



Instituto de Investigaciones Biológicas Clemente Estable
Departamento de Bioquímica y Genómica Microbianas
Grupo Microbiología Molecular

Tesis para optar al título de Doctora en Ciencias Biológicas
opción Microbiología
PEDECIBA
Área Biología
Sub-área Microbiología

TRANSFERENCIA HORIZONTAL DE GENES EN
BACTERIAS ANTÁRTICAS

VERÓNICA BEATRIZ ANTELO GUTIÉRREZ

Tutora: Dra. Silvia Batista (IIBCE)
Co-tutor: Didier Mazel (Institut Pasteur París)

Montevideo
Uruguay
2019

Tribunal

Dra. Claudia Etchebehere

(Departamento de Bioquímica y Genómica Microbianas, IIBCE)

Dra. Carolina Márquez

(Cátedra de Microbiología, Laboratorio de Microbiología Clínica,

Facultad de Química, UdelaR)

Dr. Hugo Naya

(Unidad de Bioinformática, Institut Pasteur Montevideo)

Todos alabamos la belleza del árbol, todos elogiamos el encanto de la flor, todos saboreamos el fruto maduro, todos escanciamos el zumo que fermenta... y pocos, muy pocos, se acuerdan de la oscura raíz que trabaja en profundidad... Eso ocurre con la Ciencia Pura y la Ciencia Aplicada.

Clemente Estable

A mis padres, Ulises y Beatriz

AGRADECIMIENTOS

Quiero agradecer a todas las personas que me apoyaron y me acompañaron durante este largo camino y que de diferentes maneras han hecho posible la realización de este trabajo.

En primer lugar, quiero agradecer a mi directora Silvia Batista por haberme dado la oportunidad de seguir creciendo, por guiarme y apoyarme en todo momento. También quiero reconocer especialmente su esfuerzo en las madrugadas de desvelo corrigiendo esta tesis.

A Didier Mazel por brindarme sus conocimientos durante el transcurso de este proyecto.

A todas las personas que de una forma u otra han participado en este trabajo:

Daniela Senatore
Daniela Costa
Héctor Romero
Matías Giménez
Ismael Díaz Isasa
Karen Malán
Patricia Valdespino
Gregorio Iraola
Ignacio Ferréz
Pablo Fresia
Sebastián Pita
Cecilia Salazar
José Sotelo
Martín Graña
Raúl Platero

A mis compañeros del Instituto de Investigaciones Biológicas Clemente Estable por todos los años de trabajo compartidos.

A mis compañeros del Laboratorio de Genómica Microbiana y de la Unidad de Bionformática del Institut Pasteur Montevideo.

A mi familia y amigos.

A Carmen por escucharme y alentarme siempre.

A mi hermano Rober.

Y muy especialmente quiero agradecer a Lany y mis hijos Adrián y Marcel por su gran apoyo y por alentarme a seguir adelante.

Gracias!

Resumen

Los integrones son plataformas genéticas capaces de incorporar, intercambiar y expresar nuevos genes. Mediante eventos de recombinación específica de sitio integran marcos de lectura abierta, asegurando su correcta expresión. A fines de la década del 80 se caracterizaron los primeros integrones en bacterias de origen clínico. Asimismo, estudios posteriores permitieron determinar que estas estructuras son frecuentes en el genoma bacteriano.

El análisis de diversidad de integrones en ambientes naturales sugiere que podrían desempeñar un papel fundamental en la adaptación y evolución bacteriana. Numerosos estudios han sido desarrollados en ambientes diversos empleando diferentes abordajes, sin embargo, ambientes considerados “extremos” como la Antártida han sido poco explorados.

El objetivo de este estudio fue identificar y analizar elementos de transferencia horizontal genética en bacterias antárticas terrestres que se desarrollan en la Isla Rey Jorge (Antártida marítima) y el continente antártico en general, con especial atención a los integrones que pudiesen estar asociados. El estudio incluyó el uso de técnicas dependientes e independientes del cultivo, incorporando distintos tipos de sesgo para la búsqueda de estas estructuras.

