



Universidad de Chile

Facultad de Ciencias Químicas y Farmacéuticas

Departamento de Bioquímica y Biología Molecular

“Caracterización de plásmidos en metagenomas de microbiomas de aguas oceánicas de la Fosa de Atacama”

Memoria para optar al título profesional de
Bioquímico(a)

Angela Marina Reyes Tramolao

Director

Dr. Osvaldo Ulloa Quijada
Departamento de Oceanografía e
Instituto Milenio de Oceanografía
Universidad de Concepción

Patrocinante

Dr. Dante Miranda Wilson
Dpto. Bioquímica y Biología Molecular
Facultad Ciencias Químicas y
Farmacéuticas
Universidad de Chile

Diciembre, 2022

i. Financiamiento

Agencia Nacional de Investigación y Desarrollo (Anid): Fondecyt 1191360,
“Microbial Oceanography of the Atacama Trench”.

Iniciativa Científica Milenio-Anid: Instituto Milenio de Oceanografía ICN12_019.

ii. Dedicatoria

A la amorosa memoria de mi padre Nefthalí.

Ahora eres tú la luz que ilumina mi vida.

iii. Agradecimientos

Estando en las puertas del final de esta etapa de mi vida, llega la hora de mirar los pasos dados, los momentos habitados, para reconocer en ellos la presencia y la huella de las personas que no han soltado mi mano en este camino, aquellas con las que he coincidido en tiempo, espacio y vivencias, aquellas que han nutrido de anécdotas mi vida, aquellas que han hecho más amoroso el recorrido, sobre todo cuando este no ha sido el más amable con esta estudiante. En este ejercicio recopilatorio me doy cuenta de la magnitud de mi gratitud, y en esta vasta emoción, no sé muy bien por donde comenzar a desentrañar el nudo de mi sensibilidad y mis afectos. Por lo tanto, el orden de mención en mi escritura no necesariamente implica un orden creciente o decreciente de afecto y gratitud por mi parte.

Así entonces, manifiesto mi agradecimiento y admiración por todos los profesores que vertieron sus saberes en esta mente mía, quiero creer que aquello ha dado frutos del agrado de los docentes, tanto en el pregrado como en el colegio. Quisiera mencionar a aquellos involucrados en esta etapa de tesis; así, mis gracias al profesor Dr. Dante Miranda, su apoyo fue fundamental cuando los problemas abundaron, su empatía me hizo sentir acompañada, apoyada y segura. Mis agradecimientos al Dr. Ulloa por darle espacio a las ideas de esta mente y por guiar el proceso investigativo.

Quiero agradecer a los integrantes del laboratorio, quienes nutrieron mi investigación con sus comentarios y sugerencias. Un especial gracias a la Mag. Paula Ruíz. No solo es ella una gran investigadora, un talento increíble, y una mujer de una entereza inspiradora, sino que además ha sido una luz para mí y esta tesis. No estaría en este punto sin su apoyo y colaboración. Así mismo, agradezco con gran ímpetu a la Dra. Carolina González y a la Mag. Francisca Olivares, ambas contribuyeron con sus saberes científicos a esta estudiante, e inspiraron mi corazón con la fortaleza de sus valores. Extiendo, además, mi agradecimiento a FemenIMO, por el trabajo que realizan en pro de la igualdad de género y en contra de las violencias de género en el espacio científico académico. Son ellas un ejemplo a seguir, pues un espacio seguro para todas y todos no está, nada menos, que a la altura de la noble profesión que desempeñamos en ciencias.

¿Qué sería de mi paso universitario sin los amigos que me han acompañado? Muy probablemente, no podría estar ya en esta etapa sin la compañía de varios de ellos. Un

gran gracias a Camila A. por ser mi partner. Hiciste de los trabajos y de la vida universitaria algo más agradable. Estoy agradecida por haber coincidido contigo y por qué no me hiciste falta cuando mi corazón se desvaneció. Del mismo modo, estaré yo siempre con mi corazón dispuesto a acompañar el tuyo, amiga. Un reconocimiento a los amigos que aún conservo del primer año, a Valeria V., Claudio D., Dominique J y Camila M. Son innumerables las anécdotas que nos unen, las risas que compartimos y el cariño que tengo por cada uno de ustedes. Agradezco además a mi amiga Paula T. por su compañía y apoyo en esta última etapa, por los momentos vividos y las canciones entonadas, espero poder siempre compartir, entre contralto y soprano, y que nunca nos falten canciones. A mi amigo Martin B., con quien nos hemos acompañado en estos últimos pasos de tesis y en pasos de baile también. A mi amigo Bastián Q., por siempre estar dispuesto a escucharme y a ofrecerme palabras de aliento, por su amistad que perduró más allá de su paso por la bioquímica.

A toda la generación del 2016, pues fueron todos muy buenos compañeros conmigo, tanto en estudios como en carretes, y espero seamos buenos colegas. Un agradecimiento a todos los espacios que me acogieron, a la Batupobre que nació en 2017, en un paro, y que murió cuando perdimos a nuestro amigo José H.; al CEBQ del 2019; a las chicas del paro feminista del 2018 y la toma feminista del 2019, son todas tremendas, y con ellas aprendí que cambiar el mundo requiere una entrega de tiempo, cuerpo y alma, que no hay espacio para la indiferencia ante las violencias de género, pues tal cosa no nos llevará a ningún puerto. A nunca callar ante la desigualdad y la violencia de género. Por último, agradezco al coro de la facultad, que le abrió un espacio a mi voz y me ha permitido explorar nuevos desafíos y melodías.

Finalmente, expreso mi gratitud y amor por mi familia. A mi madre Rosa por su cariño tierno y sin límites, a mis hermanos Boris, Gonzalo y Francisco, por su cariño constante. A ustedes las gracias, por la fortaleza demostrada cuando los cimientos de nuestro hogar se desmoronaron tras la muerte de nuestro papá, por multiplicar amor cuando nos faltó un trozo de corazón. A mi perrito Jacko, que sin saberlo, lleno de ternura mis días y de amor mi corazón. A toda la familia por el acompañamiento y amor entregado, en momentos de alegría y en aquellos momentos tristes.

Quisiera agradecer el amor recibido en estos últimos años de carrera, de todas aquellas personas que me ofrecieron un hombro en el cual llorar, una palabra de aliento, y

todos aquellos abrazos que me hicieron sentir menos rota y sola, y que eventualmente me permitieron afrontar la vida después del duelo.

Les agradezco a todos con el ímpetu de mi corazón lastimado, que hoy se apresta a la vida con un pedazo en el cielo y otro aquí, con aquellos a quienes amo.

Papá:

He pensado repetidamente sobre qué escribirte, cómo hacerlo, de qué forma converger aquí los pensamientos, los anhelos y las penas. En esencia, esto no es otra cosa que un ejercicio para mi corazón, un intento de compartir contigo este hito de mi vida, obviando la separación que nos marca hoy. Tal ejercicio revela la crueldad de hacer concreta tu ausencia, a la que aún no me acostumbro y cuyas dimensiones no termino de asimilar. Una y otra vez me imaginé escribiendo esto, y en cada esfuerzo lloré mares, vastos y profundos. Hoy no es diferente. Hoy aún pienso en la brutalidad de la vida que nos separó y que me privó de tu compañía y amor en este día y en el resto de los días.

Quisiera, a pesar de todo, mostrarte que fui fuerte y valiente, que cumplí con mis deberes, que trabajo honradamente, que he sido amable y severa cuando corresponde, que mis afectos son honestos y que mis valores son piedra angular de mi actuar. Y quisiera que este trabajo simbolice avance, que signifique que la tristeza no apagó mi vida ni me detuvo en el tiempo. Es mi deseo ofrecerle y dedicarle este trabajo a tu memoria, porque fuiste motor cuando la pena consumió mi voluntad y desoló mi corazón, porque soñé presentándotelo orgullosa, de un logro que es también tuyo.

Así como pensé repetidamente en qué escribir, también te he soñado repetidamente orgulloso, me he soñado en tus pupilas que ven a su única hija terminando sus estudios, los que posibilitaste a puro trabajo extenuante y honesto, he soñado tu abrazo cálido y completo, he imaginado tus manos amorosas envolviendo las mías, tu cariño ilimitado e inmenso. Y soñar no es suficiente, no alcanza a endulzar la amargura de la pérdida, no amansa la violencia de la añoranza. Hoy me haces falta, aún.

Pero prometo seguir avanzando con ímpetu y valientemente, prometo llevarte en cada paso, prometo recordarte y soñarte tiernamente.

Gracias el arduo trabajo que hiciste y que me permite hoy estudiar.

Te amo papá.

iv. Índice contenidos

Financiamiento	i
Dedicatoria	ii
Agradecimientos	iii
Índice de contenidos	vii
Índice de figuras y tablas	x
Abreviaturas	xi
Resumen	xii
Abstract	xiii
1. Introducción	1
El ambiente hadal	1
Efectos de la alta presión hidrostática en los sistemas biológicos	1
La vida procarionte a altas presiones hidrostáticas	2
Plásmidos como agentes efectores de la transferencia horizontal de genes	3
Plasmidios en ambientes marinos	4
Plasmidios en el océano profundo	5
Fosa de Atacama como sitio de estudio de plasmidios y preguntas científicas asociadas	6
2. Propósito del estudio	8
3. Objetivos general y específicos	9
4. Materiales y métodos	10
5. Resultados	13
5.1 Identificación de contigs plasmidiales	13
5.2 Anotación genómica	15
5.3 Análisis comparativo	16
5.3.1 Abundancias de categorías COG en los plasmidomas	16

5.3.2 Abundancias de genes asociados a proteínas en cada categoría COG	17
5.3.3 Genes asociados a proteínas en la categoría de metabolismo y transporte de carbohidratos	21
5.3.4 Genes asociados a proteínas en la categoría de metabolismo y transporte de lípidos	22
5.3.5 Genes asociados a proteínas en la categoría de metabolismo y transporte de aminoácidos	23
5.3.6 Genes asociados a proteínas en la categoría de metabolismo y transporte de iones inorgánicos	23
5.3.7 Genes asociados a proteínas en la categoría de tráfico intracelular, secreción y transporte vesicular	25
6. Discusión	26
6.1 Identificación de secuencias plasmidiales en la Fosa de Atacama: sobre su presencia, tamaño y otras características	26
6.2 Caracterización funcional de las secuencias plasmidiales: descubriendo el contenido genético de los plasmidios de la Fosa de Atacama	30
6.2.1 Metabolismo y transporte de carbohidratos en plasmidios de la Fosa de Atacama	32
6.2.2 Sobre el metabolismo y transporte de lípidos en plasmidios de la Fosa de Atacama	35
6.2.3 Sobre el metabolismo y transporte de aminoácidos en plasmidios de la Fosa de Atacama	37
6.2.4 Sobre el metabolismo y transporte de iones inorgánicos en plasmidios de la Fosa de Atacama	38

6.2.5 Tráfico intracelular, secreción y transporte vesicular: posibilidad de THG por conjugación de plasmidios en la Fosa de Atacama	42
6.2.6 Moviloma: profagos y transposones, otras vías de THG en la Fosa de Atacama	43
6.2.7 Sobre la biogénesis de pared, membrana y envoltura celular en los plasmidomas de la Fosa de Atacama	44
6.3 Desafíos y proyecciones del estudio de plasmidios en ambientes oceánicos profundos	45
7. Conclusiones	46
8. Referencias	49

v. Índice tablas y figuras

Tablas

Tabla 1: Contigs plasmidiales identificados en los distintos plasmidomas	13
Tabla 2: Tamaño de los contigs plasmidiales en las distintas profundidades	14
Tabla 3: Caracterización de los contigs plasmidiales hallados en los distintos plasmidomas	15
Tabla 4: Cantidad de genes anotados y proteínas hipotéticas por plasmidoma	16

Figuras

Figura 1: Mapa de locación de muestreo	10
Figura 2: <i>Heatmap</i> de abundancias relativas de las categorías COG por cada plasmidoma	17
Figura 3: Abundancia relativa de genes asociados a proteínas en cada categoría COG	18
Figura 3: (Continuación)	19
Figura 3: (Continuación)	20

vi. Abreviaturas

SAGs: Acrónimo del inglés *Single Amplified Genomes*

THG: Transferencia Horizontal de Genes

Número EC: Del inglés: *Enzyme Comisión Number*

COG: Acrónimo del inglés *Clusters of Orthologus Genes*

ORF: Acrónimo del inglés *Open Reading Frame*

AMR: Acrónimo del inglés *Antimicrobial Resistance*

CDS: Acrónimo del inglés *Coding sequences*

TRAP: Acrónimo del inglés *Tripartite ATP-independent Periplasmic transporters*

ABC: Acrónimo del inglés *ATP-binding Cassete*

HAAT: Acrónimo del inglés *Hydrophobic Amino Acid uptake Transporter*

RND: Acrónimo del inglés *Resistance-nodulation-cell División*

HME-RND: Acrónimo del inglés *Heavy Metal Efflux RND*

SST4: Sistema de Secreción Tipo 4

SBP: Acrónimo del inglés *Substrate-binding Protein*

HME: Acrónimo del inglés *Heavy Metal Efflux*

Nramp: Acrónimo del inglés *Natural resistance-associated macrophage protein*

vii. Resumen

Las fosas hadales representan los ecosistemas más recónditos del océano, con profundidades máximas de hasta 11000 metros, y donde la alta presión hidrostática es una de las características más distintivas que, junto con otras variables ambientales, impulsa los cursos evolutivos y adaptativos de los organismos procariontes que habitan estos ambientes. La transferencia horizontal de genes es un fenómeno que permite el intercambio de material genético entre los miembros de una comunidad, mediante distintos mecanismos, y donde los plasmidios son importantes determinantes, y por ende, agentes relevantes en la dinámica de los procesos adaptativos y evolutivos. En el presente trabajo se estudia y caracteriza los plasmidios presentes en metagenomas obtenidos a través de la columna de agua de la Fosa de Atacama, tanto desde la zona hadal (8081, 8079 y 7928 metros de profundidad), como de la zona mesopelágica (500 m) y batial (1000 y 2000 m). Los análisis incluyen la identificación de secuencias plasmidiales dentro de los metagenomas y anotación genómica, los cuales revelan presencia efectiva de plasmidios en los metagenomas en estudio, junto con una importante representación de funciones celulares en las secuencias encontradas. En detalle, el transporte y metabolismo de iones inorgánicos es la función celular más abundante en general, con buena representación incluso en los plasmidomas hadales. Además, el transporte y metabolismo de carbohidratos y lípidos está más presente en los plasmidomas hadales, en contraste con los plasmidomas de aguas menos profundas, mientras que la biogénesis de membranas celulares presenta una abundancia relevante en todos los plasmidomas. Otras funciones celulares de interés como el transporte y metabolismo de aminoácidos, sistemas de secreción y transporte vesicular y aquellas relacionadas al moviloma, muestran caracteres importantes para el metabolismo, la transferencia horizontal de genes y la plasticidad genómica.

viii. Abstract

Hadal trenches are one of the most remote ecosystems in the ocean, with maximum depths of up to 11000 meters, and where the high hydrostatic pressures are a distinctive characteristic which, together with other environmental variables, impulses the adaptative and evolutive curses of prokaryotic organisms inhabiting the ocean's depths. The horizontal gene transfer allows the exchange of genetic material between the members of a microbial community, making use of different mechanisms in which plasmids play an important role, and therefore, they're relevant in adaptative and evolutive processes. The present study identifies and characterizes plasmidal sequences from metagenomes from Atacama Trench's water column, including the hadal zone (7928, 8079 and 8081 meters of depth), the bathyal zone (1000 y 2000 m), and the mesopelagic zone (500 m). The results reveal effective presence of plasmidal sequences in the depths analyzed, along with the finding of strong representation of certain cellular functions codified in the sequences found. For instance, the transport and metabolism of inorganic ions is the most abundant function overall, with notorious presence even in the hadal plasmidomes. Furthermore, the transport and metabolism of carbohydrates and lipids were more present in the hadal plasmidomes than on the less deep plasmidomes, while the membrane biogenesis shows a relevant abundance in all plasmidomes. Other functional categories of interest like the transport and metabolism of amino acids, secretion and transport systems, and mobilome related, show important traits for metabolism, horizontal gene transfer and genome plasticity.

1. INTRODUCCIÓN

El ambiente hadal.

El océano es el ecosistema más extenso de la biosfera terrestre, dentro del cual la zona hadal, que comprende profundidades entre los 6.000 y 11.000 metros, constituye el ~1% del suelo marino global y el ~45% de las profundidades verticales más recónditas del océano, implicando su importante posición en el ecosistema oceánico [1-3]. La mayoría de las fosas hadales se concentran en el Océano Pacífico, entre las que se cuentan las más profundas, como la Fosa de las Marianas, Tonga, Kermadec, entre otras [4]. Las fosas o trincheras hadales se forman a partir de la subducción de placas tectónicas, lo que les otorga características topológicas de importancia, esto es; su ubicación en los límites de las placas, una conformación tipo V donde el suelo profundo está delimitado por una pendiente continental y una pendiente asociada a la planicie abisal adyacente, conformación que resulta, además, en un área elongada sobre el suelo oceánico [1]. Estas características implican una alta heterogeneidad en los hábitats existentes dentro de la zona hadal. Además, la zona hadal presenta condiciones físicas, químicas e hidrológicas extremas y propias; tales como temperaturas entre 1°C – 2.5°C; ausencia de luz solar y, por tanto, carencia de producción primaria fototrófica [1, 5, 6]; el oxígeno disuelto registrado en la Fosa de las Marianas varía entre 150-160 uM, similar a la zona abisal; del mismo modo, en esta fosa se registran concentraciones de aniones inorgánicos NO^{-3} , NO^{-2} , $(\text{PO}_4)^{-3}$ similares entre la zona hadal y abisal [6]; por su parte, la materia orgánica que suele disminuir sus concentraciones de forma inversamente proporcional a la profundidad, halla un depocentro para su acumulación en las fosas oceánicas, ya sea por sedimentación natural, transporte lateral desde las paredes de la trinchera debido a terremotos, o por efecto de las corrientes oceánicas, influyendo en la abundancia, alta biomasa, diversidad, y elevada actividad microbiana de los microorganismos de la zona hadal [7-10]. Las presiones hidrostáticas presentes en la zona hadal pueden alcanzar hasta los 110 MPa, esto es, unas 1000 veces la presión en superficie, siendo este parámetro el más distintivo de las zonas hadales sobre los hábitats más someros y el que más impacto tiene sobre los organismos que le habitan [1].

Efectos de la alta presión hidrostática sobre sistemas biológicos.

A saber, las altas presiones afectan a los sistemas biológicos por su influencia en los equilibrios químicos y volúmenes de reacción y activación, lo cual haya razón de ser en el principio de LeChâtelier, que indica que ante perturbaciones por alta presión el equilibrio

químico se modifica para compensar tal perturbación, priorizando estados que ocupen menor volumen; así también, acelera aquellos procesos donde el estado de transición tiene menor volumen que el estado inicial [11]. Asimismo, las altas presiones modifican las distancias intermoleculares y afectan las conformaciones y estructuras supramoleculares de las moléculas, perjudicando la función biológica de estas [12]. En lo relativo a los ácidos nucleicos y sus funciones asociadas, las altas presiones hidrostáticas estabilizan los puentes de hidrogeno en las moléculas de ADN, dificultando la transición de la hebra doble a la simple que es necesaria para iniciar los procesos de replicación, transcripción y posterior traducción, lo que se resume en una disminución de la síntesis de ácidos nucleicos y proteínas [13]. Otro componente sensible a los efectos de las altas presiones son las membranas lipídicas, las cuales son una matriz compleja que reúne diversas moléculas lipídicas, además de proteínas y carbohidratos, que cumplen distintas funciones como reconocimiento celular y transporte [14]. Ante elevadas presiones, los lípidos de membrana cambian su conformación y se empaquetan disminuyendo el volumen celular, lo cual conlleva a una pérdida de fluidez y de permeabilidad al agua y otras moléculas, y una disminución en las interacciones funcionales entre proteínas y lípidos [14]. Tales efectos se ven reflejados en la pérdida de motilidad celular y disminución de transporte transmembrana de solutos [14]. Por su parte, las proteínas y las estructuras multiméricas también disminuyen su volumen a modo de adaptación a la alta presión, lo cual consiguen cambiando su conformación. No se ha descrito denaturación de proteínas por altas presiones en el rango de presiones que se puede hallar en los biotopos que presentan altas presiones hidrostáticas. No obstante, las modificaciones conformacionales pueden dificultar la asociación de monómeros y el acceso a sitios catalíticos, viéndose perjudicada la función de la proteína en procesos catalíticos, metabólicos y estructurales [12, 15]. Así entonces, muchos procesos y estructuras esenciales son gravemente afectados cuando los organismos son sometidos a altas presiones, pudiendo incluso gatillar la muerte celular [16].

La vida procarionte a altas presiones hidrostáticas.

Aquellos organismos que son capaces de sobrevivir las altas presiones son denominados piezófilos, los cuales se clasifican según las presiones a las cuales presentan sus tasas de crecimiento máximo. De esta manera, se describen organismos piezotolerantes, piezófilos e hiperpiezófilos, los cuales requieren presiones menores a 10 MPa, entre 10 y 50 MPa, y mayores a 50 MPa, respectivamente [17]. Decenas de organismos procariontes piezófilos obligatorios han sido aislados desde ambientes hadopelágicos, la mayoría de ellos pertenecen a los géneros *Colwellia* [18-20], *Shewanella*

[21, 22], *Psychromonas* [23, 24] y *Moritella* [21, 25], los cuales forman parte de una comunidad bacteriana distintiva de la zona hadal, dominada por gamma-proteobacterias heterotróficas y bacteroidetes [6]. Tales organismos, que están expuestos a condiciones desafiantes, se ven en la necesidad de generar adaptaciones como las antes mencionadas, para lo cual, fenómenos como la transferencia horizontal de genes (THG) son claves a la hora de extender determinantes genéticos ventajosos entre los organismos y comunidades.

