol filogenético utilizando el Método de Máxima Verosimilitud (Figura 57). Los concatenados provenientes de las amplificaciones de las cepas CHA0 y PAO1 se ubicaron en la misma posición que sus contrapartes extraídas de bibliografía, mientras que la cepa KT2440 se ubicó dentro del complejo *P. putida*, ya que *P. putida* KT2440 no formó parte de los 126 genomas seleccionados en la base de datos. Este resultado representa la validación empírica del diseño computacional propuesto en este capítulo. El aislamiento BF61 se ubicó dentro del clado de *P. fluorescens*, mientras que el aislamiento BP01 se ubicó dentro del grupo de *P. putida*. Los aislamientos SVMP4 y SBMP6 se ubicaron dentro del grupo de *P. mosselii*, y además se encuentran filogenéticamente muy relacionadas entre sí, esto fue coherente con los resultados obtenidos del perfil de metabolitos de sus sobrenadantes de cultivo (Capítulo 3).

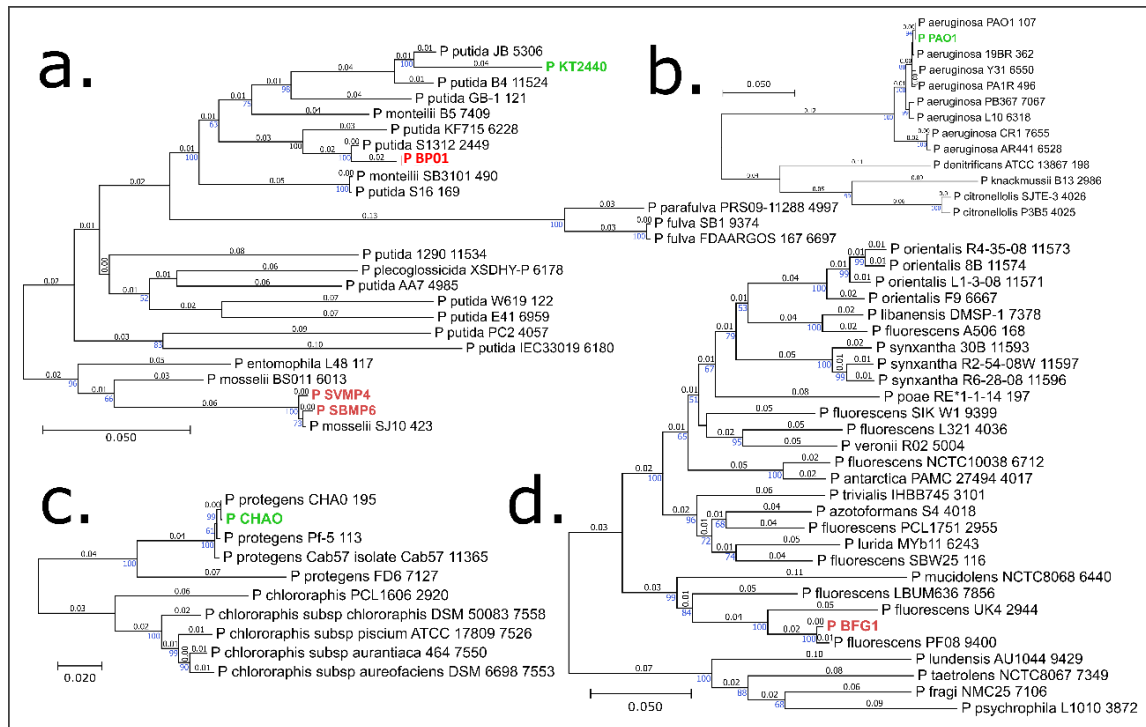


Figura 57 Posicionamiento filogenético de *Pseudomonas* caracterizadas y aislamientos no caracterizados sobre la base del concatenado *rpoD-pepN-gltX* propuesto en este Capítulo. Los nombres de las cepas en verde corresponden a las cepas caracterizadas utilizados en este estudio, mientras que los nombres en rojo corresponden a aislamientos no caracterizados utilizados. Los valores de *Bootstrapping* en azul están asociados a los nodos del árbol. Los valores de distancia TN93 se indican en negro. Los paneles a, b, c y d corresponden a los clados recortados del árbol completo donde colocaron las cepas analizadas.