La estrategia de trabajo consistió en: 1-establecer la presencia y organización genética de integrones en microorganismos antárticos aislados de muestras de sedimento-suelo y tapetes microbianos, colectadas en la Isla Rey Jorge. Los sitios de muestreo incluyeron zonas afectadas por la actividad humana (cerca de las bases), presencia animal (pingüíneras) y otras zonas presuntamente prístinas (ZAEP 125d); 2- analizar la organización y el contexto genómico de integrones clase 1 de tipo “clínico”, identificados previamente en tres enterobacterias antárticas seleccionadas en medio LB con antibiótico.; 3- analizar la diversidad de genes integrasa y genes *cassette* (GCs) mediante métodos independientes del cultivo, en muestras colectadas en la isla Rey Jorge (sedimento, tapetes microbianos, agua de lagos, cañadas y agua de mar) en sitios seleccionados con

el mismo criterio del punto 1 y en tapetes microbianos colectados en distintos sitios del continente antártico.

1-En el marco de la estrategia dependiente del cultivo, se analizaron aislamientos obtenidos a 5°C en medio específico para oligotrofos (R2A). Se emplearon cebadores dirigidos a amplificar integrasas de integrones clase 1, 2 y 3. Se obtuvieron resultados positivos únicamente en cinco aislamientos de la misma muestra (Lago Uruguay, I. Rey Jorge), un integrón de clase 1 y cuatro integrones de clase 2. El análisis en un medio no selectivo (sin presencia de antibióticos) podría explicar la dificultad para encontrar integrones en organismos cultivables.

2- Se trabajó con tres aislamientos de enterobacterias antárticas psicotolerantes (CN11, TR5 y HP19) resistentes a Ampicilina, Estreptomina y Trimetoprim, previamente obtenidos. Estos aislamientos poseen integrones clase 1 de tipo “clínico” asociados al gen de resistencia a Trimetoprim. Los ensayos de conjugación con *Escherichia coli* DH5α indicaron que los integrones estaban localizados a nivel plasmídico. Mediante secuenciación *shotgun* se determinó la secuencia de los plásmidos pCN11, pTR5 y pHP19. El plásmido pHP19 fue secuenciado de forma completa y ensamblado en 12 contigs. El análisis de secuencia a nivel nucleotídico exhibió alta similitud (99%) con el plásmido p34998-210.894kb de *Enterobacter hormaechei* subsp. *steigerwaltii* de origen clínico (CP012169.1), el cual contiene un integrón clase 1 de tipo “clínico”, similar al de nuestro aislamiento.

3- Las técnicas independientes de cultivo incluyeron la construcción y análisis de bibliotecas de amplicones del gen *intI* y el análisis *in silico* de metagenomas totales de tapetes microbianos obtenidos mediante secuenciación masiva. Para la búsqueda de integrones y GCs asociados, se empleó el programa Integron-Finder. El análisis de las bibliotecas de clones indica que los cebadores usados para amplificar integrasas clase 1, no son específicos para esta clase, dado que fueron diseñados considerando un conjunto reducido de secuencias (en ese momento se disponía sólo de secuencias de origen clínico). El estudio *in silico* de metagenomas permitió identificar una gran diversidad de integrasas y GCs asociados. Gran parte de los GCs codifican para funciones desconocidas. La mayoría de las integrasas exhibieron similitud con integrasas ambientales

descritas en diferentes regiones geográficas. En las bibliotecas, los integrones de tipo “clínico” encontrados correspondieron a integrones de clase 1, mientras que en los metagenomas los integrones “clínicos” fueron de clase 2 y clase 3. En ambos casos, este tipo de integrones representan una pequeña fracción del total de genes *intI* identificados. El análisis de los árboles filogenéticos obtenidos con las secuencias ambientales recuperadas sugiere que deberían replantearse los criterios para la clasificación de integrones.

Las distintas estrategias utilizadas indican que el análisis metagenómico resultó ser el más “exitoso” para este estudio, en tanto permitió identificar una mayor diversidad de integrones y GCs, los cuales en su mayoría codifican para funciones desconocidas.