Plásmidos como agentes efectores de la transferencia horizontal de genes.

La transferencia horizontal de genes es un proceso importante que ha dado forma a los procesos evolutivos de los organismos procariontes, permitiéndoles adaptarse a distintas condiciones ambientales a través del intercambio de ADN entre organismos, sin necesariamente tener estos un ancestro común [26]. El estudio de los elementos genéticos móviles, en conjunto llamados “el moviloma”, da luces acerca de los procesos de adaptación, especiación y ecología de los organismos [27]. Como parte del moviloma, los plásmidos destacan en el ámbito de la evolución de los microorganismos al ser los mayores determinantes de la transferencia horizontal de genes siendo, por ende, relevantes en la dinámica de los procesos adaptativos dentro y entre comunidades [27]. Los plásmidos son elementos genómicos extra-cromosomales presentes tanto en organismos procariontes como en eucariontes; capaces de replicarse de manera independiente al cromosoma celular [28, 29]. Los plásmidos son diversos en distintos ámbitos, se les puede encontrar de distintos tamaños; números de copia; contenido GC; contenido genético; distinta topología, esto es, si son lineales o circulares; distintos mecanismos de replicación; métodos de transmisión; espectro de hospederos celulares; entre otros [26]. Estos elementos genómicos pueden contener una gran variedad de genes que propician la persistencia del plásmido dentro del hospedero celular y facilitan la sobrevivencia del hospedero en amplios escenarios ambientales [30]. A la fecha, los plásmidos bacterianos han sido ampliamente estudiados, permitiendo avances en distintas áreas como la biotecnología, la microbiología, medicina, entre otras. Para expandir los nichos ecológicos de los hospederos, los plásmidos codifican distintos genes que generan fenotipos capaces de metabolizar distintos compuestos y degradar compuestos tóxicos. Los genes que tienen por función permitir la transmisión vertical y horizontal del plásmido, como los sistemas de conjugación y replicación plasmidal, se consideran elementales dentro del contenido génico plasmidal. Algunos plásmidos pueden contener genes adicionales que facilitan la transmisión y estabilidad dentro del hospedero como lo son aquellos genes que codifican proteínas de anti-restricción, además de sistemas de matanza post-segregacional y de sistemas

CRISPR-Cas tipo IV [31-33]. Los genes de virulencia y de resistencia a antibióticos también sirven al propósito de ampliar el nicho del organismo hospedador, y han sido ampliamente estudiados por su implicancia en patologías clínicas [30].

Los plásmidos son transferidos de una célula a otra a través de distintos mecanismos; esto es, conjugación, transformación, transducción [34, 35], los cuales son los mecanismos más caracterizados a la fecha; y la captación de vesículas de membrana [36] y conducción [37]. De estos métodos, la conjugación es el principal en cuanto a movilización de plásmidos y es el único que requiere establecer contacto físico entre los participantes de la transferencia, contacto mediado por un pilus, permitiendo así la transferencia entre organismos que pueden estar bastante distanciados filogenéticamente [34]. Los plásmidos conjugativos han sido motivo de estudio continuo en ámbitos clínicos por ser vehículos de genes de resistencia a antibióticos [34]. La movilización conjugativa, por su parte, permite la transferencia de plásmidos que carecen de los genes necesarios para la maquinaria de conjugación, tomando ventaja de las maquinarias de conjugación de otros elementos móviles conjugativos [38]. Los distintos mecanismos de transferencia de plásmidos pueden tomar lugar dentro de las comunidades bacterianas, sustentando la transferencia horizontal de genes, que propicia la variabilidad y plasticidad genética, y da paso al surgimiento de fenotipos favorables respecto del ambiente que habitan estos organismos.

Plásmidos en ambientes marinos.

Los ambientes naturales y las comunidades de microorganismos que los habitan no son ajenos a los efectos de este fenómeno; así, los plásmidos han sido fuente de estudio en diversos ecosistemas como suelo, biopelículas, ambientes acuáticos, entre otros [39-42]; tanto en análisis cultivo dependientes como independientes, permitiendo este último un mejor entendimiento de la distribución y presencia de plásmidos dentro de comunidades microbianas. En la era de la genómica, el estudio de plásmidos en metagenomas o plasmidomas se ha abierto a un sinnúmero de nuevas interrogantes y alternativas de análisis en el ámbito de la investigación de elementos genómicos móviles. De este modo, la llegada de las “ómicas” ha impulsado las investigaciones en los ecosistemas logrando ampliar las posibilidades de estudio en esta materia [43]. En consecuencia, estudios han logrado identificar, caracterizar genómicamente y analizar filogenéticamente plásmidos de organismos marinos procariontes y eucariontes, valiéndose de distintas estrategias experimentales y bioinformáticas. Sin ir más lejos, ha quedado en evidencia la presencia y diversidad de plásmidos en ambientes marinos en trabajos investigativos. Así, en un estudio de Paulsen y Palenik (2012) [44], mediante análisis de metagenomas de cianobacterias de

aguas costeras en California, las cuales son abundantes en *Synechococcus* sp, y con datos de *Global Ocean Survey*, se identificaron los plásmidos presentes, que en su mayoría resultaron ser de corta longitud y crípticos, siendo las proteínas de replicación el principal contenido genómico, mientras que el segundo mostró un escenario más complejo, con genes involucrados en la transferencia, movilización, estabilidad y partición de plásmidos [44]. Además, otros estudios han logrado detectar nuevas familias de plásmidos en las comunidades microbianas de ambientes marinos [45]. Otro aspecto de interés han sido los grupos bacterianos de importancia en los ambientes acuáticos, como el clado de las *Roseobacter*, cuyos plásmidos suelen contener genes clave para los ciclos bioquímicos celulares así como también para la degradación de compuestos aromáticos [46]. En esta misma línea, dentro de la familia de las *Rhodobacteraceae*, los plásmidos tipo RepA-I son predominantes, a menudo acarreado genes importantes para la formación de biopelículas [47]. Otro grupo de plásmidos presente en diversos miembros de esta familia, RepL-I, han sido analizados en otro estudio, a nivel filogenético, funcional y geográfico, logrando detectar un plásmido con 100% identidad en bacterias distantes tanto geográficamente como filogenéticamente, evidenciando el potencial de este plásmido de ser transmitido entre bacterias de distintos órdenes [48], abriendo la discusión sobre el impacto de este tipo de fenómenos sobre la evolución y adaptación ecológica de los organismos. De igual forma, distintos trabajos han identificado determinantes genéticos en plasmidios que otorgan resistencia a antibióticos y otras características de importancia como clústeres fotosintéticos, clústeres asociados a formación de biopelículas, manejo de estrés oxidativo, distintas vías metabólicas, entre otros [39, 49-51]. El análisis de la filogenia de estos plásmidos ha permitido dilucidar el proceso evolutivo de genes y procesos celulares [48, 52], así como también del organismo en sí mismo [53]; junto con determinar los mecanismos de transferencia horizontal predominante en el grupo bacteriano sujeto a estudio [54].

Plasmidios en el océano profundo.

A la fecha, en el caso de los ambientes hadales, la Fosa de las Marianas concentra la mayor parte de los trabajos relativos a la microbiología hadal, mientras que otras fosas oceánicas como la de Puerto Rico o Yap están también siendo investigadas. En lo concerniente a plasmidios en estos ambientes, éstos se han detectado en trabajos de obtención de genomas completos de organismos aislados desde el océano profundo. Tal es el caso de *Deinococcus* sp. D7000, organismo aislado desde la región Challenger Deep en la Fosa de las Marianas, a 7000 metros de profundidad, el cual cuenta con 9 plasmidios distintos además de un cromosoma [55]; o *Paraoceanicella profunda*, un organismo

caracterizado por primera vez, aislado desde la Fosa de las Marianas a 10.890 metros de profundidad, que revela un cromosoma y seis plasmidios, los cuales van desde 112.235 pb a 586.520 pb [56], por nombrar algunos. Por otro lado, algunos estudios de caracterización genómica de organismos hadales revelan genes asociados a la homeostasis plasmidial; ejemplo de esto es un análisis de múltiples SAGs (del inglés *single amplified genomes*) de *Roseobacter* y *Alteromonas* recuperados desde la Fosa de las Marianas a 8289 metros de profundidad, que expuso la existencia de agrupamientos genéticos relacionados a distintas funciones plasmidiales como replicación plasmidial y estabilización de plasmidios [57].

En suma y a la fecha reciente, las investigaciones orientadas a la identificación de plásmidos y su caracterización genética en ambientes marinos han nutrido de luz a las incógnitas que la transferencia horizontal trae consigo, siendo este fenómeno una fuerza capaz de moldear las comunidades en su camino adaptativo y evolutivo, y que encuentra en los plásmidos un elemento efector, que resulta igualmente complejo e intrigante. No obstante, es también cierto que el estudio de plásmidos en los ambientes marinos aún no se ha extendido hacia el océano profundo, lo que principalmente se atribuye a la dificultad técnica que suponen las tomas de muestras para las investigaciones.

Fosa de Atacama como sitio de estudio de plasmidios y preguntas científicas asociadas.

En la línea de ambientes hadales a explorar, la Fosa de Atacama también ha suscitado mayor interés en los últimos años, lo que se ve reflejado en una mayor cantidad de trabajos asociados a la zona, tanto en el ámbito de geología, fauna hadal, microbiología, entre otros. La Fosa de Atacama, también conocida como Fosa Perú – Chile, se origina de la subducción de la placa de Nazca bajo la placa Sudamericana, y se ubica en el sureste del océano Pacífico, extendiéndose desde el norte de Perú hasta la zona central-sur de Chile, 6° 00' S, 81° 50' W a 39° 00' S, 75° 00' W; presentando una longitud máxima de 5900 km, ancho de 64 km y una profundidad máxima de 8081 metros, siendo la fosa más extensa del océano pacífico sur [58]. Las aguas someras de la zona de la Fosa de Atacama son peculiares en cuanto son altamente productivas, además de eutróficas, por efecto del sistema de corrientes de Humboldt [59, 60], además de alojar una voluminosa zona de mínimo oxígeno [61]. Recientemente, en un trabajo nacional, se ha esclarecido la composición taxonómica de las comunidades de la columna de agua de la Fosa de Atacama, dejando ver que los grupos taxonómicos más abundantes en la columna de agua son *Bacteroidetes*, *Candidatus Marinimicrobia*, *Alfaproteobacteria* (*Pelagibacterales*, *Esfingomonadales*), *Gammaproteobacteria* (Ca. *Thioglobus*, *Alteromonadales*, *Cromatiales* y *Oceanospirillales*), mientras que en la zona hadal predominan los organismos heterótrofos,

y fueron más abundantes los organismos *Planctomycetes*, *Betaproteobacteria* (*Burkholderiales*), y otras *Deltaproteobacterias*, respecto de su presencia en otras profundidades [62]. Si bien el mencionado estudio es de alta relevancia para la microbiología de las aguas de la Fosa de Atacama, el hecho es que lo mismo no se ha replicado en el ámbito de plasmidios, por cierto que a la fecha, no se han realizado estudios de caracterización de plasmidios o estudios donde estos elementos genómicos hayan sido identificados, lo que significa un vacío en lo relativo a la presencia y rol de plasmidios en los procesos adaptativos y evolutivos de los organismos procariontes que pueblan una región del océano profundo de naturaleza única. Sin embargo, esta carencia, vista con otra perspectiva, surge como una ventana amplia que permite el planteamiento de diversas interrogantes y da espacio a distintas posibilidades de investigación. De esta forma, se abren paso preguntas como: ¿Es posible hallar plasmidios en la Fosa de Atacama? ¿Cuáles serían las características y funciones presentes, de haber plasmidios?, ¿Es posible la transferencia horizontal de genes por medio de conjugación y/u otros métodos?, ¿Qué diferencias se podrían encontrar entre los plasmidios hallados en aguas más superficiales y las más profundas en esta zona? ¿Es posible atribuir esas diferencias a las distintas condiciones ambientales de la Fosa de Atacama?, por nombrar algunas. Esta curiosidad suscitada es el pilar sobre el cual plantear los objetivos de este trabajo exploratorio, el cual podría brindar una primera visión sobre el conjunto de plásmidos de las comunidades bacterianas en la Fosa de Atacama, en comprensión de las variables ambientales inherentes de esta región oceánica que paulatinamente devela sus secretos.

2. PROPÓSITO DEL ESTUDIO

Se plantea en este proyecto hacer un estudio de caracterización de los plásmidos presentes en los microbiomas de algunos puntos de la columna de agua localizada en la Fosa de Atacama, también conocida como Fosa Chile – Perú. Así, el propósito principal es otorgar una perspectiva amplia respecto al conjunto de plásmidos presentes en las comunidades microbianas de este ambiente único, para contribuir al entendimiento de las dinámicas que la transferencia horizontal de genes tiene en las comunidades. En consideración de la carencia de trabajos relativos a plasmidios en la Fosa de Atacama, es también propósito del presente trabajo entregar un primer antecedente respecto de la presencia y caracteres genómicos contenidos en plasmidios, a modo de comenzar a comprender el rol de estos en la adaptación de los organismos y su evolución, en el contexto del ambiente hadal que habitan.

3. OBJETIVOS GENERAL Y ESPECÍFICOS

Objetivo General: Identificación y caracterización genómica de plásmidos en metagenomas de aguas oceánicas de la Fosa de Atacama.

Objetivo Específico 1: Establecer un protocolo para la identificación de plásmidos.

Objetivo Específico 2: Caracterizar funcionalmente los plásmidos de los metagenomas de aguas profundas de la Fosa de Atacama y de aguas superficiales.

Objetivo Específico 3: Establecer diferencias y similitudes entre aguas profundas y superficiales, en términos de función.

Los objetivos aquí presentados difieren en sus actividades propuestas presentadas en el Proyecto de Memoria de Título. Más en concreto, en el Objetivo Específico 3, los análisis comparativos relacionados a la similitud de secuencias, que fueron propuestos para llevarse a cabo tanto entre las mismas secuencias plasmidiales halladas en el Objetivo Específico 1, así como en contraste con bases de datos, a fin de conocer si las secuencias halladas han sido descritas previamente y dilucidar cuáles estarían presentes y su distribución en los distintos plasmidomas, no fueron llevados a cabo en consideración de la relevancia que fueron tomando los resultados de la anotación genómica; por lo cual, el camino que se siguió prioriza la comparación de las secuencias encontradas en el ámbito de las funciones en ellas codificadas.

4. MATERIALES Y MÉTODOS

Para la realización del presente trabajo, se dispuso de 6 metagenomas, previamente obtenidos de muestras de agua sacadas frente a las costas de Antofagasta (Figura 1), en las expediciones Atacamex y SO261 que tuvieron lugar entre febrero y marzo de 2018, procesadas en filtros de poro de 0,2 μm , a fin de separar las bacterias libres en la muestra de aquellas asociadas a partículas. Estos 6 metagenomas, ensamblados a través de metaSPAdes [63], corresponden a las muestras extraídas desde los 500, 1000, 2000, 7928, 8079 y 8081 metros de profundidad. Los metagenomas se agruparon en dos categorías, la primera denominada G1 que contine aquellos metagenomas meso-batipelágicos (500, 1000 y 2000 m), y la segunda, G2, donde se agrupan los metagenomas más profundos, los hadopelágicos (7928 m, 8079 m y 8081 m).

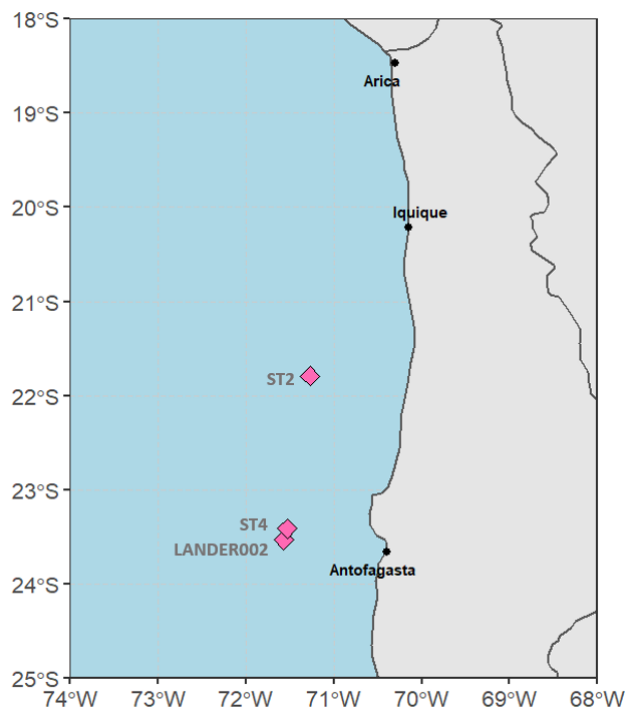


Figura 1| Mapa de locación de muestreo. Se identifican, mediante las figuras de diamantes, las distintas estaciones de muestreo. Desde *Lander002* se obtienen las muestras de 500, 1000, 2000 y 8081 metros; desde *ST4* la muestra de 8079; y desde *ST2* la de 7928 metros de profundidad.

Identificación de contigs plasmidales – Se utilizó el software Platon en Linux para la identificación y distinción de contigs plasmidales y cromosomales, en su configuración *default*. Platon basa su actividad en un análisis enfocado a las diferencias de distribución de replicones de genes codificantes de proteínas entre cromosomas y plasmidios, en su configuración *default* [64]. Así, se analizaron los 6 archivos FASTA correspondientes a los metagenomas mencionados. El software provee una caracterización de los contigs que se consideran plasmidiales, aportando información sobre la potencial circularidad, número de secuencias codificantes de ARNr, genes asociados a la replicación, movilización, y conjugación plasmidial, orígenes de transferencia OriT, además de información sobre grupos de incompatibilidad. Para todo lo recién mencionado, Platon recurre a Prodigal [65], Diamond [66], MUMmer [67], HMMER [68] e INFERNAL [69]; y realiza una búsqueda por similitud mediante BLAST+ usando la base de datos de plasmidios de NCBI [70]. Finalmente, de Platon se obtiene un archivo FASTA para los contigs plasmidiales y otro para los cromosomales, un archivo “.tsv” con información de la caracterización, y uno “.json” con los resultados e información completa sobre cada contig plasmidal. Los contigs plasmidiales obtenidos de cada metagenoma serán referidos de aquí en adelante como “plasmidomas”.

Anotación genómica – Para la identificación y etiquetamiento de los caracteres genómicos relevantes en las secuencias, se utilizó Prokka, también en Linux. Se sometieron a análisis los archivos obtenidos de Platon, específicamente, los 6 archivos que contienen las secuencias de los contigs plasmidiales en formato FASTA. Prokka coordina distintas herramientas para la detección y anotación de genes, tales como: Prodigal para las secuencias codificantes para proteínas [65], RNAmmer para los genes de ARNr [71], Aragorn para los ARN de transferencia [72], SignalP para los péptidos señal [73], e Infernal para los RNA no codificantes [69]. Prokka procede a la anotación de las secuencias codificantes CDS utilizando bases de datos de UniProt [74], RefSeq [75], Pfam [76] y TIGRFAM [77], a través de BLAST+ [70], y en cada una de las bases de datos se usa por defecto un *e-value* umbral de 10^{-6} . Si la búsqueda no arroja resultados para las CDS, se anota como “proteína hipotética”. Entre los archivos de salida, existe uno conteniendo todas las características halladas en el archivo de entrada, esto es: Locus de la secuencia anotada, tipo de secuencia, largo de la secuencia anotada, nombre del gen, número EC (del inglés: *enzyme comisión number*) [78], código COG (acrónimo del inglés: *Clusters of Orthologous Genes*) [79], producto del gen; archivo que se destina a la construcción de matrices de abundancia relativa a utilizar en los análisis posteriores.

Análisis Comparativo – Para los análisis siguientes se procedió a construir una matriz de abundancias relativas de genes anotados elaborada a partir de una normalización que considera los pares de bases plasmidiales asociados a un gen anotado (datos obtenidos del análisis con Prokka) sobre los pares de bases plasmidiales totales de cada plasmidoma. En la construcción de la matriz se deja fuera aquellas proteínas hipotéticas que no poseen datos de: nombre del gen, tipo de secuencia, EC number y código COG [78, 79]. Como análisis comparativo se construye un *heatmap* de abundancias relativas de las superfamilias COG por plasmidomas utilizando el paquete “pheatmap” en RStudio [80]. Además, se elabora un gráfico de burbujas con las abundancias relativas, en base a las distintas categorías COG, las cuales engloban los distintos procesos celulares, para los distintos plasmidomas. En la construcción del gráfico de burbujas se considera la asignación de Prokka por sobre el código COG, aunque se utilizan las categorías COG para su categorización. Así, no se suman aquellos genes con distinta asignación funcional que compartan el mismo código COG. Finalmente, se muestran los 5 genes asociados a proteínas más abundantes por categoría COG en el gráfico de burbujas.

5. RESULTADOS

5.1 Identificación de contigs plasmidiales.

En todos los metagenomas sometidos a análisis con Platon se identificaron contigs plasmidiales con distintos porcentajes de presencia y distintos tamaños (Tabla 1). Así, se revela que el porcentaje de pares de bases en contigs plasmidiales respecto al metagenoma correspondiente va desde a 0,029 a 0,81%. En consideración de la cantidad de contigs plasmidiales encontrados en los metagenomas de 1000 y 8081 m respecto a los demás, cantidad que resulta ser considerablemente más baja (al menos un orden de magnitud más bajo), no se incluyen a los plasmidomas de 1000 y 8081 metros de profundidad en los análisis siguientes por no ser estadísticamente confiables respecto a los demás. Así, se utilizan los plasmidomas de 500, 2000, 7928 y 8079 metros en los análisis siguientes.