Figura modificada de Garavaglia, Muzlera et. al. (2022)

Las cepas de *Pseudomonas* son relativamente fáciles de aislar y cultivar en el laboratorio. Por estas razones, el posicionamiento taxonómico de los aislamientos dentro del género es clave para diferentes disciplinas, desde el estudio de patógenos humanos, animales y vegetales hasta la producción agrícola, la biorremediación y la ecología. De hecho, la organización de la estructura filogenética de *Pseudomonas*, compuesta por más de 270 especies (<https://lpsn.dsmz.de/genus/pseudomonas>) (Martini et al. 2015), es un tema de constante revisión y actualización (Lalucat et al. 2020; Girard et al. 2021; Peix et al. 2018).

Sin embargo, durante la caracterización de nuevos aislamientos, su asignación de especie se lleva a cabo con frecuencia con estudios de profundidad insuficiente (p. ej., basándose solo en la secuencia parcial o completa del ADNr 16S o, en algunos casos, ni siquiera está respaldada por información basada en el ADN), lo que conduce a la asignación de especies incorrectas (Öhrman et al. 2021; Tong et al. 2018; Colston et al. 2014). Los nombres incorrectos dificultan la posibilidad de realizar un análisis comparativo sólido, pero, lo que es más importante, pueden introducir sesgos en la descripción de las especies (Espariz et al. 2016). Aun así, incluso si la información del genoma está disponible para un aislamiento dado, puede suceder que los datos genómicos terminen cargados en bases de datos públicas como entradas asociadas con especies

mal nombradas sin una asignación taxonómica completa (Öhrman et al. 2021; Tong et al. 2018; Colston et al. 2014).

El MLSA basado en el concatenado *rpoD*³⁵⁵⁻¹⁰⁹⁷-*gltX*⁴⁷⁶⁻¹³²⁵-*pepN*¹³⁵¹⁻²²⁰⁰ requiere solo de la secuenciación bidireccional de Sanger de tres productos de PCR para cada aislamiento de interés. Esto representa un costo total aproximado de purificación y secuenciación de PCR de 30-35 U\$S por aislamiento, dependiendo de la empresa. Los costos de purificación se pueden reducir aún más si se agregan etiquetas con secuencias universales en la región 5' de cada par de oligonucleótidos específicos para permitir la combinación de las tres reacciones de PCR en un solo tubo para realizar solo una purificación seguida de una secuenciación utilizando cada uno de los oligonucleótidos universales (p. ej., M13, Bluescript, T3/T7). Creemos que este es un costo razonable y accesible para obtener el posicionamiento filogenético de un aislamiento con una resolución similar a la del ANI, a un costo total de una décima parte de la secuenciación del genoma.

4.7 Conclusión

Se estableció un flujo de trabajo computacional para analizar 126 genomas completos de más de 60 especies y subespecies de *Pseudomonas* en un enfoque imparcial y sistemático, para seleccionar regiones internas de genes conservados que pueden ser objeto de PCR y posterior secuenciación de Sanger para generar MLSA. La capacidad filogenética alcanzada fue comparable al método actualmente utilizado como estándar (ANI). Encontramos que *rpoD*³⁴⁶⁻¹¹⁹⁶-*pepN*¹⁷¹¹⁻²⁵⁷¹-*gltX*⁸⁶⁻⁹⁰⁹ fue el concatenado que mejor se desempeñó en términos de robustez filogenética en comparación con ANI. El MLSA basado en dicho concatenado representa una herramienta filogenética rápida, accesible y sólida para aislamientos del género *Pseudomonas*, con un poder de resolución taxonómico comparable a los enfoques basados en el genoma completo.