Palabras claves: Antártida; bacteria; integrones

Tabla de contenido

CAPÍTULO 1	15
1.1-Introducción general	15
1.1.1- Antártida.....	15
1.1.2- Regiones biogeográficas de la Antártida	15
1.1.3-Antártida Marítima: Isla Rey Jorge	17
1.1.4-Tapetes microbianos	19
1.1.5-Suelos antárticos	19
1.1.6-Estrategias para el estudio de comunidades microbianas antárticas	21
1.2-Transferencia Horizontal Genética	22
1.2.1-Mecanismos de transferencia horizontal genética.....	23
1.2.2-Elementos genéticos móviles del genoma bacteriano	25
1.2.3-Conjugación	26
1.3-Integrones	27
1.3.1-Introducción.....	27
1.3.2-Estructura de un integrón	28
1.3.3-Clasificación.....	29
1.3.4-Sitio <i>attI</i>	30
1.3.5-Sitios <i>attC</i>	30
1.3.6-La enzima integrasa IntI.....	32
1.3.7-Reacciones de recombinación de la enzima IntI	34
1.3.8-Expresión de la integrasa (IntI)	36
1.3.9- Genes <i>cassette</i> (GCs)	39
1.3.10-Integrones cromosómicos sedentarios (ICSs)	43
1.3.11-Integrones móviles (IM).....	45
1.3.12-Origen de los integrones clase 1 clínicos.....	49
1.3.13- Origen viral de los integrones?	51
CAPÍTULO 2	52
2-Estrategias dependientes de cultivo	52
2.1- Antecedentes e Introducción	53
Objetivo General	59
Objetivo Específico	59
2.2-Materiales y Métodos	60
2.2.1-Procesamiento de muestras y aislamientos de microorganismos heterótrofos para la búsqueda de integrones clase 1 (<i>intI1</i>), clase 2 (<i>intI2</i>) y clase 3 (<i>intI3</i>).....	60
2.2.2-Búsqueda de genes <i>intI1</i> , <i>intI2</i> e <i>intI3</i> en organismos cultivables.	61
2.3-Resultados	61
2.3.1- Búsqueda de integrones clase 1, 2 y 3:.....	61
2.3.2.-Conclusiones y Perspectivas	63

2.4 -Materiales y Métodos.....	64
2.4.1-Ensayos de conjugación bacteriana.....	64
2.4.2- Secuenciación plasmídica en plataforma Ion Torrent	64
2.4.3- Secuenciación genómica en plataforma Illumina	65
2.4.4-Procesamiento de datos, anotación y análisis de secuencias.....	65
2.5-Resultados	66
2.5.1-Ensayos de conjugación.....	66
2.5.2- Secuenciación plasmídica	66
2.6-Discusión	78
2.7-Conclusiones y Perspecivas	82
2.8-Conclusiones generales	83
CAPÍTULO 3.....	84
3. Estrategias independientes de cultivo	84
3.1 Introducción.....	85
3.1.1- Integrones en ecosistemas naturales	85
3.1.2- Descripción de las herramientas empleadas para el análisis in silico de metagenomas	85
i) Servidor MG-RAST	85
ii) Programa Integron Finder para la identificación y anotación de integrones.....	87
iii) Anotación funcional de genes con eggNOG mapper.....	90
Objetivo General.....	93
Objetivo Específico	93
3.2- Construcción de bibliotecas de amplicones intI.....	94
3.2.1-Materiales y Métodos	94
3.2.2 Resultados	97
3.3-Análisis taxonómico y funcional <i>in silico</i> de metagenomas de tapetes microbianos y sedimento/suelo antárticos. Identificación de genes integrasas intI y GCs asociados.	100
3.3.1-Materiales y Métodos.....	100
i) Sitios de muestreo	100
ii) Búsqueda de genes integrón integrasa (<i>intI</i>) en metagenomas de tapetes microbianos y sedimento/suelo completamente secuenciados:	101
iii) Anotación de GCs de resistencia a antibióticos (Base de datos Resfam).....	101
iv) Análisis filogenético de integrasas antárticas con secuencias IntI de referencia	102
v) Contexto genómico de los integrones completos identificados por el programa Integron-Finder	102
vi) Anotación funcional genómica de GCs mediante la asignación de ortología con eggNOG-mapper.....	103
vii) Composición taxonómica de la comunidad en metagenomas antárticos	103
.....	104