Tabla 1 | Contigs plasmidiales identificados en los distintos metagenomas. Se informa la cantidad de contigs y los pares de bases totales correspondientes al metagenoma (gris claro) y aquellos identificados como plasmidiales (gris oscuro). Los pares de bases de plasmidomas se normalizan y se ofrece el porcentaje de ellas a cada profundidad.

	N° de contigs	Pares de bases totales	pb plasmidios / pb metagenoma	%
G1_0500	28113	221522341	0,0053	0,53
	59	1167841		
G1_1000	25621	213435033	0,00057	0,057
	7	122310		
G1_2000	28980	244660752	0,0020	0,20
	15	487276		
G2_7928	38260	362699890	0,0077	0,77
	136	2783180		
G2_8079	34851	328806277	0,0081	0,81
	112	2671303		
G2_8081	43880	344814162	0,00029	0,029
	12	101639		

De este modo, G2_8079 posee la mayor presencia de contigs plasmidiales, mientras que G1_2000 posee la menor cantidad. Respecto a los tamaños de los contigs identificados como plasmidiales en los distintos metagenomas, el promedio es de 197934 pb, 32485 pb,

20465 pb, 23851 pb, para las profundidades de 500, 2000, 7928 y 8079 metros, respectivamente; hallándose el contig más largo en G1_0500, con 280522 pb (Tabla 2).

Tabla 2 | Tamaño de los contigs plasmidiales en las distintas profundidades.

Se muestra, en pares de bases, los tamaños del contig más largo, el promedio del tamaño de los contigs de un plasmidoma y la desviación estándar para cada muestra.

	Tamaño contig más largo (pb)	Tamaño promedio (pb)	Desviación estándar del tamaño
<i>G1_0500</i>	280522	19793,9	196227,8
<i>G1_2000</i>	255192	32485,1	64958,27
<i>G2_7928</i>	179643	20464,6	33792,22
<i>G2_8079</i>	255498	23850,9	38808,49

Adicionalmente, como se muestra en la Tabla 3, los resultados de Platon muestran que en la mayoría de los contigs no se logró la circularización, en todos los plasmidomas. En cuanto a los marcos de lectura abiertos, estos son abundantes de forma proporcional a la cantidad de contigs hallados, de esta manera, el plasmidoma de 7928 metros, es en el que más marcos de lectura abiertos se identificaron. Este plasmidoma es también el que posee la mayor cantidad de genes asociados a la replicación plasmidial, seguido por el plasmidoma de 8079 metros. Continuando con las secuencias de orígenes de transferencia OriT, Platon identifica tres en total, repartidos en los plasmidomas de 500, 2000 y 7928 metros. Fueron identificados en todos los plasmidomas, genes asociados a la conjugación, de forma proporcional a la cantidad de contigs de los plasmidomas. También, se identifican genes de resistencia a antibióticos en los plasmidomas de 500, 7928 y 8079, y en este mismo orden se dan sus abundancias. No se registran genes asociados a ARNr ni movilización. La búsqueda en bases de datos por similitud realizada por Platon halla hits en todos los plasmidomas y el número de hallazgos es mayor en los plasmidomas con mayor cantidad de contigs.

Tabla 3 | Caracterización de los contigs plasmidiales hallados en los distintos plasmidomas. Se informa sobre circularidad; marcos de lectura abiertos (ORF, del inglés *open reading frame*); cantidad de genes asociados a la replicación, movilización, orígenes de transferencia (OriT), conjugación, resistencia a antibióticos (AMR, del inglés *Antimicrobial resistance*), ARN ribosomal; e hits con bases de datos.

	Circular	No circular	ORFs	Replicación	Movilización	OriT	Conjugación	AMRs	ARNr	Hits base de datos
G1_0500	4	56	1089	6	0	1	19	14	0	19
G1_2000	2	13	536	5	0	1	4	0	0	8
G2_7928	5	131	2697	56	0	1	36	10	0	76
G2_8079	2	110	2632	53	0	0	21	6	0	70

5.2 Anotación genómica

Los archivos fasta que contienen las secuencias de los contigs plasmidiales obtenidos por Platon fueron sometidos a análisis en Prokka para anotar los genes existentes. Se hallaron 9624 genes en total, de los cuales 1064 se presentan en el plasmidoma de 500 m, 525 en el de 2000 m, 5440 en el de 7928 m y 2595 en el de 8079 m. Se logró anotar 2836 genes codificantes para proteína de forma exitosa, mientras que 4062 genes quedan como hipotéticos. La búsqueda de genes no codificantes arroja 7 genes asociados a ARNt. Los genes que no codifican para proteína y se hayan sin predicción están en el apartado “Otros” (Tabla 4).

La mayor cantidad de proteínas predichas exitosamente por Prokka se encuentra en el plasmidoma de 8079 metros con 1141, mientras que la menor cantidad se halla en el de 2000 metros, con 109 proteínas predichas. Por otro lado, el plasmidoma de 7928 metros concentra la totalidad de los genes sin predicción, además de contener la mayor cantidad de genes codificantes para ARNt.

Tabla 4 | Cantidad de genes anotados y proteínas hipotéticas por plasmidoma.

El apartado “CDS” (acrónimo del inglés: *Coding sequences*) reúne a aquellos genes codificantes para alguna proteína calificadas como “Anotadas” e “Hipotéticas”, además de incluir el total por plasmidomas. Se informa la cantidad de secuencias identificadas como codificantes para ARNt, además de aquellos genes sin predicción en el apartado “Otros”.

	CDS			ARNt	Otros
	Anotadas	Hipotéticas	Total		
G1_0500	512	552	1063	1	0
G1_2000	109	416	525	0	0
G2_7928	1074	1642	2716	4	2720
G2_8079	1141	1452	2593	2	0
Total	2836	4062	6897		

5.3 Análisis comparativo

5.3.1 Abundancias de categorías COG en los plasmidomas

Como parte de los análisis comparativos a nivel de función que se plantearon, se ofrece en la Figura 2 un *heatmap* que muestra la abundancia relativa de las 22 categorías COG identificadas, según profundidad. Se nota la ausencia de la categoría B, correspondiente a “Estructura y dinámica de la cromatina”, la cual, en la base de datos, tiene dos definiciones correspondientes a histidina en organismos arqueas, y proteínas tipo histidina en organismos procarionte. Las categorías “W” de “estructuras extracelulares”, “Y” de “estructura nuclear” y “Z” de “citoesqueleto”, tampoco están presentes en los plasmidomas.

En el grupo G1, las categorías “C”, “E”, “G”, “L”, y “M” se presentan con destacada abundancia en el plasmidoma más superficial, mientras que en el plasmidoma de 2000m se imponen las categorías “P”, “O”, “C”. En los plasmidomas profundos son las categorías “G”, “L”, “P” y “M” las más abundantes. Las categorías “P” y “M” tienen una abundancia relativa importante en todos los plasmidomas.

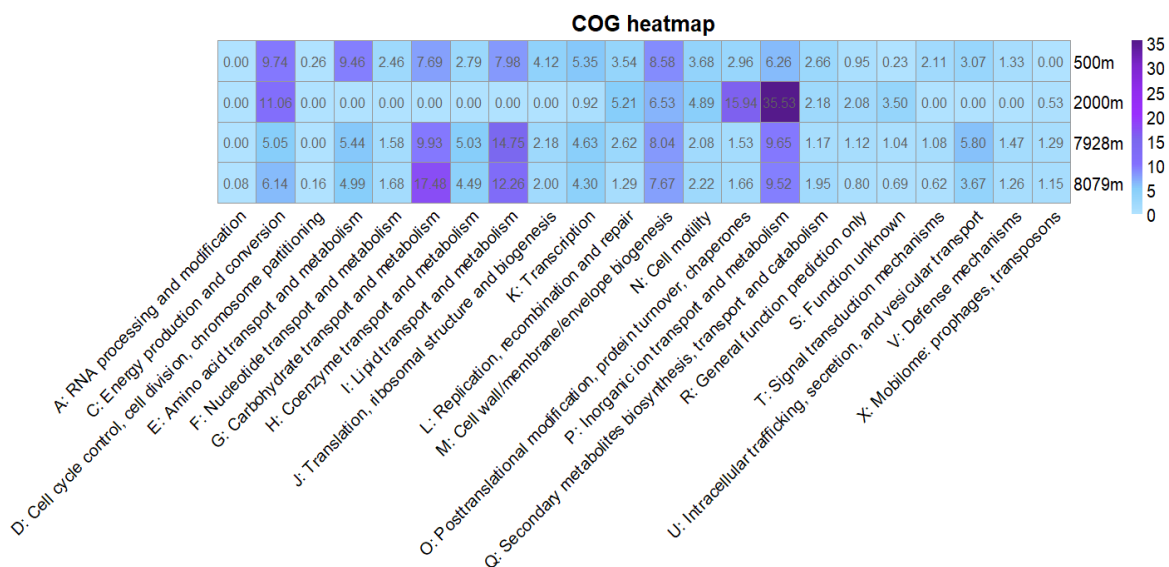


Figura 2| Heatmap de abundancias relativas de las categorías COG por cada plasmidoma. Las columnas muestran a las categorías que corresponden a superfamilias, según la nomenclatura COG. Las filas muestran los plasmidomas del más superficial al más profundo. Las abundancias relativas se muestran dentro de cada celda, las cuales están coloreadas según la guía de color adjunta.

En la anotación genómica surgieron genes asociados a proteínas que fueron asignados a más de una categoría COG, donde las más abundantes en la zona profunda corresponden a las categorías “DMP”, “EP”, “EQ”, “UN” y “OC”. No obstante, estas abundancias están muy por debajo de las halladas en las categorías troncales; valga por caso, comparando las abundancias más altas, la abundancia de la categoría “P: transporte y metabolismo de iones inorgánicos” es 60 veces más alta que la abundancia de la categoría compuesta “DMP”, aproximadamente, razón por la cual el enfoque es en las categorías troncales ilustradas en la Figura 3.

5.3.2 Abundancia de genes asociados a proteínas en cada categoría COG.

Para analizar y comparar en más detalle las características funcionales de los plasmidomas en estudio, se presenta un gráfico con las abundancias relativas de los 5 genes más abundantes por categoría COG (Figura 3). De las 22 categorías presentes, las que presentan menor número de genes asociados son las “A” (procesamiento y modificación de ARN) y “D” (control de ciclo celular, división celular y fraccionamiento de cromosomas), en concordancia con lo revelado en la sección 5.3.1. Dentro de las

categorías con mayor presencia, se desglosa, en el caso de la categoría “P” (metabolismo y transporte de iones inorgánicos), que el mayor aporte es de la proteína CusA del sistema de eflujo de cationes (COG3696) presente en el plasmidoma de 2000 metros. Así, en cada categoría se visualizan los aportes en abundancia de cada gen asociado a proteína.

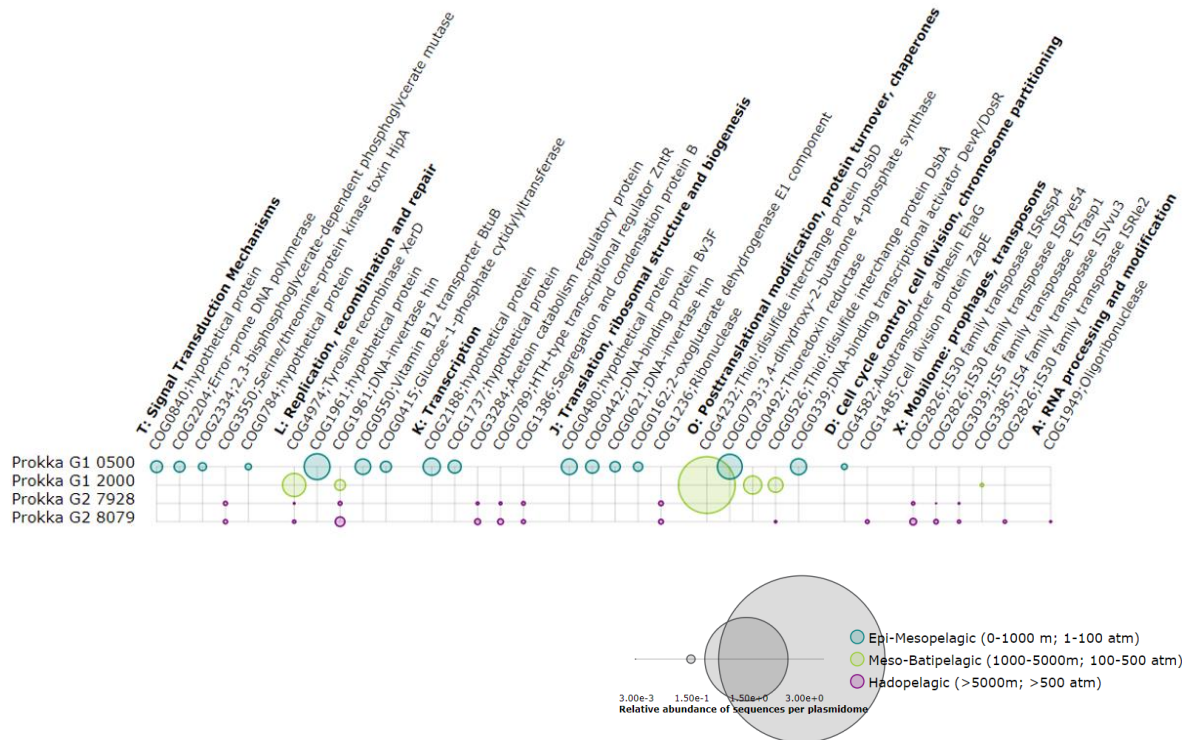


Figura 3 | Abundancia relativa de genes asociados a proteínas en cada categoría COG. Gráfico de burbujas que informa, por cada categoría COG identificada, los 5 genes asociados a proteínas más abundantes; luego desglosa las abundancias relativas por cada plasmidoma, donde el diámetro de cada burbuja es proporcional a dicha abundancia. Los colores de cada burbuja indican la zona de la columna de agua que representan los plasmidomas, esto es, epi-mesopelágico, meso-batipelágico y hadopelágico, según indica la leyenda.

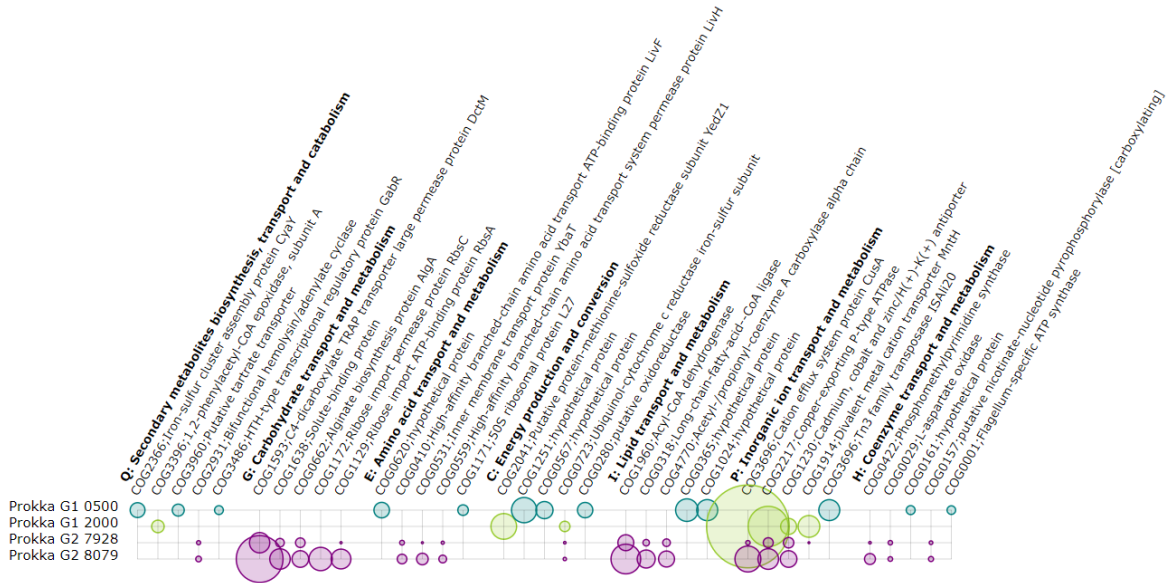


Figura 3 | Abundancia relativa de genes asociados a proteínas en cada categoría COG (continuación). Gráfico de burbujas que informa, por cada categoría COG identificada, los 5 genes asociados a proteínas más abundantes; luego desglosa las abundancias relativas por cada plasmidoma, donde el diámetro de cada burbuja es proporcional a dicha abundancia. Los colores de cada burbuja indican la zona de la columna de agua que representan los plasmidomas, esto es, epi-mesopelágico, meso-batipelágico y hadopelágico, según indica la leyenda.

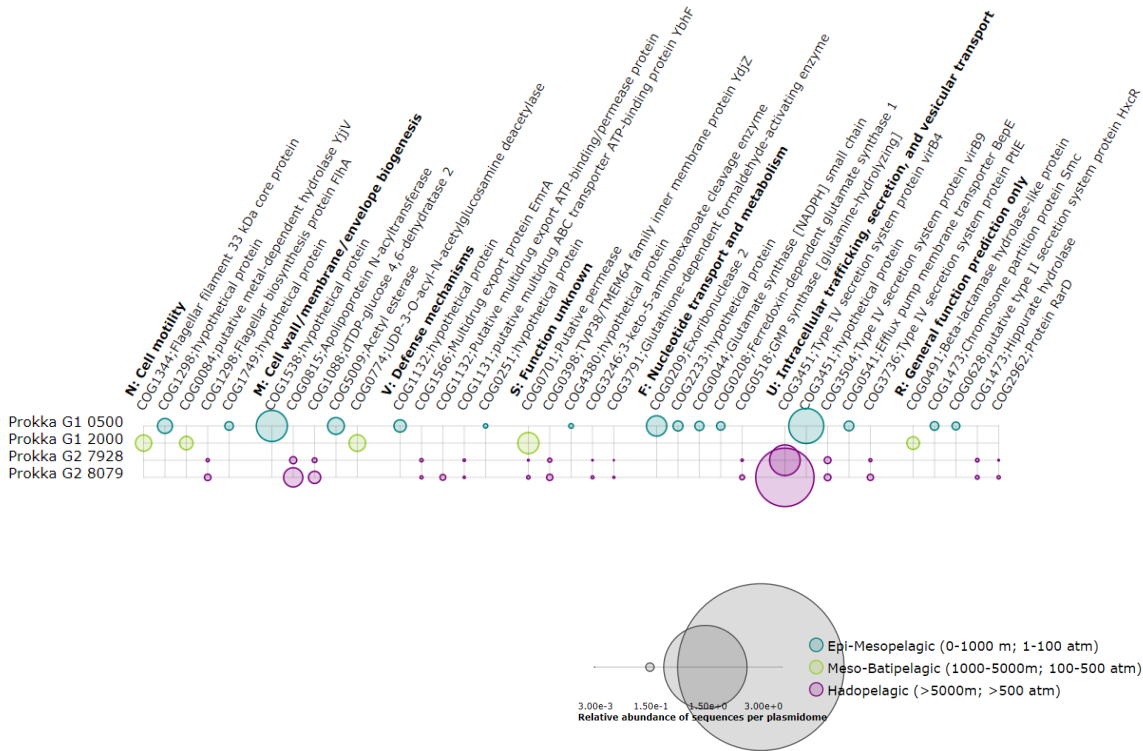


Figura 3 | Abundancia relativa de genes asociados a proteínas en cada categoría COG (continuación). Gráfico de burbujas que informa, por cada categoría COG identificada, los 5 genes asociados a proteínas más abundantes; luego desglosa las abundancias relativas por cada plasmidoma, donde el diámetro de cada burbuja es proporcional a dicha abundancia. Los colores de cada burbuja indican la zona de la columna de agua que representan los plasmidomas, esto es, epi-mesopelágico, meso-batipelágico y hadopelágico, según indica la leyenda.

A partir de los resultados exhibidos en las Figuras 2 y 3, y con la finalidad de profundizar en los caracteres genéticos más relevantes, se desglosa a continuación el detalle de los resultados obtenidos en las siguientes categorías COG: “G”, “I”, “E”, “P”, “X” y “U”. La elección de las primeras tres, es debido a la alta presencia de las proteínas predichas en la zona profunda, y su relación con el comportamiento de la materia orgánica; la elección de la categoría “P” es por ser la que mayor abundancia de genes asociados a proteínas recoge; mientras que las últimas dos han sido seleccionadas para ser revisadas en profundidad por su importancia en la transferencia horizontal de genes. Esto, siempre con especial interés y énfasis en la zona hadal.

5.3.3 Genes asociados a proteínas en la categoría de metabolismo y transporte de carbohidratos.

Según se ve en la Figura 3, en el apartado de la categoría “G: transporte y metabolismo de carbohidratos”, es evidente la gran representación de esta categoría en los plasmidomas más profundos. Esta representación está distribuida entre los genes: Proteína permeasa grande DctM del transportador TRAP para C4-dicarboxilato (COG1593); Proteína de Unión a Solutos (COG1638); Proteína de Biosíntesis de Alginato AlgA (COG0662); Proteína Permeasa RbsC de Importe de Ribosa (COG1172); Proteína de Unión de ATP para el Importe de Ribosa RbsA (COG1129).

El gen más abundante de esta categoría corresponde a la proteína permeasa grande DctM del transportador TRAP (acrónimo del inglés *tripartite ATP-independent periplasmic transporters*) para C4-dicarboxilato. El transportador TRAP para C4-dicarboxilato posee en su estructura a DctM, una proteína de 12 hélices transmembrana [81], que forma parte del sistema DctPQM caracterizado en *Rhodobacter capsulatus* [82]. DctM sería parte de la traslocación del sustrato a través membrana, pudiendo ser este, fumarato o malato [83]. Este gen asociado a proteína es hallado en abundancia en el grupo G2, más específicamente en el plasmidoma de 8079 metros.