4.8 Bibliografía

- Agaras, B., et al. (2012). "Specific enumeration and analysis of the community structure of culturable pseudomonads in agricultural soils under no-till management in Argentina." 61: 305-319.
- Agaras, B. C., et al. (2015). "Quantification of the potential biocontrol and direct plant growth promotion abilities based on multiple biological traits distinguish different groups of *Pseudomonas* spp. isolates." 90: 173-186.
- Altschul, S. F., et al. (1990). "Basic local alignment search tool." 215(3): 403-410.
- Anzai, Y., et al. (2000). "Phylogenetic affiliation of the pseudomonads based on 16S rRNA sequence." 50(4): 1563-1589.
- Belda, E., et al. (2016). "The revisited genome of *Pseudomonas putida* KT2440 enlightens its value as a robust metabolic chassis." 18(10): 3403-3424.
- Bergey, D. (1923). "Manual of Determinative Bacteriology, 1st ed."
- Campbell, V., et al. (2011). "The performance of the Congruence Among Distance Matrices (CADM) test in phylogenetic analysis." 11(1): 1-15.
- Cohn, F. J. Ö. B. Z. (1872). "Conspectus familiarum cryptogamarum secundum methodum naturalem dispositarum." 346-349.
- Colston, S. M., et al. (2014). "Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case." 5(6): e02136-02114.
- Colwell, R., et al. (1965). "Deoxyribonucleic acid base composition and Adansonian analysis of heterotrophic aerobic pseudomonads." 90(4): 1148-1149.
- Colwell, R. and M. J. J. o. B. Mandel (1964). "Adansonian analysis and deoxyribonucleic acid base composition of some Gram-negative bacteria." 87(6): 1412-1422.
- Deakin, G., et al. (2018). "Spatial structuring of soil microbial communities in commercial apple orchards." 130: 1-12.
- Doudoroff, M., et al. (1974). "*Pseudomonas* Migula 1894." 217-243.
- Espariz, M., et al. (2016). "Taxonomic identity resolution of highly phylogenetically related strains and selection of phylogenetic markers by using genome-scale methods: the *Bacillus pumilus* group case." 11(9): e0163098.
- García-Valdés, E. and J. Lalucat (2016). *Pseudomonas: molecular phylogeny and current taxonomy*. *Pseudomonas: molecular and applied biology*, Springer, Cham: 1-23.
- Garavaglia JM, Muzlera A and Valverde CF, *Molecular Phylogenetics & Evolution* 2022. 'Comparative genomics and informational content analysis uncovered internal regions of the core genes *rpoD*, *pepN* and *gltX* for an MLSA with genome-level resolving power within the genus *Pseudomonas*', En segunda revision.
- Girard, L., et al. (2021). "The ever-expanding *Pseudomonas* genus: description of 43 new species and partition of the *Pseudomonas putida* group." 9(8): 1766.
- Jain, C., et al. (2018). "High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries." 9(1): 1-8.
- Janda, J. M. and S. L. J. J. o. c. m. Abbott (2007). "16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls." 45(9): 2761-2764.