Adicionalmente, se predice la proteína AlgA, una enzima monomérica con función fosfomanosa isomerasa y GDP-manosa pirofosforilasa [84], ambas funciones cruciales en la biosíntesis de alginato, un polisacárido que se conforma por distintas proporciones de ácido beta-D-manurónico unido por enlace 1,4-glucosídico al ácido alfa-L-gulurónico y que forma geles cuando está en medios con cationes divalentes, tales como Ca⁺², Sr⁺², Ba⁺² [85]. La mencionada proteína predicha se concentra más en el plasmidoma de 7928 m.

Por otra parte, figura dentro de los genes más abundantes en esta categoría un transportador para ribosa. El importador de ribosa bacteriano es parte de la superfamilia de transportadores ABC (acrónimo del inglés *ATP-binding cassette*) y se compone de tres proteínas, una proteína citosólica tipo ABC con dominios de unión a nucleótidos llamada RbsA [86], un homodímero de transmembrana RbsC [87], y una proteína de unión a sustrato en el periplasma RbsB [88, 89]. Más en específico, en el grupo G2, se identifican las proteínas predichas RbsA y RbsC, la primera con mayor abundancia relativa en el plasmidoma de 8079 y la última solo con presencia en dicho plasmidoma.

Así, estos genes asociados a proteínas están orientados al transporte de solutos que participan luego en el metabolismo central del carbono o en el metabolismo de glúcidos, y a la síntesis de biopolímeros, dejando ver su abundancia en los plasmidomas más profundos.

5.3.4 Genes asociados a proteínas en la categoría de metabolismo y transporte de lípidos

En la categoría “I: Transporte y metabolismo de lípidos”, según la Figura 3, los 5 genes asociados a proteínas con mayor abundancia son: acil-CoA deshidrogenasa (COG1960), ligasa de ácidos grasos-CoA de cadena larga (COG0318), cadena alfa de la carboxilasa de acetil/propionil-coenzima A (COG4770), proteína hipotética (COG0365), proteína hipotética (COG1024).

Primeramente, las acil-CoA deshidrogenasas son flavoproteínas que participan en el catabolismo de ácidos grasos y aminoácidos, y se clasifican según el sustrato con el que tengan mayor especificidad [90]. La beta oxidación de ácidos grasos saturados ocurre en 4 pasos enzimáticos repetitivos, y las Acil-CoA deshidrogenasas participan en parte del primer paso, catalizando la conversión de acil-CoA a enoil-CoA [91]. Esta función está presente en los plasmidomas más profundos, 70928 y 8079, y en este último la abundancia es aún mayor, según lo indicado en la Figura 3, apartado “I”.

Por su parte, las ligasas de ácidos grasos-CoA de cadena larga, catalizan la formación de acil-CoA a partir de ácidos grasos, ATP y coenzima A [91, 92]. El gen asociado a esta función se halla en los plasmidomas más profundos de 70928 y 8079 metros.

En el caso de cadena alfa de la carboxilasa de acetil/propionil-coenzima A (COG4770), identificada en el grupo G2, con mayor abundancia en el plasmidoma de 8079 metros de profundidad, se trata de una proteína con actividad acetil-CoA carboxilasa y propionil-CoA carboxilasa dependiente de biotina, que participaría en la fijación de carbono mediante el ciclo del 3-hidroxiopropionato [93]. La cadena alfa es la subunidad responsable de la actividad biotina carboxilasa necesaria para la actividad enzimática [94].

El COG0365 refiere a la familia de las acil-CoA sintetisas/AMP-ácido graso ligasa, involucradas en el metabolismo de lípidos, activando el acetato formando acil-CoA, en los primeros pasos de la beta oxidación [95]. Por su parte, COG1024 define a la enoil-CoA hidratasa/carnitina racemasa, parte de la biosíntesis de ácidos grasos [96]. Los dos genes que fueron asignados con estos códigos tienen presencia solo en el plasmidoma de 500m.

5.3.5 Genes asociados a proteínas en la categoría de metabolismo y transporte de aminoácidos.

Dentro de esta categoría, como muestra la Figura 3, apartado “E: Transporte y metabolismo de aminoácidos”, los genes con mayor abundancia corresponden a: proteína hipotética (COG0620), proteína LivF de unión a ATP para transporte de alta afinidad de aminoácidos de cadena ramificada (COG0410), proteína de transporte YbaT de membrana interna (COG0559), proteína permeasa LivH del sistema de transporte de alta afinidad para aminoácidos de cadena ramificada (COG0559), proteína ribosomal 50s L27 (COG1171).

Dentro del grupo G2, se predice la proteína LivF de unión a ATP para transporte de alta afinidad de aminoácidos de cadena ramificada, perteneciente a la familia de transportadores HAAT (acrónimo del inglés: *hydrophobic amino acid uptake transporter*), de la super familia de transportadores ABC [97, 98]. LivF es una proteína citosólica que forman parte de un transportador con alta afinidad a leucina, isoleucina y valina. Este gen asociado a la proteína LivF presenta mayor abundancia en el plasmidoma de 8079 metros. Conjuntamente, se identifica a LivH, una proteína integral de membrana de la misma familia HAAT [98], con alta abundancia en los plasmidomas de 7928 y 8079 metros.

La proteína de membrana interna YbaT, se predice en el grupo G2. Esta proteína tiene función putativa de transporte de aminoácidos y aminas [99]. YbaT también está en el plasmidoma de 8079 con mayor abundancia respecto al plasmidoma de 7928 metros de profundidad.

La proteína hipotética con el código asignado COG0620, según la definición de este mismo código disponible en la página de COG database en NCBI [100], podría tener funciones metionina sintetasa independiente de cobalamina. Esta proteína hipotética se predice únicamente en el plasmidoma de 500 metros de profundidad, al igual que la proteína ribosomal 50s L27.

5.3.6 Genes asociados a proteínas en la categoría de transporte y metabolismo de iones inorgánicos.

En la categoría “P: Transporte y metabolismo de iones inorgánicos”, según ilustra la Figura 3, las 5 proteínas predichas con mayor abundancia son: proteína del sistema de eflujo de cationes CusA (COG3696), ATPasa de tipo P exportadora de cobre (COG2217), contra-transportador de cadmio, cobalto y zinc/H(+) – K(+) (COG1230), transportador de

cationes metálicos divalentes MntH (COG1914), transposasa ISAli20 de la familia Tn3 (COG3696).

CusA es parte de la superfamilia RND (acrónimo del inglés: *resistance-nodulation-cell división*), en su subfamilia HME-RND (acrónimo del inglés: *Heavy metal efflux RND*), del cual es el único miembro [101, 102]. CusA es una proteína homotrimérica de transmembrana que interactúa con CusB, CusC y CusF, formando un sistema de eflujo con alta afinidad para Cu^+ y Ag^+ , desde el citoplasma hacia el periplasma [103]. Se predice esta función en los plasmidomas de 2000, 7928 y 8079 metros de profundidad, donde la mayor abundancia se registra en el plasmidoma de 2000 metros.

La proteína predicha ATPasa de tipo P exportadora de cobre, se presenta también en los plasmidomas de 2000, 7928 y 8079 metros de profundidad. En *E. coli*, *copA* es el gen que codifica una ATPasa exportadora de Cu de tipo P [104]. La proteína CopA transporta el cobre cuproso Cu(I) citoplasmático hacia el periplasma donde es oxidado a cobre cúprico Cu(II) por CuO, una multicobre oxidasa [105]. Al igual que en el caso de la proteína predicha anterior, esta ATPasa también se haya en mayor abundancia en el plasmidoma de 2000 metros de profundidad.

El contra-transportador de cadmio, cobalto y zinc/H(+) – K(+) está asociado a la superfamilia de clústeres de dominios proteicos cl40534 que tiene como único miembro a CzcD, un transportador de la familia de proteínas facilitadoras de difusión de cationes que se ubica en la membrana citoplasmática, con al menos cuatro alfa hélices de transmembrana [106]. Su acción está orientada a la disminución de las concentraciones intracelulares de Cd^{+2} , Zn^{+2} y Co^{+2} [107, 108]. Esta proteína pronosticada se presenta en los plasmidomas de 2000, 7928 y 8079 metros de profundidad.

En los plasmidomas de 2000 y 7928 se hace presente el gen asociado a la proteína transportador de cationes metálicos divalentes MntH (COG1914). MntH es el miembro procarionte de la familia Nramp de transportadores de cationes divalentes [109]. MntH tiene afinidad preferencial por Mn^{+2} , no obstante, también es capaz de transportar Fe^{+2} , ambas funciones energizadas por el gradiente de protones [109]. Entre los dos plasmidomas mencionados, el de 2000 metros de profundidad posee la mayor abundancia de este gen asociado.

5.3.7 Genes asociados a proteínas en la categoría de tráfico intracelular, secreción y transporte vesicular.

En la categoría “U” correspondiente a “Tráfico intracelular, secreción y transporte vesicular”, como se deja ver en la Figura 3, los 5 genes asociados a proteínas más abundantes son: proteína del sistema de secreción tipo 4 VirB4 (COG3451), proteína hipotética (COG3451), proteína del sistema de secreción tipo 4 VirB9 (COG3504), transportador de la bomba de eflujo de membrana BerE (COG0541), proteína del sistema de secreción tipo 4 PtlE (COG3736).

Dentro de estas proteínas predichas, tres se identifican como parte del sistema de secreción tipo 4 (SST4), presentes y abundantes solo en los plasmidomas del grupo G2, Las mencionadas proteínas permiten transportar ADN, proteínas y otras macromoléculas, previo establecimiento de contacto físico entre los participantes [110]. En esta categoría, la mayor abundancia corresponde a la proteína del sistema de secreción tipo 4 VirB4, la cual es parte del conjunto de proteínas primordiales para la conformación de un sistema de transporte mínimo y funcional, y está presente en todos los SST4 [111]. En este mismo contexto, notamos a la proteína hipotética informada, la cual comparte su código COG con VirB4. En los resultados obtenidos, en el contexto del SST4, a VirB4 le sigue VirB9, la cual es parte de la estructura del canal, en específico del ensamble ubicado en el periplasma y membrana externa [112, 113]. En este mismo contexto, se identifica el gen asociado a la proteína PtlE, ortóloga de VirB8 identificada en *Bordetella pertussis*, proteína con actividad transglicosilasa [114].

Por otro lado, se predice la proteína BepE, una bomba de membrana de la superfamilia RND, la cual posee alta presencia en bacterias gram negativas [115]. BepE es componente de un sistema de eflujo junto a BepD y BepC, descritas en *Brucella suis* [116]. A este sistema de eflujo se le atribuye el transporte de un amplio espectro de antibióticos y otros tóxicos como bromuro de etidio y desoxicolato, entre otros [117]. El gen asociado a esta proteína se presenta únicamente en el plasmidoma de 500 metros de profundidad.

6. DISCUSIÓN

6.1 *Identificación de secuencias plasmidiales en la Fosa de Atacama: sobre su presencia, tamaño y otras características.*

A raíz de los análisis realizados a los distintos metagenomas que comprenden las profundidades de 500, 2000, 7928 y 8079 metros, dentro de la Fosa de Atacama, en consideración del primer objetivo planteado en esta tesis, sale a la luz la efectiva presencia de secuencias plasmidiales en las muestras de aguas, lo cual está en concordancia con el hallazgo de plasmidios en distintos organismos, inclusive los aislados desde la zona hadal de las fosas oceánicas más estudiadas, como se expuso anteriormente [55, 56]. Tras la normalización efectuada, se informa en la Tabla 1 el porcentaje de pares de bases plasmidiales sobre los pares de bases del metagenoma correspondiente. De este modo, se desglosan porcentajes diferentes para cada profundidad, diferencia aún más manifiesta en los metagenomas de 1000 y 8081 metros, los que no fueron incluidos en los análisis venideros por presentar muy pocos contigs identificados como plasmidiales. Una filtración de las muestras deficiente y/o una posterior extracción de ácidos nucleicos insuficiente pueden ser motivo de la menor presencia de secuencias plasmidiales en los metagenomas mencionados. No obstante, ¿Podría esto ser resultado de un menor porte de plasmidios por parte de las comunidades bacterianas de dichas profundidades? A la luz de los resultados obtenidos para los metagenomas restantes, esto es poco plausible, hallándose los metagenomas de 1000 y 8081 metros cercanos físicamente a los de 500 y 8079 metros de profundidad respectivamente, los cuales presentan una cantidad consistente de contigs plasmidiales con el resto de los metagenomas en estudio. Por su parte, los metagenomas que sí fueron considerados óptimos en lo que respecta a cantidad de secuencias plasmidiales, presentan también diferencias en este aspecto. En concreto, como se observa en la Tabla 1, el porcentaje de pb plasmidiales es mayor en los metagenomas más profundos, mientras que el metagenoma de 2000 metros es el que cuenta con la menor cantidad de pb plasmidiales dentro de los metagenomas incluidos en el análisis. Las mencionadas diferencias son muy atribuibles a distintos parámetros, dígame de las condiciones ambientales del nicho, la composición de las comunidades bacterianas y la estructuración espacial de las comunidades, la tasa de pérdida de plasmidios, el costo de mantención del plasmidio, e incluso de la etapa de crecimiento de las bacterias [118-122]. Primeramente, las condiciones ambientales que caracterizan cada ambiente a las distintas profundidades presentarán diferentes presiones selectivas que moldearán fenotipos bacterianos favorables, los que se extenderán a través de la comunidad, en parte gracias

a los plasmidios y otros elementos genómicos móviles. Los plasmidios que proporcionen aquellos caracteres fenotípicos favorables serán mejor preservados y distribuidos por las comunidades, dejando entrever una correlación entre la presión selectiva del ambiente y la prevalencia de plasmidios [122]. Respecto a la composición de las comunidades, se ha descrito que la velocidad de diseminación de plasmidios conjugativos y la magnitud de ésta, incrementa cuando la composición de la comunidad es más heterogénea, esto pues aumentaría la cantidad de cepas donadoras eficientes; mientras que al mismo tiempo, en algunas cepas, la velocidad y magnitud de la transferencia es menor, y ambos efectos provocarían una estabilización de la presencia del plasmidio [123]. Al mismo tiempo, la transferencia entre individuos de la misma especie está favorecida aparentemente por los mecanismos de defensa comunes entre los participantes de la transferencia, dígase de sistemas de matanza post-segregacional, sistemas CRISPR-Cas, entre otros [124-126]. En contraste, factores como el estrés o la densidad celular tienden a disminuir las tasas de transferencia horizontal [127, 128]. En cuanto a la estructuración espacial de las comunidades, es cierto que afecta a la prevalencia y permanencia de los plasmidios, esto cuando dicha permanencia depende de la frecuencia de intercambios; este tipo de presión selectiva basada en la frecuencia varía en comunidades que crecen en superficies [129, 130] y en aquellas que crecen en medios líquidos [131]. Ante el hecho de que en las fosas hadales el movimiento de sedimentos se dé por deslizamientos laterales, terremotos o por sedimentación, además de que se tiende a la acumulación de material sedimentario en las fosas oceánicas [7, 8, 132, 133], la heterogeneidad de ambientes se vería potenciada sin desmedro de las condiciones hidrográficas que caracterizan la zona, lo que podría favorecer la permanencia de plasmidios en las profundidades, pudiéndose esto reflejar en lo hallado en los plasmidomas aquí analizados. Otro aspecto por tener en consideración a la hora de evaluar la persistencia de plasmidios en una comunidad es la compensación entre los costos asociados a la mantención fisiológica de plasmidios y la ventaja fenotípica que aporten los plasmidios alojados, pues, según se postula, el número de plasmidios dentro de una comunidad puede ser satisfactoriamente mantenido si las bacterias que los hospedan tienen tantas aptitudes fisiológicas como aquellas que no albergan ningún plásmido [119]. Del mismo modo, es de importancia el número de copias que genera un plasmidio dentro de una célula hospedera y los costos de síntesis y mantenimiento que genera esta carga genética adicional [26, 134]. Así entonces, distintos son los parámetros que influyen en la mantención, distribución y persistencia de plasmidios en una población y

que son necesarios ponderar para una correcta interpretación de los resultados obtenidos en esta materia.

Los plasmidios muestran diversos tamaños, presentando un enorme rango entre el más pequeño y el más largo; pongamos por caso, desde el plasmidio de 5.51 Mb del organismo *Bacillus* sp. N5-665 al plasmidio de 117 bases de *Mesorhizobium japonicum* MAFF 303099, según lo registrado en GenBank a la fecha de revisión [135]. Más aun, en algunos organismos aislados desde las profundidades oceánicas, se reportan plasmidios de 112 kb a 586 kb en *Deinococcus* sp. D7000 [55] y *Paraoceanicella profunda* [56]. En el suceso de los resultados expuestos en la Tabla 2, se aprecian distintos tamaños promedio de los contigs definidos como plasmidiales para las distintas profundidades. En cada plasmidoma hay un amplio rango de tamaños, según se devela en las altas desviaciones estándar y, en concreto, el tamaño promedio de los contigs en los distintos plasmidomas varía, siendo más alto en el de 2000 metros de profundidad con cerca de 32 Kb aproximadamente, mientras que el más bajo se halla en el plasmidoma de 500 metros, con 19 Kb, aproximadamente (Tabla 2). Más allá de los aspectos que influyen en la construcción de un contig y su longitud final, dígame de la cantidad de reads superpuestos, su longitud y profundidad de secuenciación [136], y al igual que en el caso de la prevalencia de plasmidios en una comunidad; existe evidencia que apunta a la influencia que tendrían los ambientes sobre el tamaño que presentan los plasmidios de un conjunto bacteriano [137, 138]. Valga como ejemplo, un estudio de Slater et al. (2008) donde la tendencia indica que las secuencias plasmidiales más extensas son más abundantes en los ambientes más heterogéneos no asociados a animales [122]. Si bien aquel estudio arrastra las limitaciones de los sesgos de los muestreos en los distintos ambientes para obtener las secuencias disponibles en las bases de datos, en sí mismo pudiera significar una inclinación, tal como se presenta en el tamaño de cromosomas, y el genoma en general, respecto de la complejidad del ambiente y de la composición de la comunidad [137, 138]. Así, en términos de heterogeneidad y complejidad de un ambiente, se deben considerar todos los factores ambientales, tales como materia orgánica, iones, luz, oxígeno, pH, entre otros, los cuales varían en distintas magnitudes y están sujetos a sus propias relaciones y dependencias, a lo largo de la columna de agua oceánica. De este modo, los tamaños de secuencias plasmidiales podrían estar reflejando la complejidad de los ambientes oceánicos, además de poder responder a requerimientos fenotípicos que el ambiente habitado impone. La distribución y tamaños de plasmidios dentro de una comunidad también se ven influenciados por el tipo de plasmidios; esto es, si se trata de plásmidos conjugativos,

movilizables o no movilizables. En el primer caso porque los plasmidios conjugativos tienden a ser más grandes, al reunir diferentes operones y cerca de 20 genes necesarios para la transferencia y regulación del proceso de conjugación, además de otros genes asociados a la estabilidad del plasmidio, lo que en suma resultaría en un tamaño mínimo cercano a los 25 kb [139]. Estos tamaños y carga genética implican, además, un mayor compromiso energético por parte del hospedero [139]. Por otro lado, los plásmidos movilizables se consideran relativamente pequeños al no contener el sistema de transferencia en su totalidad; sin embargo, suelen tener también un mayor número de copias que aporta a su permanencia en los organismos [139]. Se ha mostrado que este fenómeno ha resultado en una distribución bimodal de plasmidios en las poblaciones; valga por caso, análisis basados en los datos de Plasmid Database, revelan que los plasmidios conjugativos se concentran alrededor de los 100 kb y los movilizables cerca de los 5 kb y luego cerca de 150 kb [34]. En contraparte, los plasmidios no movilizables poseen una distribución multimodal, siendo estas cercanas a los 5, 35 y 300 kb [34]. En suma, los tamaños informados en la Tabla 2 estarían influenciados tanto por las presiones adaptativas, la composición de la comunidad bacteriana, los aspectos que definen la permanencia de plásmidos en una comunidad, el contenido genético de las secuencias, y la capacidad de conjugación y/o movilización. En este último aspecto, cabe destacar que en los contigs identificados como plasmidiales, Platon halló orígenes de transferencia, genes para la conjugación, y no informa genes para la movilización (Tabla 3), lo que implicaría que en las comunidades de la Fosa de Atacama hay una parte importante de plasmidios no movilizables y una porción conjugable. Esta dualidad presente en los plasmidomas influiría en la distribución de tamaños plasmidiales como se expuso anteriormente. Si bien estarían dominando los plasmidios no movilizables en los plasmidomas analizados, no necesariamente se reflejaría esto en las tasas de transferencia horizontal dentro de las comunidades, pues otros mecanismos para estos fines pueden estar llevándose a cabo, por ejemplo, a través de la transducción o transformación [34]. De este modo, tanto la presencia efectiva de secuencias plasmidiales en los metagenomas de la Fosa de Atacama, como la información respecto de sus características de distribución y tamaño, además de la naturaleza mayoritariamente no movilizable de estos, quedan expuestos en el marco del primer objetivo planteado en esta tesis.

6.2 Caracterización funcional de las secuencias plasmidiales: Descubriendo el contenido genético de los plasmidios de la Fosa de Atacama

Seguidamente a la identificación de secuencias plasmidiales, con la intención de desvelar los caracteres genéticos contenidos en dichas secuencias, se realiza la anotación genómica, a razón del desarrollo del segundo objetivo específico planteado. Mientras Prokka identifica 6897 CDS (Tabla 4), Platon informa 6954 marcos de lectura abierto, a la vez que no incluye una herramienta de búsqueda de ARNt, como sí las tiene para identificar ARNr, los cuales no reporta, al igual que Prokka. En cuanto a los marcos de lectura abierto, en el entendimiento de que estos se conforman entre codones de inicio y término, y que en procariontes una CDS constituye un ORF [140], se evidencian diferencias entre lo informado por Prokka y Platón, sin desmedro de que los datos anunciados son bastante cercanos en el número de secuencias codificantes CDS; si bien es cierto que ambos softwares invocan a PRODIGAL para la predicción de CDS, la diferencia se halla en distintas configuraciones para el uso de PRODIGAL por parte de Prokka y Platon, donde en el primero algunos CDS son removidos cuando se sobreponen con características de secuencias de ARNt o ARNr. Incluso así, se considera que la diferencia no es suficiente para cambiar los métodos, y se prosigue con la anotación genómica. Por otro lado, a la luz de los resultados de Prokka respecto de secuencias de ARNt hallados en los plasmidomas de este trabajo, estos podrían hallar su razón de presencia en la consideración de que la traducción de proteínas codificadas en plasmidios suponen una carga energética y fisiológica adicional para la célula, motivando la búsqueda de que la traducción de estas proteínas no se vea impedida por una limitación de los ARNt ya presentes en la célula, a la vez de no perjudicar la traducción de las proteínas codificadas en el/los cromosomas [120].

Como resultado de la anotación genómica, fue posible organizar los genes anotados en categorías según el sistema COG. Tal como se evidencia en la Figura 3, las más altas abundancias de genes asociados a proteínas dentro de las categorías que representan funciones celulares mayores, y que fueron identificadas en los plasmidomas en análisis, corresponden a: metabolismo y transporte de carbohidratos (G), metabolismo y transporte de lípidos (I), y metabolismo y transporte de iones inorgánicos (P). Según lo reportado en análisis de metagenomas del océano profundo, como la Fosa de Puerto de Rico, en estos ambientes abundan los genes asociados a mecanismos de transducción de señales y regulación transcripcional [141]. Cuando se trata de entender las diferencias halladas entre los plasmidomas aquí presentados y los estudios disponibles de metagenomas del océano

profundo, es preciso hacer notar la prevalencia que tienen los genes considerados esenciales o “*housekeeping*” – pudiéndose catalogar como tales a reguladores transcripcionales o elementos de mecanismos de transducción de señales, por dar ejemplo - en los cromosomas por sobre los plasmidios [142], por lo cual no se debe establecer un paralelo entre plasmidomas y metagenomas de buenas a primeras. Ante la diferencia expuesta, se podría, entonces, entender la ausencia de las categorías asociadas a la estructura nuclear “Y” o a la estructura y dinámica de la cromatina “B”, por ser estas predominantemente de naturaleza eucarionte, o incluso más allá, la carencia de abundancias en categorías asociadas a metabolismos de carbohidratos, lípidos y aminoácidos a cierta profundidad -2000 m-, siendo plausible que dichas funciones estén integradas en los cromosomas, reflejando en su posible presencia en dicha estructura la necesidad de estabilidad, tanto de expresión como de transferencia vertical, condicionada por la disponibilidad de nutrientes y variables ambientales dentro del ambiente habitado. En esta misma línea, el hallazgo de abundancias destacables en las categorías “G”, “I” y “P” en los plasmidomas presentados, podría obedecer a las necesidades o presiones adaptativas del ambiente de la Fosa de Atacama, más que a requerimientos esenciales para la fisiología celular. En los metagenomas de la Fosa de Puerto Rico, se destaca la gran abundancia de transportadores como una característica adaptativa relevante, y se evidencia que buena parte de estos son transportadores de cationes, además de hallar diversas porinas de membrana externa, y transportadores de tri y dicarboxilato [141]. Ciertamente, en las categorías más abundantes de la zona hadal de la Fosa de Atacama – “G”, “I” y “P”- los transportadores son protagonistas. Así se revela en la Figura 3 donde, dentro de las categorías mencionadas, de los 5 genes asociados a proteínas más abundantes por categoría COG, muchos corresponden a transportadores. Podría ser que, al igual que en la Fosa de Puerto Rico, la presencia de transportadores sea indicadora de un mecanismo adaptativo que se podría originar en respuesta a los cambios en la fluidez y permeabilidad de las membranas con razón de las grandes presiones hidrostáticas. De ser así, entonces, al menos en la Fosa de Atacama, una parte de esta adaptación esta codificada en los plasmidios de las comunidades de las profundidades analizadas, reflejando además la necesidad y posibilidad de transferir estos caracteres genómicos. Por supuesto, el tipo de solutos a transportar ya sea hacia o desde el medio intracelular, está vinculado estrechamente a la disponibilidad de moléculas en el nicho y su potencial función o efecto en el ambiente celular.

Dentro de las categorías existentes, en esta tesis se pone especial énfasis en: aquellas con mayor abundancia, con gran presencia en la zona hadal, las que son relevantes para la adaptabilidad a las altas presiones, y las necesarias para la transferencia horizontal de genes. Así, a continuación se discute en profundidad la presencia de los genes hallados en las categorías “G”, “I”, “E”, “P”, “U” y “M”, en los distintos plasmidomas, incluyendo datos de otras fosas oceánicas y organismos piezófilos aislados, en un intento de comprender la relación entre la funcionalidad que ofrecen los genes hallados y las condiciones del ambiente único en el que se emplaza este trabajo.

6.2.1 *Metabolismo y transporte de carbohidratos en plasmidios de la Fosa de Atacama*

En la zona hadal, la categoría “G” que refiere al metabolismo y transporte de carbohidratos, es la más representada en los plasmidomas de la zona hadal de la Fosa de Atacama. Tanto el transporte como el metabolismo de carbohidratos se relaciona con la disponibilidad de materia orgánica en el ambiente; y justamente, la Fosa de Atacama se describe como una zona eutrófica de alta productividad primaria, consecuencia del sistema de corrientes de Humboldt, característica que la diferencia de otras fosas oceánicas, además de acumularse esta materia orgánica en las trincheras [143]. A este respecto, se describen genes asociados a proteínas directamente vinculados al manejo fisiológico de carbohidratos, en todos los plasmidomas analizados, con excepción del plasmidoma de 2000 m, registrándose la mayor abundancia de proteínas predichas en esta categoría en el plasmidoma de 8079 metros.

La proteína permeasa DctM es el gen asociado a proteína más abundante. Ésta es parte del transportador de C4-dicarboxilato, de la familia de transportadores TRAP (acrónimo del inglés *tripartite ATP-independent periplasmic transporters*), los cuales se caracterizan por asociarse estructuralmente con proteínas de unión a sustrato SBP (del inglés *substrate-binding protein*) [144], las que también fueron identificadas en esta categoría, siendo las segundas proteínas predichas más abundantes. Los transportadores TRAP han sido destacados en distintos trabajos enfocados en la microbiología hadal. Estudios comparativos con distintas cepas de *Photobacterium profundum*, esto es, la cepa de referencia piezotolerante SS9, la cepa piezotolerante DSJ4, y la cepa piezosensible 3TCK, muestran presencia de transportadores TRAP y que, además, estos están sobreexpresados a 45 MPa [145]. Por otro lado, un análisis de un metagenoma hadopelágico de la Fosa de Puerto Rico resalta la gran presencia de transportadores de di

y tricarboxilato, entre otros; en concordancia con el hecho de que el movimiento de solutos a través de la membrana es influenciado por la presión hidrostática [141]. Adicionalmente, un análisis genómico de 2016 de bacterias del clado *Roseobacter*, esto es, *Thiobacimonas profunda* JLT2016 y *Pelagibaca abyssi* JLT2014, aisladas desde 2571 y 2000 metros de profundidad, respectivamente, en el suroeste del océano Pacífico, identificó en ambas una importante cantidad de transportadores, donde los TRAP son los segundos más abundantes, y se encuentran tanto en los cromosomas como en los plasmidios [146]. Es evidente, entonces, que este tipo de transportadores TRAP están representados en buena cantidad en aquellos organismos piezófilos estudiados, y se deja entrever una expresión diferenciada dependiente de la presión hidrostática en algunos casos, dando luces de lo que podría ser una adaptación de los organismos a las altas presiones. El hallazgo de transportadores TRAP en plasmidios, mencionado en el estudio de 2016, indica que en los organismos del clado *Roseobacter* hay propensión a la transferencia vertical y horizontal de tales caracteres genéticos [146], estampando en dichos caracteres la naturaleza adaptativa de este fenotipo en herencia. Esto último, puede también estar siendo el caso en los resultados mostrados aquí.

La tercera proteína predicha más abundante en la categoría “G”, corresponde a AlgA, involucrada en la biosíntesis de alginato. Al menos dos géneros bacterianos pueden producir alginatos como exopolisacáridos, estos son, *Pseudomonas* [147] y *Azotobacter* [148]. Los alginatos producidos han demostrado ser de alta importancia para la formación de biopelículas densas muy estructuradas en *Pseudomonas aeruginosa*, mientras que en *Azotobacter*, los alginatos son altos en ácido alfa-L-gulurónico, lo que le permite generar geles aún más densos y formar quistes resistentes a la desecación [85]. En lo relacionado al océano profundo y funciones celulares asociadas al alginato, un estudio de 2016 describe una proteína alginato liasa en *Nitratiruptor* sp. SB155-2, una bacteria aislada desde una fumarola hidrotermal a 1000 metros de profundidad [149], indicando que el alginato podría ser utilizado por los organismos en ambientes profundos con altas temperaturas, aprovechando las propiedades fisicoquímicas que posee el polímero y que entrega a la biopelícula formada, esto es, mayor resistencia a temperaturas y presión, y en el caso particular de esta enzima liasa, sería necesario, además, el desacople de células desde las matrices de alginato [150]. Más en detalle, los geles de alginato e ión calcio son altamente resistentes y estables entre los 0°C y 100°C; además, las propiedades de los geles de alginato y ión calcio tales como estructura, conductividad eléctrica y fuerza, no sufren cambios significativos a presiones bajo los 150 MPa (1480.4 atm) [151, 152]. De esto se

deriva el pensar a los exo-polisacáridos, como el alginato, como adaptaciones que permiten resistencia a las altas presiones, evitando los efectos de estas sobre la arquitectura celular, su volumen y morfología en general.

Junto a los genes asociados a proteínas ya mencionados, se presentan dos miembros de un importador de ribosa, estos son RbsC y RbsA. Como resultado del importe de ribosa se puede sintetizar ribosa-5-fosfato, en cuyo catabolismo participa la rama no oxidativa de la vía de las pentosas fosfato [153]. De dicha vía, a la cual ingresan también la ribulosa-5-fosfato y la xilulosa-5-fosfato, resultan intermediarios de la vía glicolítica [153, 154]. La ribosa-5-fosfato es de alta importancia en distintos procesos anabólicos, por ejemplo, es sustrato para la síntesis de fosforibosil difosfato, el cual es precursor para la síntesis de nucleótidos y de los aminoácidos histidina y triptófano [155]. En un estudio genómico de la cepa piezófila SS9 de *P. profundum*, se revela una proteína asociada al metabolismo de la ribosa, un represor putativo RbsR del operón ribosa que, en un análisis comparativo mediante microarray donde se evaluaron los perfiles transcriptómicos en condiciones de 0.1 MPa y 28 MPa, dicha proteína se muestra subexpresada a 0.1 MPa, indicio del efecto que tiene una condición de presión poco óptima para el organismo sobre los procesos metabólicos del carbono, lo que podría resultar en una situación de estrés celular [145].

Que la mayor abundancia de genes asociados al metabolismo y transporte de carbohidratos en plasmidomas se halle en el más profundo aquí estudiado podría obedecer, primero, a la presencia de plásmidos en los taxas de la zona. Sin ir más lejos, según los resultados obtenidos por Olivares [62] sobre taxonomías en la columna de agua de la Fosa de Atacama, en la zona hadal, el filo *Planctomycetes* es más abundante respecto de su presencia en aguas suprayacentes y, según la literatura, la familia *Isosphaeraceae*, dentro de los *Planctomycetes*, posee hasta 4 plasmidios donde el más largo en su mayoría codifica genes para la biosíntesis de carbohidratos [156]. A pesar de solo haberse identificado plasmidios de la familia indicada en las secuenciaciones de genomas de organismos cultivados del filo *Planctomycetes*, no se descarta que esta observación pueda extenderse a otras familias del filo que pudiesen estar presentes en la profundidad de la Fosa. Sumado a esto, es relevante la disponibilidad de los carbohidratos en el medio, tópico en el cual se ha registrado una disminución de la abundancia relativa de polisacáridos en la materia orgánica disuelta, a medida que incrementa la profundidad [157]. Si bien desde aquí se podría extraer que aquello podría motivar un aumento de transportadores para satisfacer

las necesidades celulares, junto con la presión adaptativa que suponen las altas presiones hidrostáticas, lo cierto es que hay insuficiente información acerca del comportamiento de carbohidratos en la materia orgánica en la zona abiso-hadal del océano o en la Fosa de Atacama en particular.

6.2.2 *Sobre el metabolismo y transporte de lípidos en plasmidios de la Fosa de Atacama*

Los ácidos grasos y lípidos son esenciales para la estructura de membranas plasmáticas y además son importante fuente de energía para los organismos, de ahí la relevancia de los procesos metabólicos de lípidos. En bacterias se ha descrito como los ácidos grasos más abundantes a aquellos de 14 a 18 carbonos de largo, considerados de longitud mediana a larga [158]. En las aguas de la Fosa de Atacama, los lípidos presentes forman parte de la materia orgánica lábil y particulada a la que contribuyen los organismos fototróficos de aguas someras, así como los organismos heterótrofos de las zonas subyacentes [159, 160], además del transporte lateral de sedimentos, sobre todo en la zona abisal y hadal [7]; lo cual contribuye a que la categoría “I” presente la segunda mayor abundancia relativa en los plasmidomas hadales.

Esta categoría la domina la proteína predicha acil-CoA deshidrogenasa. Éstas participan de la beta oxidación de ácidos grasos saturados, donde catalizan la formación de enoil-CoA que será luego procesado para generar acetil-CoA, que participa en otras vías metabólicas [91]. Evidencia de esta función aparece en un análisis genómico del organismo piezófilo *Shewanella violácea*, aislado desde la capa superficial de los sedimentos de la Fosa Ryukyu a 5110 metros de profundidad, y el cual crece óptimamente a 30 MPa, que revela en su genoma múltiples proteínas de la familia de las acil-CoA deshidrogenasas, además de ligasas de ácidos grasos CoA de cadena larga [161], halladas también en los plasmidomas aquí analizados; un antecedente que podría dar algunas luces acerca de la relación entre la presencia de genes asociados a un propósito funcional particular y el metabolismo de ácidos grasos en condiciones de alta presión.

Las proteínas predichas como ligasas de ácidos grasos-CoA de cadena larga, también conocidas como acil-CoA ligasas o acil-CoA sintetasas, intervienen en distintos procesos, tales como; degradación de ácidos grasos, biosíntesis de lípidos, metabolismo intermediario y expresión génica, incorporación de triglicéridos y transducción de señales [91, 158], aportando desde la formación de acil-CoA a partir de ácidos grasos, ATP y

coenzima A [162], y se hallan abundantes en los plasmidomas hadales aquí trabajados. En *E. coli*, la proteína FadD, perteneciente al operón *fadIJ*, es la responsable de la función acil-CoA sintetasa y le permite al organismo el crecimiento mediante el uso de ácidos grasos como única fuente de carbono, participando de una beta-oxidación anaeróbica [163]. En el análisis transcriptómico comparativo de la cepa SS9 de *P. profundum*, se identifica una acil-CoA ligasa sobreexpresada en condiciones de alta presión, pudiéndose extraer de aquí que a tales presiones, esta función se potencia como mecanismo adaptativo, pudiendo incrementar el rol del metabolismo de lípidos. Las acil sintetetas participan además en la activación de los ácidos grasos poliinsaturados de cadena larga en la etapa de terminación de la biosíntesis de tales ácidos grasos, los cuales quedan disponibles para su incorporación a fosfolípidos [164]. Sin ir más lejos, se ha reportado que una importante fracción de ácidos grasos poliinsaturados forman parte de los fosfolípidos de organismos piezófilos aislados desde la Fosa de Las Marianas [165]. Incluso, en la Fosa de Atacama se evidencia una abundancia mayor de ácidos grasos poliinsaturados en los sedimentos respecto de las aguas, fenómeno que se atribuye en parte a la producción por parte de los organismos piezófilos del nicho [160].

A estas proteínas le sigue la acetil/propionil CoA carboxilasa, la cual tiene la particularidad de presentar las funciones acetil CoA carboxilasa y propionil-CoA carboxilasa en una misma estructura proteica, funciones que suelen encontrarse en proteínas separadas; sin embargo, hay registros de acetil/propionil CoA carboxilasa en algunas arqueas, siempre dependientes del cofactor biotina y ATP [93, 94, 166]. Las funciones carboxilasas mencionadas son relevantes en el proceso de fijación de carbono mediante el ciclo del 3-hidroxiacetil-CoA [167, 168], aportando además a la biosíntesis de ácidos grasos y metabolitos secundarios [169]. La enzima se postula como la principal respecto a la fijación autotrófica de carbono. En la Fosa de Puerto Rico se hallaron componentes de diversas vías de fijación de carbono autotrófica, entre ellas el ciclo del 3-hidroxiacetil-CoA, en la cual participan las carboxilasas mencionadas anteriormente [141]. Además, el metabolismo general de ácidos grasos está presente con una abundancia relativamente alta en la Fosa de Puerto Rico [141], lo cual aporta un antecedente del hallazgo de enzimas involucradas con vías de fijación de carbono autotróficas y con el metabolismo de lípidos en el océano profundo.

El hecho de que las proteínas predichas más abundantes dentro de esta categoría estén dominando los plasmidomas de 7928 y 8079 metros de profundidad, según se

muestra en la Figura 3, y se relacionen además con la biosíntesis y beta-oxidación de ácidos grasos, y la fijación de carbono, podría suponer que los organismos hadales de las profundidades aquí analizadas tengan tendencia a orientar los esfuerzos adaptativos hacia el metabolismo de ácidos grasos, así también lo sugieren los estudios metagenómicos y el análisis transcripcional revisados.

6.2.3 Sobre el metabolismo y transporte de aminoácidos en plasmidios de la Fosa de Atacama.

Los aminoácidos son esenciales para diversos procesos y estructuras celulares, son intermediarios tanto en el metabolismo del carbono como en el del nitrógeno. Debido a su indispensable rol, y debido a las distintas maquinarias de transporte que se requieren para la variedad de aminoácidos existentes -los cuales son distintos en sus propiedades químicas-, los mecanismos funcionales y regulatorios de su metabolismo son de alta importancia [98]. En el océano se ha reportado que el filo *Bacteroidetes* y la clase *Gamma-proteobacterias* juegan un rol en la asimilación de aminoácidos, sobre todo cuando estas macromoléculas se encuentran en altas concentraciones [170, 171]. Justamente, ambos están presentes a lo largo de la columna de agua en la Fosa de Atacama, según Olivares et al [62].

En este plano, el gen asociado a proteína más abundante de esta categoría corresponde a una proteína hipotética que se concentra en el plasmidoma de 500 metros. A partir del código COG asignado (COG0620), se le asocia a la función de metionina sintetasa independiente de cobalamina. La metionina es requerida en la biosíntesis de proteínas y poliaminas y es, además, un precursor de S-adenosilmetionina, principal donador de metil para diversas metiltransferasas [172]. El último paso de la biosíntesis de metionina está altamente conservado y es llevado a cabo por metionina sintetasa, cuya actividad enzimática puede depender o no de la presencia de cobalamina (Vitamina B12) u otro cofactor orgánico [173, 174].

Adicionalmente, YbaT, identificada en los plasmidomas más profundos, está asociada a la respuesta celular frente al estrés ácido dependiente de glutamato en la cual aporta desde el transporte de glutamina, la cual es transformada a glutamato mediante hidrólisis por YbaS [99]. Sin embargo, YbaT no resulta esencial en la sobrevivencia de *E. coli* en medios con pH disminuidos [99]. En *E. coli*, se ha descrito el rol que tiene en la resistencia al cobre en medios ácidos, mediante la respuesta dependiente de glutamato,

que se manifiesta en la sobre expresión de YbaT y YbaS en condiciones de altas concentraciones de cobre y bajo pH, y la baja resistencia celular a tales condiciones cuando se realiza delección de los genes de ambas proteínas [175]. En el contexto de su hallazgo en los plasmidomas hadales, podría asociarse a los requerimientos fisiológicos en línea con la presencia del metal en el ambiente.

Junto con esto, se identifican en la zona hadal dos proteínas predichas asociadas al desplazamiento de aminoácidos a través de la membrana, LivF y LivH. Dentro de los transportadores de aminoácidos, la familia de transportadores HAAT perteneciente a la super familia de transportadores ABC, se encarga del transporte de aminoácidos de cadena radical ramificada [98]. La familia HAAT consiste en un solo grupo de transportadores, LIV-I. En *E. coli*, LIV-I consiste en dos proteínas extracelulares de unión a solutos, LivK y LivJ; dos proteínas integrales de membrana, LivH y LivM; y dos proteínas citosólicas de unión a ATP, LivF y LivG. Particularmente, LivK tiene especificidad solo con leucina, mientras LivJ tiene alta afinidad por leucina, isoleucina y valina, y se une con menor afinidad a treonina, alanina y serina [97, 176]. En la Fosa de Puerto Rico se destaca la mayor abundancia de algunos transportadores ABC, entre ellos los HAAT, por sobre lo hallado en un metagenoma de aguas superficiales, lo cual indicaría que en la adaptación a las altas presiones, los transportadores ABC estarían marcando una presencia relevante, que también se relaciona con las adaptaciones a nivel de composición de membrana plasmática [141]. El encontrar estos transportadores en secuencias plasmidiales refleja la capacidad de transferencia de un carácter genético ventajoso, dadas las presiones adaptativas de la zona hadal de la Fosa de Atacama. A lo último se suma la naturaleza lábil de los aminoácidos contenidos en la materia orgánica, pues estos tienden a la degradación temprana en comparación al carbono orgánico sedimentario [177], lo que podría derivar en mecanismos para una captación adecuada y suficiente para los requerimientos fisiológico.