- Johnson, J. L. and E. J. J. J. o. B. Ordal (1968). "Deoxyribonucleic acid homology in bacterial taxonomy: effect of incubation temperature on reaction specificity." 95(3): 893-900.
- Konstantinidis, K. T. and J. M. J. P. o. t. N. A. o. S. Tiedje (2005). "Genomic insights that advance the species definition for prokaryotes." 102(7): 2567-2572.
- Kumar, S., et al. (2018). "MEGA X: molecular evolutionary genetics analysis across computing platforms." 35(6): 1547.
- Lalucat, J., et al. (2020). "Genomics in bacterial taxonomy: impact on the genus *Pseudomonas*." 11(2): 139.
- Lopez-Fernandez, H., et al. (2020). "SEDA: a desktop tool suite for FASTA files processing."
- Loveridge, E. J., et al. (2017). "Reclassification of the specialized metabolite producer *Pseudomonas mesoacidophila* ATCC 31433 as a member of the *Burkholderia cepacia* complex." 199(13): e00125-00117.
- Martini, M. C., et al. (2015). "Characterization of a collection of plasmid-containing bacteria isolated from an on-farm biopurification system used for pesticide removal." 80: 16-23.
- Martini, M. C., et al. (2018). "Novel environmental class 1 integrons and cassette arrays recovered from an on-farm bio-purification plant." 94(3): fix190.
- Meier-Kolthoff, J. P., et al. (2013). "Genome sequence-based species delimitation with confidence intervals and improved distance functions." 14(1): 1-14.
- Mulet, M., et al. (2009). "An *rpoD*-based PCR procedure for the identification of *Pseudomonas* species and for their detection in environmental samples." 23(3-4): 140-147.
- Mulet, M., et al. (2010). "DNA sequence-based analysis of the *Pseudomonas* species." 12(6): 1513-1530.
- Öhrman, C., et al. (2021). "Reorganized genomic taxonomy of Francisellaceae enables design of robust environmental PCR assays for detection of *Francisella tularensis*." 9(1): 146.
- Orla-Jensen, S. J. Z. B. P. I. H. A. I. (1909). "Die hauptlinien des natürlichen bakteriensystems." 22: 305-346.
- Palleroni, N. J. B. s. m. o. s. b. (1984). "Genus I. *Pseudomonas*." 1: 141-199.
- Parks, D. H., et al. (2018). "A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life." 36(10): 996-1004.
- Parte, A. C., et al. (2020). "List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ." 70(11): 5607.
- Peix, A., et al. (2009). "Historical evolution and current status of the taxonomy of genus *Pseudomonas*." 9(6): 1132-1147.
- Peix, A., et al. (2018). "The current status on the taxonomy of *Pseudomonas* revisited: an update." 57: 106-116.
- Ramos, J.-L. (2011). *Pseudomonas: Volume 1 Genomics, Life Style and Molecular Architecture*, Springer Science & Business Media.
- Richter, M. and R. J. P. o. t. N. A. o. S. Rosselló-Móra (2009). "Shifting the genomic gold standard for the prokaryotic species definition." 106(45): 19126-19131.
- Saxena, R., et al. (2017). "Metagenomic analysis of hot springs in Central India reveals hydrocarbon degrading thermophiles and pathways essential for survival in extreme environments." 7: 2123.

- Schildkraut, C. L., et al. (1961). "The formation of hybrid DNA molecules and their use in studies of DNA homologies." 3(5): 595-IN516.
- Silby, M. W., et al. (2011). "Pseudomonas genomes: diverse and adaptable." 35(4): 652-680.
- Slater, G. S. C. and E. J. B. b. Birney (2005). "Automated generation of heuristics for biological sequence comparison." 6(1): 1-11.
- Sosa, M. F., et al. (2020). "A black-pigmented pseudomonad isolate with antibacterial activity against phyllospheric pathogens." 15: 100207.
- Stackebrandt, E., et al. (2002). "Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology." 52(3): 1043-1047.
- Stover, C. K., et al. (2000). "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen." 406(6799): 959-964.
- Stutz, E., et al. (1986). "Naturally occurring fluorescent *Pseudomonads* involved in suppression." 76(2): 181-185.
- Tong, W., et al. (2018). "Genomic insight into the taxonomy of *Rhizobium* genospecies that nodulate *Phaseolus vulgaris*." 41(4): 300-310.
- Whitman, W. B., et al. (2015). *Bergey's manual of systematics of Archaea and Bacteria*, Wiley Online Library.
- Woese, C., et al. (1984). "The phylogeny of purple bacteria: the alpha subdivision." 5(3): 315-326.
- Yoon, H. and T. J. B. Leitner (2015). "PrimerDesign-M: a multiple-alignment based multiple-primer design tool for walking across variable genomes." 31(9): 1472-1474.
- Zhao, H., et al. (2021). "*Vibrio marinisediminis* sp. nov., Isolated from Marine Sediment." 78(2): 810-815.