6.2.4 Sobre el metabolismo y transporte de iones inorgánicos en plasmidios de la Fosa de Atacama

Entre las categorías COG identificadas en los plasmidomas aquí en estudio, la categoría "P", referente al metabolismo y transporte de iones inorgánicos, es la que cuenta con la mayor abundancia, concentrándose sobre todo en el plasmidoma de 2000 metros. Según las proteínas predichas aquí encontradas, algunos de los metales presentes asociados son el cobre, cadmio, níquel, cobalto, zinc y manganeso, los cuales, con excepción del cadmio, tienen conocidas y esenciales funciones en procesos biológicos, esto

es; procesos enzimáticos, de transporte, transducción de señales, entre otros [178]. De ahí que su homeostasis es regulada por las células, a fin de contar con los requerimientos fisiológicos de disponibilidad de dichos metales, a la vez de evitar los efectos de la toxicidad por estos. Tanto el ingreso como el egreso de los metales es regulado por proteínas metaloreguladoras y sistemas de transporte [179]. En este último punto es importante la acción de los sistemas de eflujo, los cuales fueron identificados en los plasmidomas aquí estudiados.

Las dos proteínas predichas más abundantes en esta categoría están relacionadas al exporte de cobre (Figura 3). El cobre es un micronutriente esencial para los sistemas biológicos, de ahí que su homeostasis en los organismos sea altamente regulada, y los mecanismos que regulan la concentración intracelular de este metal y previenen la intoxicación son claves para la célula [180]. Tanto CusA como la ATPasa tipo P aquí descritas, tienen roles en procesos de detoxificación fisiológica. En el caso de CusA, su función de transporte es de alta importancia para la detoxificación, aún más en condiciones de bajo oxígeno por estar limitada la oxidación a Cu^{+2} , estado menos tóxico, por parte de CuO, una multicobre oxidasa [181]. En específico, respecto a la presencia de CusA en organismos de aguas profundas, se ha reportado un análisis transcriptómico en *Nitratiruptor* sp. SB155-2 [182], aislado desde una fumarola hidrotermal a 1000 m de profundidad, el cual revela que ni CusA ni CusB son diferencialmente expresados en condiciones de altas concentraciones no letales de cobre, sugiriendo que el componente *cus* podrían no ser esencial para la detoxificación, considerando las condiciones micro-aeróbicas del estudio; sin embargo, el componente *cus* sí es dominante ante altas concentraciones de cobre bajo condiciones aeróbicas en *E. coli* [183], organismo piezosensible. Ante esto, es posible que la estrategia de detoxificación del organismo piezófilo *Nitratiruptor* sp. SB155-2 se apoye más en sistemas de eflujo, como lo revela el estudio de Ares et al. (2021) [182]. Los sistemas de eflujo de cobre compuestos por ATPasas de tipo P son parte de la categoría de mecanismos de resistencia más diseminado entre los organismos [184]. En *E. coli*, *copA* es el gen que codifica una ATPasa exportadora de Cu de tipo P, como la hallada en los plasmidomas en estudio. En *E. coli*, la expresión del ARNm de CopA aumenta ante la adición de 1 mM de CuCl_2 , alcanzando hasta 350 copias por célula tras 2 minutos de exposición en condiciones aeróbicas [185]. En condiciones anaeróbicas, dicho incremento es de hasta 1200 copias por célula, dejando ver la expresión diferenciada de esta ATPasa de tipo P en presencia y ausencia de oxígeno [185], lo cual también se describe para los sistemas de eflujo de *Nitratiruptor* sp. SB155-2 ante estrés por cobre [182]. Sin desmedro

de esto, también es cierto que la presencia de estos genes se relaciona con la presencia y distribución del metal en cuestión.

A las proteínas predichas asociadas al transporte de cobre, les sigue un anti-transportador de cadmio, cobalto y zinc/H(+)-K(+), que tiene por función la detoxificación de Cd^{+2} , Zn^{+2} y Co^{+2} [106]; sin embargo, el efecto de resistencia que provee es menor respecto a otro tipo de sistemas de resistencia a metales, como CzcCBA [108] [107]. En concreto, en el océano profundo, un análisis de metagenomas de aguas profundas en Matapan - Vavilov (de profundidades menores a 4500m, en la parte más profunda del Mar Mediterráneo) se halló un gran grupo de proteínas deducidas pertenecientes a categorías relacionadas a la resistencia a metales pesados, como el sistema de eflujo para Co/Zn/Cd CzcD [186]. Además, se identifica en los plasmidomas analizados el gen asociado a MntH, responsable del importe de Mn^{+2} , cuya presencia en los plasmidomas de 2000 y 7928 metros se relaciona, desde luego, con la presencia del metal en las aguas.

Bajo la idea de que la presencia de estos y otros genes se debe en parte a las presiones adaptativas que suponen las condiciones ambientales, es relevante considerar la distribución de los distintos metales a lo largo de la columna de agua. La disponibilidad de los metales en el océano, en toda la extensión de la columna de agua, se genera desde distintas fuentes, como pueden ser: remineralización de pellets fecales, aporte de ríos, flujos desde fumarolas hidrotermales, aportes desde la actividad humana, transporte de sedimentos, entre otros [187-191]. En la zona de la Fosa de Atacama, según eGEOTRACES, la concentración de Cu disuelto aumenta desde los 1.25 nmol/kg aproximadamente, hasta los 4 nmol/kg aproximadamente, entre los 500 y 3000 metros, profundidad a partir de la cual las concentraciones se mantienen entre los 4 y 4.5 nmol/kg hasta los 5000 m [192]. La medición de metales en el océano también incorpora el biomonitoreo con organismos como *Eurythenes gryllus* y *Hirondella* spp., anfípodos lisianasoides ampliamente distribuidos en el océano, desde los hábitats intermariales hasta los hadales [193]. Esta estrategia fue utilizada en un estudio de 2018, donde se informan concentraciones de Cu y otros metales en la Fosa de Atacama mediante *Eurythenes gryllus*. Más en detalle, las mediciones a las profundidades de 4602, 5329 y 6173 metros dan concentraciones de Cu de 24.73, 15.21, 22.23 $\mu\text{g/g}$ de peso seco; mientras que para Zn son de 188.43, 230.2, 120.27 $\mu\text{g/gr}$ de peso seco; y para Cd son de 15.56, 6.19, 18.25 $\mu\text{g/gr}$ de peso seco, respectivamente. Si bien, las concentraciones halladas pudiesen estar influenciadas por la fisiología misma del organismo anfípodo, también son reflejo de la

biodisponibilidad del metal [194]. En otras fosas del océano pacífico, el cobre tiende a aumentar su concentración a medida que aumenta la profundidad entre los 2000 y 8000 metros [195] [196]. No obstante, se describe el comportamiento de la concentración de cobre en la columna de agua como híbrido, al ser dependiente de procesos de reciclamiento y recolección [197]. En resumen, según los antecedentes revisados, las concentraciones de cobre son mayores a medida que aumenta la profundidad hasta un punto en el cual la concentración se mantiene en un rango estable, lo cual podría explicar, en parte, las abundancias de los genes asociados a transportadores de cobre en las secuencias plasmidiales descritas aquí, las cuales se presentan entre los 2000 y 8079 metros de profundidad; esto sumado a la preferencia que tienen los organismos por recurrir a los transportadores a la hora de aminorar los efectos de las altas presiones hidrostáticas en las membranas plasmáticas [16]. En el caso de los metales Cd y Zn, se ha mencionado en otros trabajos que estos poseen un comportamiento tipo nutrientes, lo que implicaría bajas concentraciones en las aguas someras en contraste con las aguas profundas, donde estos metales suelen oxidarse y remineralizarse [197], presentando además un máximo de concentración en profundidades medias. El zinc mantiene su concentración relativamente estable a partir de los 3000 metros hasta los 10000 metros, según se ha visto en otras fosas del océano Pacífico [196, 198, 199], y según eGEOTRACES [192], las concentraciones de zinc aumentan en las proximidades del continente Sudamericano. Según eGEOTRACES, el cobalto disuelto está presente en concentraciones altas en la superficie en comparación con las aguas profundas, más en concreto, cerca de 80 pmol/kg en los primeros 100 metros, aumentando a 120 pmol/kg a los 300 metros, luego decayendo hasta los 10 pmol/kg a los 3500 m, y bordeando los 30 pmol/kg más allá de los 5000 metros [192]. El manganeso disuelto, según eGEOTRACES, se mantiene estable en su concentración desde los 1000 metros hacia la profundidad [192], mientras que según otros autores, la concentración de Mn alcanza su punto más alto en los 1700 metros aproximadamente, en el Océano Pacífico, disminuyendo en las profundidades a medida que Mn⁺² es absorbido en partículas en precipitación, además de irse incorporando a los sedimentos [200]. Puede ser que a razón de estas concentraciones se vea una mayor abundancia del transportador de Manganeso en el plasmidoma de los 2000 metros. En suma, estos metales están distribuidos en la columna de agua, de tal forma que es posible su influencia en el hallazgo de genes asociados a la movilización de estos metales

A lo ya revisado, se suma el antecedente de la Fosa de Puerto Rico de sobrerrepresentación de transportadores miembros de la familia HME (acrónimo del inglés:

Heavy metal efflux), como lo es CusA, y aquellos pertenecientes a la superfamilia de las ATPasas tipo P, también hallada en los plasmidomas aquí expuestos; por sobre las muestras de aguas someras y en comparación con las aguas del mar Sargaso, según reportan Eloe y equipo (2011) [141]. Esto podría suponer una estrategia basada en transportadores asociados a ATP y a H⁺, orientada a la detoxificación de metales pesados compartida en ambos ambientes hadales, y en el caso particular de este trabajo, parte de esta potencial estrategia esta codificada en las secuencias plasmidales de la Fosa de Atacama, en las profundidades aquí estudiadas.

6.2.5 Trafico intracelular, secreción y transporte vesicular: Posibilidad de THG por conjugación de plasmidios en la Fosa de Atacama

En los plasmidomas se identificaron genes asociados a proteínas en la categoría “U: tráfico intracelular, secreción y transporte vesicular”, función encargada del transporte de solutos dentro del espacio celular y hacia el exterior o hacia otra célula. Esta es una categoría relevante pues comprende las maquinarias requeridas para la conjugación plasmidial, proceso que suele recurrir a instrumentos sofisticados como los SST4, los cuales están involucrados en la biogénesis del pilus y en el transporte de moléculas o factores de virulencia hacia otras células [201, 202]. Esta maquinaria multiproteica se diferencia en subfamilias cuando se trata de sistemas conjugativos o translocadores efectores, y tiene muchas similitudes entre bacterias gran negativas y positivas en su estructura y proceso de transferencia [201, 203]. Los SST4 conjugativos conforman el MPF (sigla del inglés *mating pair formation*), los genes que forman parte de este sistema llevan el nombre de genes MPF, y son requeridos para que un plásmido sea auto-transmisible [110, 204]. La mayor abundancia de esta categoría se presenta en el plasmidoma de 7928 metros, mientras que en el plasmidoma de 2000 metros no se hayan genes asociados, lo podrían deberse a la presencia de plasmidios no conjugables y/o plasmidios no movilizables que predominaría por sobre aquellos conjugativos.

Entre los 5 genes asociados a proteínas más abundantes dentro de esta categoría destacan distintos miembros de la maquinaria de SST4, primeramente VirB4, seguido de VirB9, PtlE, e incluso, posiblemente, la proteína hipotética pues su COG asociado se corresponde con VirB4, como muestra la Figura 3. VirB4 es una de 3 ATPasas que forman un núcleo energético en la base de la maquinaria en el citoplasma [205], y se considera además parte del ensamble de la membrana interna, interactuando con proteínas del canal en sí mismo, dígame; VirB3, VirB6, VirB8, y el dominio N-terminal de VirB10 [206]. Por su

parte, VirB9 es una proteína hidrofílica que interactúa con VirB7, VirB8 y VirB10, estabilizando la estructura del canal de secreción [112, 113]. Los genes mencionados se distribuyen solo en los plasmidomas hadopelágicos (7928 y 8079 metros), lo que además está en concordancia con la abundancia que informa Platon sobre genes asociados a la conjugación (Tabla 3), la cual halla sus máximos en los plasmidomas de 7928 y 8079 metros; esto sin desestimar la presencia de otros genes asociados a esta categoría COG en los otros plasmidomas, más allá de los 5 más abundantes que son mostrados en la Figura 3. La presencia de SST4 en los ambientes marinos ha sido documentada previamente; sin ir más lejos, en un estudio de 2007 se identificaron SST4 en el género *Roseobacter* [207], y en otro trabajo se identifican, dentro de un conjunto de organismos marinos, SST4 tipo VirB/D4 solo en las alphaproteobacterias incluidas en dicho estudio [208]. Lo hallado en este último estudio mencionado se distribuye entre el clado de las *Roseobacter*, en las familias de *Eritrobacteraceae* del orden de las *Sphingomonadales*, además de la familia *Caulobacteraceae*, todas ellas presentes en las aguas de la Fosa de Atacama [62]. Esto estaría indicando una influencia de los taxos presentes en las comunidades para la presencia de genes destinados a la constitución de un pili que podría mediar en el intercambio de ADN.

Si bien la abundancia de componentes para el ensamblaje de SST4 y la cantidad de genes para la conjugación informados por Platon concuerdan entre ambos, en términos de abundancia, lo cierto es que podría esperarse que a este hallazgo le acompañe un número similar de secuencias OriT, en la concepción de que estos son también necesarios para la conjugación. Sin embargo, Platon reporta 3 secuencias OriT, una por los plasmidomas de 500, 2000 y 7928 metros, respectivamente. Se podría esto deber a una poco probable pero no descartable sub-identificación de secuencias por parte del software Platon, o podría indicar que los SST4 hallados en los plasmidomas están orientados o especializados en la transferencia de proteínas y otras macromoléculas [34]. Esto último estaría en concordancia, además, con la nula cantidad de genes para la movilización de plasmidios informados por Platon, y con la posterior suposición de que la mayoría de los plasmidios aquí reportados son no conjugables/no movilizables.

6.2.6 Moviloma: Profagos y transposones, otras vías de THG en la Fosa de Atacama

Además, por su rol en la movilidad de genes, se revisa la categoría “X: Moviloma: profagos y transposones”, donde la presencia de genes anotados en los plasmidomas de 2000, 7928 y 8079 metros es efectiva (Figura 3 y 2). En esta categoría, la mayor parte

corresponde a transposasas de distintas familias de secuencias de inserción. Los elementos transposables no solo son destacables por su papel en la plasticidad y dinámica de los genomas [209], sino que también porque se ha reportado en distintos análisis de metagenomas profundos una presencia importante de elementos transposables, como en la Fosa de Puerto Rico [141], en el Océano Pacífico Norte [210][211] y en el Mar Mediterráneo [212], así como también se ha reportado una gran abundancia de genes asociados a profagos y transposones en SAGs de organismos del grupo *Roseobacter* y del género *Alteromonas* aislados desde la zona hadal de la Fosa de las Marianas [213]. Si lo mismo fuese cierto en las comunidades procariontes de la Fosa de Atacama, entonces parte de aquella abundancia está distribuida en las secuencias plasmidiales en las profundidades aquí estudiadas.

6.2.7 Sobre la biogénesis de pared, membrana y envoltura celular en los plasmidomas de la Fosa de Atacama

Adicionalmente, por haberse reportado como una de las estructuras celulares en las que las altas presiones ejercen un gran efecto [14], se presta atención a la categoría “M: Biogénesis de pared, membrana y envoltura celular”, la cual presenta abundancias similares entre los plasmidomas, encontrándose la más grande en el plasmidoma de 500 metros. Entre los 5 genes más abundantes se hallan dos genes anotados en los plasmidomas más profundos, una apolipoproteína-N-acetiltransferasa y una dTDP-glucosa 4,6-dehidratasa 2 (Figura 3). La primera es una proteína integral de membrana que cataliza el último paso de la modificación post-transcripcional por la que pasan las lipoproteínas de bacterias [214]. Dichas lipoproteínas son importantes componentes de las membranas celulares y poseen distintas funciones entre las que podemos contar aquellas de naturaleza enzimática -como es el caso de las betalactamasas- [215], o aquellas con actividad activadora/inhibitoria hacia otras enzimas [216, 217]. Esta enzima en particular N-acila una cisteína de apolipoproteína usando fosfatidiletanolamina como fuente de lípidos, generando últimamente una lipoproteína triacilada [214]. La segunda de ellas, dTDP-glucosa-4,6-dehidratasa, participa en la síntesis de desoxiazúcares como la inusual 6-desoxihexosa [218], la cual es necesaria para la glicosilación de distintas moléculas, afectando su actividad, toxicidad y solubilidad [219]. La mayoría de las 6-desoxihexosas están involucradas en el metabolismo secundario y algunas de ellas son consideradas importantes dentro de polisacáridos accesorios a la pared celular de algunas bacterias[218]. Estas proteínas predichas reflejan potenciales adaptaciones a nivel de modificaciones post

transcripcionales, a fin de habilitar funciones de reconocimiento molecular, y a nivel de modificaciones a la pared celular y otras moléculas [220]. Si bien es plausible esperar una mayor presencia de genes asociados a la biogénesis de membranas en la zona hadal, por ser necesaria la modificación de éstas con razón de la alta presión hidrostática, como se ha señalado previamente; lo cierto es que es igualmente plausible que los caracteres genéticos que posibilitan dichas modificaciones estén contenidos en los cromosomas, resultando en la abundancia encontrada en los plasmidomas profundos aquí analizados.

6.3 Desafíos y proyecciones del estudio de plasmidios de ambientes oceánicos profundos

A la fecha, la exploración de las profundidades del océano y los estudios científicos asociados van al alza, luego de superar varios desafíos tecnológicos que supone el muestreo de este ambiente extremo; así lo reflejan los artículos científicos que van en aumento. Sin embargo, esto no se ha extendido al estudio de plasmidios, de la misma forma en la que estos son estudiados en los ambientes marinos superficiales. Podría esto deberse a la temprana etapa de caracterización biológica y molecular de las fosas hadales, más que a un desinterés general en materia de plásmidos, considerando el papel que juegan en las dinámicas evolutivas y adaptativas de las comunidades procariontes. En este sentido, el trabajo aquí realizado responde al interés mencionado, con una primera identificación y caracterización funcional de secuencias plasmidiales en un reducido número de muestras y profundidades. Sin dudas, resultaría necesario y enriquecedor poder extender este trabajo a la totalidad de la columna de agua y sedimentos, y así abarcar e integrar de mejor forma los resultados a hallar con las características del nicho en estudio, favoreciendo, además, la comparativa en términos de función en relación con las variaciones físicas y químicas que se dan a lo largo de la columna, desde las aguas someras hasta las grandes profundidades oceánicas, pudiendo incluso establecer vínculos con las funciones codificadas en cromosomas. Igualmente, este primer esfuerzo podría replicarse en otras fosas hadales a fin de comprender mejor la adaptación y evolución en ambientes de altas presiones hidrostáticas, pudiéndose ahondar en las distintas funciones celulares a las cuales pudieran aportar los plasmidios; por dar ejemplo y desde los resultados aquí expuestos, sería valioso profundizar en aquellas funciones celulares más representadas; dígame del metabolismo de iones inorgánicos o el metabolismo de carbohidratos, por nombrar algunos. Adicionalmente, resultaría de gran relevancia lograr asociar plasmidios y taxonomías de una comunidad, pudiéndose revelar relaciones entre taxas portadoras de

plasmidios y la persistencia y distribución de estos últimos en la comunidad, enriqueciéndose aún más nuestra comprensión del vínculo entre plasmidios, organismos y ambiente habitado.

7. CONCLUSIONES

El océano profundo engloba un conjunto de características que, no solo aportan a la naturaleza única, y por lo demás, extrema del ambiente, sino que estas suponen una serie de presiones adaptativas que fuerzan a los organismos piezófilos a incurrir en distintas estrategias de adecuación cuyo entendimiento aún está en curso. En este sentido, los plasmidios son ampliamente reconocidos por aportar a la adaptabilidad y evolución de los organismos procariontes pues juegan un importante rol en la THG; fenómeno del cual los ambientes marinos no se hayan exentos de efectos. A pesar de esto, sin embargo, en el océano profundo aún es escasa la información sobre la presencia y funciones codificadas en estos replicones independientes. Desde esta carencia se emplazan distintas preguntas científicas enfocadas en la Fosa de Atacama, una trinchera oceánica ubicada frente a las costas de nuestro país, cuya profundidad es máxima a los 8081 metros y que presenta alta productividad en sus aguas someras, y que no reúne estudios sobre los plásmidos de los organismos que le habitan. En esta línea, de las preguntas científicas previamente planteadas surgen los objetivos pilares de esta memoria. Así entonces, se buscó identificar las secuencias plasmidiales de las comunidades de la Fosa de Atacama y caracterizarlas genómicamente en miras de comprender la relación entre su contenido y funciones codificadas con el ambiente estudiado. En primera instancia, en el marco del Objetivo Específico 1, se logra establecer un protocolo que permitió la identificación satisfactoria de contigs plasmidiales a partir de metagenomas de distintos puntos de la columna de agua, siendo estos a los 500, 1000, 2000, 7928, 8079 y 8081 metros de profundidad, de los cuales, tras analizar la cantidad de contig plasmidiales, fueron considerados estadísticamente relevantes solo los de 500, 2000, 7928 y 8079 metros. Como parte del Objetivo Específico 2, se logra la anotación genómica de las secuencias plasmidiales identificadas, echando mano de la metodología propuesta. A partir de estos datos obtenidos, se procede a un análisis comparativo basado en las distintas categorías COG, cuyo código de clasificación se incluye en la anotación generada por Prokka. Avanzando hacia el Objetivo Específico 3, como se deja en evidencia en el *heatmap* (Figura 2) y en el gráfico de burbujas (Figura 3), es manifiesta la diferencia de contenido en los plasmidomas, donde las 22 categorías COG identificadas se exhiben con distintas abundancias relativas, destacando en general la

categoría “P: Metabolismo y transporte de iones inorgánicos” con fuerte presencia en todos los plasmidomas, y predominando en la zona hadal las categorías “G: Metabolismo y transporte de carbohidratos” e “I: Metabolismo y transporte de lípidos” en contraste con su presencia en los plasmidomas de aguas menos profundas, mientras que la categoría “M: biogénesis de pared, membrana y envoltura celular” presenta abundancias destacables a todas las profundidades estudiadas, incluso en la zona hadal. Las categorías ya mencionadas fueron analizadas más acabadamente, junto con “E: Metabolismo y transporte de aminoácidos”, “U: Tráfico intracelular, secreción y transporte vesicular” y “X: Moviloma: profagos, transposones”, todas con presencia en la zona hadal, profundizando en el detalle de los 5 genes más abundantes por categoría en cada profundidad estudiada.

Los resultados otorgan una perspectiva acerca de las funciones que están representadas en los plasmidios de las comunidades procariontes de las profundidades aquí estudiadas, pudiendo ser estas un reflejo de las necesidades adaptativas de los organismos que surgen a partir de las presiones adaptativas del ambiente, dígase de la disponibilidad de materia orgánica y la composición de esta, las altas presiones hidrostáticas, las características fisicoquímicas e hidrológicas de las profundidades aludidas, la presencia y distribución de distintos metales pesados, la capacidad de movilización y conjugación de los plasmidios, por nombrar algunas. En suma, los resultados presentados se hallan en conformidad con el Objetivo General de este trabajo, logrando exitosamente la identificación y caracterización genómica de plásmidos en metagenomas de las aguas oceánicas de la Fosa de Atacama.

Si bien no todas las funciones fueron analizadas en detalle, y aun cuando no se ha analizado la totalidad de la columna de agua, lo expuesto en este trabajo exploratorio supone una primera base, un primer punto en el estudio de plásmidos de las comunidades procariontes de la Fosa de Atacama, área que sin duda ofrece un amplio marco de posibilidades de investigación, de preguntas e inquietudes a resolver que, de permitírseles el ejercicio investigativo, sin duda nutrirán los saberes a cerca de la adaptación y evolución de los organismos hadales y del entramado de su relación con diversas presiones adaptativas, aún más en virtud del ambiente único de la Fosa de Atacama, el cual aviva las curiosidades más profundas y que paulatinamente se apresta a la exploración de sus enigmas.

Finalmente, el conjunto de resultados aquí expuestos permite llegar a las siguientes conclusiones:

Conclusión 1: Existen secuencias de naturaleza plasmidial en todos los metagenomas analizados provenientes de distintas profundidades, incluso en aquellos de profundidades mayores (7928, 8079 y 8081 metros de profundidad) dentro la Fosa de Atacama.

Conclusión 2: Las funciones codificadas mejor representadas en los plasmidomas hadales fueron: Metabolismo y transporte de iones inorgánicos (P), Metabolismo y transporte de carbohidratos (G) y metabolismo y transporte de lípidos (I)

Conclusión 3: Existen diferencias a nivel de tipo de función celular codificada y de la abundancia relativa de tales funciones en los plasmidomas de las distintas profundidades, donde en los plasmidomas someros abundan los genes asociados a la producción y conversión de energía "C", mientras que en los plasmidomas hadales estudiados destaca el metabolismo y transporte de carbohidratos "G", por dar ejemplo.

8. REFERENCIAS

1. Jamieson, A.J.J.e., *Ecology of deep oceans: hadal trenches*. 2011.
2. Liu, R., et al., *The hadal biosphere: Recent insights and new directions*. 2018. **155**: p. 11-18.
3. Jamieson, A.J., et al., *Hadal trenches: the ecology of the deepest places on Earth*. 2010. **25**(3): p. 190-197.
4. Du, M., et al., *Geology, environment, and life in the deepest part of the world's oceans*. 2021. **2**(2): p. 100109.
5. Kawagucci, S., et al., *Hadal water biogeochemistry over the Izu–Ogasawara Trench observed with a full-depth CTD-CMS*. 2018. **14**(4): p. 575-588.
6. Nunoura, T., et al., *Hadal biosphere: Insight into the microbial ecosystem in the deepest ocean on Earth*. 2015. **112**(11): p. E1230-E1236.
7. Ichino, M.C., et al., *The distribution of benthic biomass in hadal trenches: a modelling approach to investigate the effect of vertical and lateral organic matter transport to the seafloor*. 2015. **100**: p. 21-33.
8. Glud, R.N., et al., *High rates of microbial carbon turnover in sediments in the deepest oceanic trench on Earth*. 2013. **6**(4): p. 284-288.
9. Luo, M., et al., *Sources, degradation, and transport of organic matter in the New Britain Shelf - Trench continuum, Papua New Guinea*. 2019. **124**(6): p. 1680-1695.
10. Luo, M., et al., *Benthic carbon mineralization in hadal trenches: insights from in situ determination of benthic oxygen consumption*. 2018. **45**(6): p. 2752-2760.
11. Smeller, L.J.B.e.B.A.-P.S. and M. Enzymology, *Pressure–temperature phase diagrams of biomolecules*. 2002. **1595**(1-2): p. 11-29.
12. Balny, C., P. Masson, and K.J.B.e.b.a. Heremans, *High pressure effects on biological macromolecules: from structural changes to alteration of cellular processes*. 2002. **1595**(1-2): p. 3-10.
13. Macgregor Jr, R.B.J.B.e.B.A.-P.S. and M. Enzymology, *The interactions of nucleic acids at elevated hydrostatic pressure*. 2002. **1595**(1-2): p. 266-276.
14. Winter, R. and C.J.S.M. Jeworrek, *Effect of pressure on membranes*. 2009. **5**(17): p. 3157-3173.
15. Northrop, D.B.J.B.e.B.A.-P.S. and M. Enzymology, *Effects of high pressure on enzymatic activity*. 2002. **1595**(1-2): p. 71-79.
16. Oger, P.M. and M.J.R.i.m. Jebbar, *The many ways of coping with pressure*. 2010. **161**(10): p. 799-809.
17. Fang, J., L. Zhang, and D.A.J.T.i.m. Bazyliniski, *Deep-sea piezosphere and piezophiles: geomicrobiology and biogeochemistry*. 2010. **18**(9): p. 413-422.
18. Kusube, M., et al., *Colwellia marinimaniae sp. nov., a hyperpiezophilic species isolated from an amphipod within the Challenger Deep, Mariana Trench*. 2017. **67**(4): p. 824-831.
19. Yayanos, A.A., A.S. Dietz, and R.J.P.o.t.N.A.o.S. Van Boxtel, *Obligately barophilic bacterium from the Mariana trench*. 1981. **78**(8): p. 5212-5215.
20. Nogi, Y., et al., *Colwellia piezophila sp. nov., a novel piezophilic species from deep-sea sediments of the Japan Trench*. 2004. **54**(5): p. 1627-1631.
21. Kato, C., et al., *Extremely barophilic bacteria isolated from the Mariana Trench, Challenger Deep, at a depth of 11,000 meters*. 1998. **64**(4): p. 1510-1513.
22. Lauro, F.M., et al., *The unique 16S rRNA genes of piezophiles reflect both phylogeny and adaptation*. 2007. **73**(3): p. 838-845.

23. Nogi, Y., et al., *Psychromonas kaikoe sp. nov., a novel from the deepest piezophilic bacterium cold-seep sediments in the Japan Trench*. 2002. **52**(5): p. 1527-1532.
24. Nogi, Y., et al., *Psychromonas hadalis sp. nov., a novel piezophilic bacterium isolated from the bottom of the Japan Trench*. 2007. **57**(6): p. 1360-1364.
25. Zhang, W.-J., et al., *Comparative genomic analysis of obligately piezophilic Moritella yanosii DB21MT-5 reveals bacterial adaptation to the Challenger Deep, Mariana Trench*. 2021. **7**(7).
26. Rodriguez-Beltran, J., et al., *Beyond horizontal gene transfer: the role of plasmids in bacterial evolution*. Nat Rev Microbiol, 2021. **19**(6): p. 347-359.
27. Brito, I.L., *Examining horizontal gene transfer in microbial communities*. Nat Rev Microbiol, 2021.
28. Bouma, J.E. and R.E. Lenski, *Evolution of a bacteria/plasmid association*. Nature, 1988. **335**(6188): p. 351-2.
29. del Solar, G., et al., *Replication and control of circular bacterial plasmids*. Microbiol Mol Biol Rev, 1998. **62**(2): p. 434-64.
30. Banu, H. and K.P. Prasad, *Role of Plasmids in Microbiology*. Journal of Aquaculture Research & Development, 2017. **08**(01).
31. Hernandez-Arriaga, A.M., et al., *Conditional Activation of Toxin-Antitoxin Systems: Postsegregational Killing and Beyond*. Microbiol Spectr, 2014. **2**(5).
32. Pinilla-Redondo, R., et al., *Type IV CRISPR-Cas systems are highly diverse and involved in competition between plasmids*. Nucleic Acids Res, 2020. **48**(4): p. 2000-2012.
33. Tock, M.R. and D.T.J.C.o.i.m. Dryden, *The biology of restriction and anti-restriction*. 2005. **8**(4): p. 466-472.
34. Smillie, C., et al., *Mobility of plasmids*. Microbiol Mol Biol Rev, 2010. **74**(3): p. 434-52.
35. Chen, J., et al., *Genome hypermobility by lateral transduction*. 2018. **362**(6411): p. 207-212.
36. Clark, A.J. and G.J. Warren, *Conjugal transmission of plasmids*. Annu Rev Genet, 1979. **13**: p. 99-125.
37. Erdmann, S., et al., *A plasmid from an Antarctic haloarchaeon uses specialized membrane vesicles to disseminate and infect plasmid-free cells*. Nat Microbiol, 2017. **2**(10): p. 1446-1455.
38. Norman, A., L.H. Hansen, and S.J. Sorensen, *Conjugative plasmids: vessels of the communal gene pool*. Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1527): p. 2275-89.
39. Abe, K., N. Nomura, and S.J.F.m.e. Suzuki, *Biofilms: hot spots of horizontal gene transfer (HGT) in aquatic environments, with a focus on a new HGT mechanism*. 2020. **96**(5): p. f1aa031.
40. Aminov, R.I.J.F.i.m., *Horizontal gene exchange in environmental microbiota*. 2011. **2**: p. 158.
41. Giménez, M., et al., *Horizontal gene transfer elements: plasmids in Antarctic microorganisms*, in *The Ecological Role of Micro-organisms in the Antarctic Environment*. 2019, Springer. p. 85-107.
42. Mootapally, C., M.S. Mahajan, and N.M. Nathani, *Sediment Plasmidome of the Gulfs of Kathiawar Peninsula and Arabian Sea: Insights Gained from Metagenomics Data*. Microb Ecol, 2021. **81**(2): p. 540-548.
43. Saak, C.C., C.B. Dinh, and R.J. Dutton, *Experimental approaches to tracking mobile genetic elements in microbial communities*. FEMS Microbiol Rev, 2020. **44**(5): p. 606-630.

44. Ma, Y., I.T. Paulsen, and B. Palenik, *Analysis of two marine metagenomes reveals the diversity of plasmids in oceanic environments*. Environ Microbiol, 2012. **14**(2): p. 453-66.
45. Palenik, B., et al., *Coastal Synechococcus metagenome reveals major roles for horizontal gene transfer and plasmids in population diversity*. 2009. **11**(2): p. 349-359.
46. Petersen, J., et al., *Extrachromosomal, extraordinary and essential—the plasmids of the Roseobacter clade*. 2013. **97**(7): p. 2805-2815.
47. Michael, V., et al., *Biofilm plasmids with a rhamnose operon are widely distributed determinants of the 'swim-or-stick' lifestyle in roseobacters*. 2016. **10**(10): p. 2498-2513.
48. Petersen, J., et al., *A marine plasmid hitchhiking vast phylogenetic and geographic distances*. Proc Natl Acad Sci U S A, 2019. **116**(41): p. 20568-20573.
49. Cuadrat, R.R.C., et al., *Global ocean resistome revealed: Exploring antibiotic resistance gene abundance and distribution in TARA Oceans samples*. Gigascience, 2020. **9**(5).
50. Liu, Y., et al., *Characteristics and Evolutionary Analysis of Photosynthetic Gene Clusters on Extrachromosomal Replicons: from Streamlined Plasmids to Chromids*. mSystems, 2019. **4**(5).
51. Wang, S., et al., *Conjugative transfer of Megaplasmids pND6-1 and pND6-2 enhancing naphthalene degradation in aqueous environment: characterization and bioaugmentation prospects*. Appl Microbiol Biotechnol, 2020. **104**(2): p. 861-871.
52. Dorrell, R.G., et al., *Phylogenomic fingerprinting of tempo and functions of horizontal gene transfer within ochrophytes*. 2021. **118**(4).
53. Brinkmann, H., et al., *Horizontal operon transfer, plasmids, and the evolution of photosynthesis in Rhodobacteraceae*. ISME J, 2018. **12**(8): p. 1994-2010.
54. Petersen, J. and I. Wagner-Dobler, *Plasmid Transfer in the Ocean - A Case Study from the Roseobacter Group*. Front Microbiol, 2017. **8**: p. 1350.
55. Zhang, R.-Y., et al., *The complete genome of extracellular protease-producing Deinococcus sp. D7000 isolated from the hadal region of Mariana Trench Challenger Deep*. 2021. **57**: p. 100832.
56. Liu, P., et al., *Physiological and genomic features of Paraoceanicella profunda gen. nov., sp. nov., a novel piezophile isolated from deep seawater of the Mariana Trench*. 2020. **9**(2): p. e966.
57. Chen, M., et al., *Genomic Characteristics and Potential Metabolic Adaptations of Hadal Trench Roseobacter and Alteromonas Bacteria Based on Single-Cell Genomics Analyses*. 2020. **11**.
58. Lemenkova, P., *Geomorphological modelling and mapping of the Peru-Chile Trench by GMT*. Polish Cartographical Review, 2019. **51**(4): p. 181-194.
59. Angel, M.V.J.E., *Ocean trench conservation*. 1982. **2**(1): p. 1-17.
60. Ahumada, R.J.B.P., *Producción y destino de la biomasa fitoplanctónica en un sistema de bahías en Chile central: una hipótesis*. 1989. **18**: p. 53-66.
61. Schneider, W., et al., *Extensión vertical y horizontal de la zona de mínima oxígeno en el Pacífico Sur Oriental*. 2006. **70**: p. 79-82.
62. Olivares, F., *Caracterización taxonómica de las bacterias y arqueas planctónicas de la Fosa de Atacama*, in *Facultad de Ciencias Naturales y Oceanográficas*. 2022, Universidad de Concepción: Concepción, Chile.
63. Nurk, S., et al., *metaSPAdes: a new versatile metagenomic assembler*. Genome Res, 2017. **27**(5): p. 824-834.

64. Schwengers, O., et al., *Platon: identification and characterization of bacterial plasmid contigs in short-read draft assemblies exploiting protein sequence-based replicon distribution scores*. *Microb Genom*, 2020. **6**(10).
65. Hyatt, D., et al., *Prodigal: prokaryotic gene recognition and translation initiation site identification*. 2010. **11**(1): p. 1-11.
66. Buchfink, B., C. Xie, and D.H.J.N.m. Huson, *Fast and sensitive protein alignment using DIAMOND*. 2015. **12**(1): p. 59-60.
67. Kurtz, S., et al., *Versatile and open software for comparing large genomes*. 2004. **5**(2): p. 1-9.
68. Mistry, J., et al., *Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions*. 2013. **41**(12): p. e121-e121.
69. Nawrocki, E.P., *Annotating functional RNAs in genomes using Infernal*, in *RNA sequence, structure, and function: computational and bioinformatic methods*. 2014, Springer. p. 163-197.
70. Camacho, C., et al., *BLAST+: architecture and applications*. 2009. **10**(1): p. 1-9.
71. Lagesen, K., et al., *RNAMmer: consistent and rapid annotation of ribosomal RNA genes*. 2007. **35**(9): p. 3100-3108.
72. Laslett, D. and B.J.N.a.r. Canback, *ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences*. 2004. **32**(1): p. 11-16.
73. Almagro Armenteros, J.J., et al., *SignalP 5.0 improves signal peptide predictions using deep neural networks*. 2019. **37**(4): p. 420-423.
74. research, U.C.J.N.a., *UniProt: a worldwide hub of protein knowledge*. 2019. **47**(D1): p. D506-D515.
75. O'Leary, N.A., et al., *Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation*. 2016. **44**(D1): p. D733-D745.
76. Mistry, J., et al., *Pfam: The protein families database in 2021*. 2021. **49**(D1): p. D412-D419.
77. Haft, D.H., et al., *TIGRFAMs: a protein family resource for the functional identification of proteins*. 2001. **29**(1): p. 41-43.
78. Webb, E.C., *Enzyme nomenclature 1992. Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes*. 1992: Academic Press.
79. Galperin, M.Y., et al., *COG database update: focus on microbial diversity, model organisms, and widespread pathogens*. 2021. **49**(D1): p. D274-D281.
80. Kolde, R. and M.R.J.R.P. Kolde, *Package 'pheatmap'*. 2018. **1**.
81. Forward, J.A., et al., *TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of Rhodobacter capsulatus and by homologs in diverse gram-negative bacteria*. 1997. **179**(17): p. 5482-5493.
82. Kelly, D.J., et al., *High-Affinity C4-Dicarboxylate Uptake in Rhodobacter Capsulatus Is Mediated By A 'Trap'Transporter, A New Type of Periplasmic Secondary Transport System Widespread in Bacteria*, in *The Phototrophic Prokaryotes*. 1999, Springer. p. 573-582.
83. Ullmann, R., et al., *Transport of C4-dicarboxylates in Wolinella succinogenes*. 2000. **182**(20): p. 5757-5764.
84. Shinabarger, D., et al., *Purification and characterization of phosphomannose isomerase-guanosine diphospho-D-mannose pyrophosphorylase. A bifunctional enzyme in the alginate biosynthetic pathway of Pseudomonas aeruginosa*. 1991. **266**(4): p. 2080-2088.
85. Hay, I.D., et al., *Genetics and regulation of bacterial alginate production*. 2014. **16**(10): p. 2997-3011.

86. Buckel, S., et al., *An analysis of the structure of the product of the rbsA gene of Escherichia coli K12*. 1986. **261**(17): p. 7659-7662.
87. Park, Y. and C.J.J.o.b. Park, *Topology of RbsC, a membrane component of the ribose transporter, belonging to the AraH superfamily*. 1999. **181**(3): p. 1039-1042.
88. Björkman, A.J. and S.L.J.J.o.m.b. Mowbray, *Multiple open forms of ribose-binding protein trace the path of its conformational change*. 1998. **279**(3): p. 651-664.
89. Clifton, M.C., et al., *In vitro reassembly of the ribose ATP-binding cassette transporter reveals a distinct set of transport complexes*. 2015. **290**(9): p. 5555-5565.
90. Kim, J.J.P. and R.J.E.J.o.B. Miura, *Acyl - CoA dehydrogenases and acyl - CoA oxidases: Structural basis for mechanistic similarities and differences*. 2004. **271**(3): p. 483-493.
91. Fujita, Y., H. Matsuoka, and K.J.M.m. Hirooka, *Regulation of fatty acid metabolism in bacteria*. 2007. **66**(4): p. 829-839.
92. Jatana, N., et al., *Molecular modeling studies of fatty acyl-CoA synthetase (FadD13) from Mycobacterium tuberculosis—a potential target for the development of antitubercular drugs*. 2011. **17**(2): p. 301-313.
93. Chuakrut, S., et al., *Characterization of a bifunctional archaeal acyl coenzyme A carboxylase*. 2003. **185**(3): p. 938-947.
94. Hügler, M., et al., *Characterization of acetyl - CoA/propionyl - CoA carboxylase in Metallosphaera sedula: Carboxylating enzyme in the 3 - hydroxypropionate cycle for autotrophic carbon fixation*. 2003. **270**(4): p. 736-744.
95. Gulick, A.M.J.A.c.b., *Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase*. 2009. **4**(10): p. 811-827.
96. Elssner, T., et al., *Involvement of coenzyme A esters and two new enzymes, an enoyl-CoA hydratase and a CoA-transferase, in the hydration of crotonobetaine to L-carnitine by Escherichia coli*. 2001. **40**(37): p. 11140-11148.
97. Adams, M.D., et al., *Nucleotide sequence and genetic characterization reveal six essential genes for the LIV-I and LS transport systems of Escherichia coli*. 1990. **265**(20): p. 11436-11443.
98. Hosie, A.H. and P.S.J.R.i.m. Poole, *Bacterial ABC transporters of amino acids*. 2001. **152**(3-4): p. 259-270.
99. Tucker, D.L., N. Tucker, and T.J.J.o.B. Conway, *Gene expression profiling of the pH response in Escherichia coli*. 2002. **184**(23): p. 6551-6558.
100. *Database of Clusters of Orthologous Genes (COGs)* 2022 March, 2022 [cited 2022; Available from: <https://www.ncbi.nlm.nih.gov/research/cog/cog/COG0620/>].
101. Tseng, T.-T., et al., *The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins*. 1999. **1**(1): p. 107-125.
102. Nies, D.H.J.F.m.r., *Efflux-mediated heavy metal resistance in prokaryotes*. 2003. **27**(2-3): p. 313-339.
103. Franke, S., et al., *Molecular analysis of the copper-transporting efflux system CusCFBA of Escherichia coli*. 2003. **185**(13): p. 3804-3812.
104. Rensing, C., et al., *CopA: an Escherichia coli Cu (I)-translocating P-type ATPase*. 2000. **97**(2): p. 652-656.
105. Singh, S.K., et al., *Cuprous oxidase activity of CueO from Escherichia coli*. 2004. **186**(22): p. 7815-7817.
106. Liu, H., et al., *The connection between czc and cad systems involved in cadmium resistance in Pseudomonas putida*. 2021. **22**(18): p. 9697.

107. Anton, A., et al., *CzcD is a heavy metal ion transporter involved in regulation of heavy metal resistance in Ralstonia sp. strain CH34*. 1999. **181**(22): p. 6876-6881.
108. Barnawi, H., et al., *RNA-based thermoregulation of a Campylobacter jejuni zinc resistance determinant*. 2020. **16**(10): p. e1009008.
109. Hohle, T.H. and M.R.J.M.m. O'Brian, *The mntH gene encodes the major Mn²⁺ transporter in Bradyrhizobium japonicum and is regulated by manganese via the Fur protein*. 2009. **72**(2): p. 399-409.
110. Cabezón, E., et al., *Towards an integrated model of bacterial conjugation*. 2015. **39**(1): p. 81-95.
111. Alvarez-Martinez, C.E., P.J.J.M. Christie, and M.B. Reviews, *Biological diversity of prokaryotic type IV secretion systems*. 2009. **73**(4): p. 775-808.
112. Beaupre, C.E., et al., *Interactions between VirB9 and VirB10 membrane proteins involved in movement of DNA from Agrobacterium tumefaciens into plant cells*. 1997. **179**(1): p. 78-89.
113. Bayliss, R., et al., *NMR structure of a complex between the VirB9/VirB7 interaction domains of the pKM101 type IV secretion system*. 2007. **104**(5): p. 1673-1678.
114. Rambow-Larsen, A.A. and A.A.J.J.o.b. Weiss, *The PtlE protein of Bordetella pertussis has peptidoglycanase activity required for Ptl-mediated pertussis toxin secretion*. 2002. **184**(11): p. 2863-2869.
115. Martin, F.A., et al., *Interplay between two RND systems mediating antimicrobial resistance in Brucella suis*. 2009. **191**(8): p. 2530-2540.
116. Wang, C., et al., *Bartonella quintana type IV secretion effector BepE - induced selective autophagy by conjugation with K63 polyubiquitin chain*. 2019. **21**(4): p. e12984.
117. Li, X.-Z. and H.J.D. Nikaido, *Efflux-mediated drug resistance in bacteria*. 2009. **69**(12): p. 1555-1623.
118. Simonsen, L.J.M., *Dynamics of plasmid transfer on surfaces*. 1990. **136**(6): p. 1001-1007.
119. Bergstrom, C.T., M. Lipsitch, and B.R.J.G. Levin, *Natural selection, infectious transfer and the existence conditions for bacterial plasmids*. 2000. **155**(4): p. 1505-1519.
120. San Millan, A. and R.C. MacLean, *Fitness Costs of Plasmids: a Limit to Plasmid Transmission*. Microbiol Spectr, 2017. **5**(5).
121. Levin, B.R., F.M. Stewart, and V.A.J.P. Rice, *The kinetics of conjugative plasmid transmission: fit of a simple mass action model*. 1979. **2**(2): p. 247-260.
122. Slater, F.R., et al., *Progress towards understanding the fate of plasmids in bacterial communities*. FEMS Microbiol Ecol, 2008. **66**(1): p. 3-13.
123. Dionisio, F., et al., *Plasmids spread very fast in heterogeneous bacterial communities*. 2002. **162**(4): p. 1525-1532.
124. Dimitriu, T., et al., *Bacteria from natural populations transfer plasmids mostly towards their kin*. 2019. **286**(1905): p. 20191110.
125. Mamontov, V., et al., *Persistence of plasmids targeted by CRISPR interference in bacterial populations*. 2022. **119**(15): p. e2114905119.
126. Kozlowski, B.K., M. Dworkin, and G.M.J.I.j.o.m.m. Dunny, *Pheromone-inducible conjugation in Enterococcus faecalis: a model for the evolution of biological complexity?* 2006. **296**(2-3): p. 141-147.
127. Oger, P. and S.K. Farrand, *Two opines control conjugal transfer of an Agrobacterium plasmid by regulating expression of separate copies of the quorum-sensing activator gene traR*. 2002, Am Soc Microbiol.

128. Sherburne, C.K., et al., *The complete DNA sequence and analysis of R27, a large IncHI plasmid from Salmonella typhi that is temperature sensitive for transfer*. 2000. **28**(10): p. 2177-2186.
129. Chao, L. and B.R.J.P.o.t.N.A.o.S. Levin, *Structured habitats and the evolution of anticompetitor toxins in bacteria*. 1981. **78**(10): p. 6324-6328.
130. Molin, S. and T.J.C.o.i.b. Tolker-Nielsen, *Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure*. 2003. **14**(3): p. 255-261.
131. Slater, F.R., et al., *Heterogeneous selection in a spatially structured environment affects fitness tradeoffs of plasmid carriage in pseudomonads*. 2008. **74**(10): p. 3189-3197.
132. Oguri, K., et al., *Hadal disturbance in the Japan Trench induced by the 2011 Tohoku–Oki Earthquake*. 2013. **3**(1): p. 1-6.
133. Turnewitsch, R., et al., *Recent sediment dynamics in hadal trenches: evidence for the influence of higher-frequency (tidal, near-inertial) fluid dynamics*. 2014. **90**: p. 125-138.
134. Münch, K., et al., *Evolutionary model for the unequal segregation of high copy plasmids*. 2019. **15**(3): p. e1006724.
135. Sayers, E.W., et al., *GenBank*. 2021. **49**(D1): p. D92-D96.
136. Koren, S. and A.M.J.C.o.i.m. Phillippy, *One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly*. 2015. **23**: p. 110-120.
137. Bentley, S.D. and J.J.A.r.o.g. Parkhill, *Comparative genomic structure of prokaryotes*. 2004. **38**(1): p. 771-791.
138. Raes, J., et al., *Prediction of effective genome size in metagenomic samples*. 2007. **8**(1): p. 1-11.
139. de Toro, M., M. Pilar Garcillán-Barcia, and F. de la Cruz, *Plasmid Diversity and Adaptation Analyzed by Massive Sequencing of Escherichia coli Plasmids*, in *Plasmids*. 2015. p. 219-235.
140. Sieber, P., M. Platzer, and S.J.T.i.G. Schuster, *The definition of open reading frame revisited*. 2018. **34**(3): p. 167-170.
141. Eløe, E.A., et al., *Going deeper: metagenome of a hadopelagic microbial community*. PLoS One, 2011. **6**(5): p. e20388.
142. Tazzyman, S.J., S.J.M.b. Bonhoeffer, and evolution, *Why there are no essential genes on plasmids*. 2014. **32**(12): p. 3079-3088.
143. Fernández - Urruzola, I., et al., *Plankton respiration in the Atacama Trench region: Implications for particulate organic carbon flux into the hadal realm*. 2021. **66**(8): p. 3134-3148.
144. Mulligan, C., M. Fischer, and G.H.J.F.m.r. Thomas, *Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea*. 2011. **35**(1): p. 68-86.
145. Vezzi, A., et al., *Life at depth: Photobacterium profundum genome sequence and expression analysis*. 2005. **307**(5714): p. 1459-1461.
146. Tang, K., et al., *Genomic, physiologic, and proteomic insights into metabolic versatility in Roseobacter clade bacteria isolated from deep-sea water*. 2016. **6**(1): p. 1-12.
147. Rehm, B.H.J.N.R.M., *Bacterial polymers: biosynthesis, modifications and applications*. 2010. **8**(8): p. 578-592.
148. Clementi, F.J.C.r.i.b., *Alginate production by Azotobacter vinelandii*. 1997. **17**(4): p. 327-361.

149. Inoue, A., et al., *Discovery of a novel alginate lyase from Nitratiruptor sp. SB155-2 thriving at deep-sea hydrothermal vents and identification of the residues responsible for its heat stability*. 2016. **291**(30): p. 15551-15563.
150. Boyd, A. and A.M. Chakrabarty, *Role of alginate lyase in cell detachment of Pseudomonas aeruginosa*. 1994. **60**(7): p. 2355-2359.
151. Schwertfeger, M., *Alterations of calcium alginate gel dispersions by high pressure treatment*, in *Advances in high pressure bioscience and biotechnology*. 1999, Springer. p. 337-340.
152. Oates, C. and D.J.F.H. Ledward, *Studies on the effect of heat on alginates*. 1990. **4**(3): p. 215-220.
153. Fraenkel, D.J.E.c., S.t. cellular, and W. molecular biology. American Society for Microbiology, DC, *Glycolysis, pentose phosphate pathway, and Entner-Doudoroff pathway*. 1987: p. 142-150.
154. Sprenger, G.A.J.A.o.m., *Genetics of pentose-phosphate pathway enzymes of Escherichia coli K-12*. 1995. **164**(5): p. 324-330.
155. Hove-Jensen, B., *Mutation in the phosphoribosylpyrophosphate synthetase gene (prs) that results in simultaneous requirements for purine and pyrimidine nucleosides, nicotinamide nucleotide, histidine, and tryptophan in Escherichia coli*. 1988. **170**(3): p. 1148-1152.
156. Wiegand, S., M. Jogler, and C. Jogler, *On the maverick Planctomycetes*. FEMS Microbiology Reviews, 2018. **42**(6): p. 739-760.
157. Benner, R., et al., *Bulk chemical characteristics of dissolved organic matter in the ocean*. 1992. **255**(5051): p. 1561-1564.
158. Dong, Y., et al., *Characterization of two long-chain fatty acid CoA ligases in the Gram-positive bacterium Geobacillus thermodenitrificans NG80-2*. 2012. **167**(10): p. 602-607.
159. Wakeham, S.G., et al., *Molecular indicators of diagenetic status in marine organic matter*. 1997. **61**(24): p. 5363-5369.
160. Flores, E., et al., *Intact polar lipids in the hadal seabed of the Atacama Trench point to lateral sediment transport and in situ production as 2 key sources of labile organic matter 3*. 2021. **2021**: p. 1-42.
161. Aono, E., et al., *Complete genome sequence and comparative analysis of Shewanella violacea, a psychrophilic and piezophilic bacterium from deep sea floor sediments*. 2010. **6**(7): p. 1216-1226.
162. Ghisla, S. and C.J.E.j.o.b. Thorpe, *Acyl - CoA dehydrogenases: a mechanistic overview*. 2004. **271**(3): p. 494-508.
163. Morgan-Kiss, R.M. and J.E.J.J.o.B.C. Cronan, *The Escherichia coli fadK (ydiD) gene encodes an anerobically regulated short chain acyl-CoA synthetase*. 2004. **279**(36): p. 37324-37333.
164. Yoshida, K., et al., *Bacterial long-chain polyunsaturated fatty acids: their biosynthetic genes, functions, and practical use*. 2016. **14**(5): p. 94.
165. Fang, J., et al., *Biochemical implications and geochemical significance of novel phospholipids of the extremely barophilic bacteria from the Marianas Trench at 11,000 m*. 2000. **47**(6): p. 1173-1182.
166. Burton, N.P., T.D. Williams, and P.R.J.A.o.m. Norris, *Carboxylase genes of Sulfolobus metallicus*. 1999. **172**(6): p. 349-353.
167. Menendez, C., et al., *Presence of acetyl coenzyme A (CoA) carboxylase and propionyl-CoA carboxylase in autotrophic Crenarchaeota and indication for operation of a 3-hydroxypropionate cycle in autotrophic carbon fixation*. 1999. **181**(4): p. 1088-1098.

168. STRAUSS, G. and G.J.E.J.o.B. FUCHS, *Enzymes of a novel autotrophic CO₂ fixation pathway in the phototrophic bacterium Chloroflexus aurantiacus, the 3 - hydroxypropionate cycle*. 1993. **215**(3): p. 633-643.
169. Rodriguez, E., et al., *Role of an essential acyl coenzyme A carboxylase in the primary and secondary metabolism of Streptomyces coelicolor A3 (2)*. 2001. **67**(9): p. 4166-4176.
170. Alonso, C. and J.J.E.m. Pernthaler, *Roseobacter and SAR11 dominate microbial glucose uptake in coastal North Sea waters*. 2006. **8**(11): p. 2022-2030.
171. Alonso-Sáez, L., J.M.J.A. Gasol, and e. microbiology, *Seasonal variations in the contributions of different bacterial groups to the uptake of low-molecular-weight compounds in northwestern Mediterranean coastal waters*. 2007. **73**(11): p. 3528-3535.
172. Price, M.N., A.M. Deutschbauer, and A.P.J.P.g. Arkin, *Four families of folate-independent methionine synthases*. 2021. **17**(2): p. e1009342.
173. Goulding, C.W., D. Postigo, and R.G.J.B. Matthews, *Cobalamin-dependent methionine synthase is a modular protein with distinct regions for binding homocysteine, methyltetrahydrofolate, cobalamin, and adenosylmethionine*. 1997. **36**(26): p. 8082-8091.
174. Gonzalez, J.C., *Studies on cobalamin-independent methionine synthase from Escherichia coli: A zinc metalloenzyme*. 1996: University of Michigan.
175. Djoko, K.Y., et al., *Interplay between tolerance mechanisms to copper and acid stress in Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America, 2017. **114**(26): p. 6818-6823.
176. Landick, R., et al., *Regulation of high - affinity leucine transport in escherichia coli*. 1980. **14**(4): p. 527-537.
177. Gaye, B., et al., *What can we learn from amino acids about oceanic organic matter cycling and degradation?* 2022. **19**(3): p. 807-830.
178. Bánfalvi, G., *Cellular effects of heavy metals*. 2011: Springer.
179. Chen, P.R. and C.J.C.o.i.c.b. He, *Selective recognition of metal ions by metalloregulatory proteins*. 2008. **12**(2): p. 214-221.
180. Halliwell, B. and J.J.B.j. Gutteridge, *Oxygen toxicity, oxygen radicals, transition metals and disease*. 1984. **219**(1): p. 1.
181. Delmar, J.A., C.C. Su, and E.W.J.P.S. Yu, *Heavy metal transport by the Cus CFBA efflux system*. 2015. **24**(11): p. 1720-1736.
182. Ares, Á., et al., *Sequestration and efflux largely account for cadmium and copper resistance in the deep sea epsilonproteobacterium, Nitratiruptor sp. SB155-2*. 2021. **28**: p. 29.
183. Outten, F.W., et al., *The independent cue and cusSystems confer copper tolerance during aerobic and anaerobic growth in Escherichia coli*. 2001. **276**(33): p. 30670-30677.
184. Argüello, J.M., E. Eren, and M.J.B. González-Guerrero, *The structure and function of heavy metal transport P1B-ATPases*. 2007. **20**(3): p. 233-248.
185. Thieme, D., et al., *Sandwich hybridization assay for sensitive detection of dynamic changes in mRNA transcript levels in crude Escherichia coli cell extracts in response to copper ions*. 2008. **74**(24): p. 7463-7470.
186. Smedile, F., et al., *Metagenomic analysis of hadopelagic microbial assemblages thriving at the deepest part of Mediterranean Sea, Matapan - Vavilov Deep*. 2013. **15**(1): p. 167-182.
187. Koschinsky, A., et al., *Reactions of the heavy metal cycle to industrial activities in the deep sea: An ecological assessment*. 2003. **88**(1): p. 102-127.

188. Chu, N.-C., et al., *Evidence for hydrothermal venting in Fe isotope compositions of the deep Pacific Ocean through time*. 2006. **245**(1-2): p. 202-217.
189. Bennett, S.A., et al., *The distribution and stabilisation of dissolved Fe in deep-sea hydrothermal plumes*. 2008. **270**(3-4): p. 157-167.
190. German, C., et al., *Heat, volume and chemical fluxes from submarine venting: A synthesis of results from the Rainbow hydrothermal field, 36 N MAR*. 2010. **57**(4): p. 518-527.
191. Hauton, C., et al., *Identifying toxic impacts of metals potentially released during deep-sea mining—a synthesis of the challenges to quantifying risk*. 2017. **4**: p. 368.
192. Schlitzer, R. *eGEOTRACES - Electronic Atlas of GEOTRACES Sections and Animated 3D Scenes*. 2021 [cited 2022; Available from: <http://www.egeotraces.org>].
193. Fairey, R., et al., *An evaluation of methods for calculating mean sediment quality guideline quotients as indicators of contamination and acute toxicity to amphipods by chemical mixtures*. 2001. **20**(10): p. 2276-2286.
194. Reid, W.D., N.J. Cuomo, and A.J.J.D.S.R.P.I.O.R.P. Jamieson, *Geographic and bathymetric comparisons of trace metal concentrations (Cd, Cu, Fe, Mn, and Zn) in deep-sea lysianassoid amphipods from abyssal and hadal depths across the Pacific Ocean*. 2018. **138**: p. 11-21.
195. Nozaki, Y., et al., *The distribution of radionuclides and some trace metals in the water columns of the Japan and Bonin trenches*. 1998. **21**(3): p. 469-484.
196. Bruland, K.W.J.E. and P.S. Letters, *Oceanographic distributions of cadmium, zinc, nickel, and copper in the North Pacific*. 1980. **47**(2): p. 176-198.
197. Bruland, K., M.J.T.o. Lohan, and m. geochemistry, *Controls of trace metals in seawater*. 2006. **6**: p. 23-47.
198. Obata, H., et al., *Dissolved iron and zinc in Sagami Bay and the Izu-Ogasawara Trench*. 2017. **73**(3): p. 333-344.
199. John, S.G., J. Helgoe, and E.J.M.C. Townsend, *Biogeochemical cycling of Zn and Cd and their stable isotopes in the Eastern Tropical South Pacific*. 2018. **201**: p. 256-262.
200. Johnson, K.S., et al., *On the formation of the manganese maximum in the oxygen minimum*. 1996. **60**(8): p. 1291-1299.
201. Li, Y.G., B. Hu, and P.J.J.M.s. Christie, *Biological and structural diversity of type IV secretion systems*. 2019. **7**(2): p. 7.2. 30.
202. Waksman, G.J.E.r., *From conjugation to T4S systems in Gram - negative bacteria: a mechanistic biology perspective*. 2019. **20**(2): p. e47012.
203. Cascales, E. and P.J.J.N.R.M. Christie, *The versatile bacterial type IV secretion systems*. 2003. **1**(2): p. 137-149.
204. Guglielmini, J., et al., *Evolution of conjugation and type IV secretion systems*. 2013. **30**(2): p. 315-331.
205. Arechaga, I., et al., *ATPase activity and oligomeric state of TrwK, the VirB4 homologue of the plasmid R388 type IV secretion system*. 2008. **190**(15): p. 5472-5479.
206. Walldén, K., et al., *Structure of the VirB4 ATPase, alone and bound to the core complex of a type IV secretion system*. 2012. **109**(28): p. 11348-11353.
207. Moran, M.A., et al., *Ecological Genomics of Marine Roseobacters*. 2007. **73**(14): p. 4559-4569.
208. Persson, O.P., et al., *High abundance of virulence gene homologues in marine bacteria*. 2009. **11**(6): p. 1348-1357.
209. Siguier, P., E. Gourgouyere, and M.J.F.m.r. Chandler, *Bacterial insertion sequences: their genomic impact and diversity*. 2014. **38**(5): p. 865-891.

210. DeLong, E.F., et al., *Community genomics among stratified microbial assemblages in the ocean's interior*. 2006. **311**(5760): p. 496-503.
211. Konstantinidis, K.T., et al., *Comparative Metagenomic Analysis of a Microbial Community Residing at a Depth of 4,000 Meters at Station ALOHA in the North Pacific Subtropical Gyre*. 2009. **75**(16): p. 5345-5355.
212. Martin-Cuadrado, A.-B., et al., *Metagenomics of the deep Mediterranean, a warm bathypelagic habitat*. 2007. **2**(9): p. e914.
213. Chen, M., et al., *Genomic characteristics and potential metabolic adaptations of hadal trench Roseobacter and Alteromonas bacteria based on single-cell genomics analyses*. 2020. **11**: p. 1739.
214. Hillmann, F., M. Argentini, and N.J.J.o.B.C. Buddelmeijer, *Kinetics and phospholipid specificity of apolipoprotein N-acyltransferase*. 2011. **286**(32): p. 27936-27946.
215. Nielsen, J. and J.J.J.o.b. Lampen, *Glyceride-cysteine lipoproteins and secretion by Gram-positive bacteria*. 1982. **152**(1): p. 315-322.
216. Paradis-Bleau, C., et al., *Lipoprotein cofactors located in the outer membrane activate bacterial cell wall polymerases*. 2010. **143**(7): p. 1110-1120.
217. Fu, Y., *Structure and dynamics of Pseudomonas aeruginosa ICP*, in *College of Medical, Veterinary and Life Sciences*. 2009, University of Glasgow: Scotland.
218. Trefzer, A., A. Bechthold, and J.A.J.N.p.r. Salas, *Genes and enzymes involved in deoxysugar biosynthesis in bacteria*. 1999. **16**(3): p. 283-299.
219. Kirschning, A., et al., *Chemical and biochemical aspects of deoxysugars and deoxysugar oligosaccharides*. 1997: p. 1-84.
220. Poulsen, S.M., C. Kofoed, and B.J.J.o.m.b. Vester, *Inhibition of the ribosomal peptidyl transferase reaction by the mycarose moiety of the antibiotics carbomycin, spiramycin and tylosin*. 2000. **304**(3): p. 471-481.