3.3.2.Resultados	105
i)Identificación de genes que codifican para integrasas de integrones (IntI) en metagenomas antárticos.....	105
ii)Anotación de GCs de resistencia a antibióticos (Base de datos Resfam)	106
iii)Análisis filogenético de IntI antárticas con IntI de referencia	106
iv) Contexto genómico de integrones de metagenomas de tapetes microbianos y sedimento antárticos.	118
v) Anotación funcional de GCs y ORFs con eggNOG-mapper	131
vi) Composición taxonómica de la comunidad de metagenomas antárticos.	136
3.3.3- Discusión	143
3.3.4-Conclusiones	146

Material suplementario

1.-Tablas	148
2.-Mapas sitios de muestreo	158
3.- Composición taxonómica de la comunidad.....	161
4.-Composición de medios de cultivo:.....	165
Referencias Bibliográficas	166

CAPÍTULO 1

1.1-Introducción general

1.1.1- Antártida

La Antártida es el continente más austral de la Tierra, posee un área de 12,3 millones de Km², lo que representa aproximadamente el 10% de la superficie terrestre mundial (Campbell et al., 2009). La mayor parte del continente antártico posee una cobertura permanente de hielo, en tanto las áreas expuestas (rocas, suelo/sedimento) representan alrededor del 0,34% de la superficie total (Convey & Stevens, 2007). La capa de hielo se extiende por encima del océano, alcanzando los 14 millones de km² en verano. Debe destacarse, sin embargo, que la capa de hielo es un indicador importante del cambio climático y el principal responsable del incremento del nivel del mar. Mediante estudios satelitales y análisis por modelado (Sheperd et al., 2018), se estima que la Antártida perdió 2,720±1,390 miles de millones de toneladas de hielo entre 1992 y 2017, lo que corresponde a un aumento en el nivel medio del mar de 7.6 ± 3.9 mm. Durante este período, la Antártida Occidental ha incrementado la pérdida de hielo en un 300%; y en tanto en la Península Antártica, el aumento fue de 470%.

1.1.2- Regiones biogeográficas de la Antártida

En la Zona de Convergencia Antártica clásicamente se han identificado tres regiones que difieren en cuanto al clima y biota asociada: Antártida Marítima, Antártida Continental y región Subantártica (Convey, P., 2010)(Figura 1.1). Estas tres regiones se diferencian principalmente por las temperaturas medias anuales, precipitaciones y vientos, lo que se asocia a una biota característica. La Antártida Continental abarca la mayor parte del continente antártico. Dentro de la región Continental se destaca una zona especialmente inclemente, McMurdo Dry Valleys, considerada como desierto polar. Las mayores precipitaciones se registran en la Antártida Marítima (400-600 mm/año), con temperaturas medias

anuales entre 1°C y -21°C. Esta región, con un clima menos riguroso que el Continental, abarca el NW de la Península Antártica e islas aledañas. La biota incluye líquenes, briofitas, algunas plantas vasculares, invertebrados, aves y mamíferos marinos. También se desarrollan tapetes microbianos de composición y apariencia diversa.

La región Subantártica incluye varias islas ubicadas entre 35°S y 60°S. El clima de estas islas permite el establecimiento de sitios de cría de aves y mamíferos marinos, generándose así un paisaje que contrasta con el resto del continente. La incorporación de nuevos datos y criterios de clasificación ha permitido la identificación de 16 Regiones Biogeográficas de Conservación de la Antártida (Terauds & Lee, 2016). Estas bio-regiones, ubicadas en las áreas sin cobertura de hielo permanente, son reconocidas como sitios con características singulares y estudiadas con diferentes objetivos, incluido el análisis de biodiversidad.

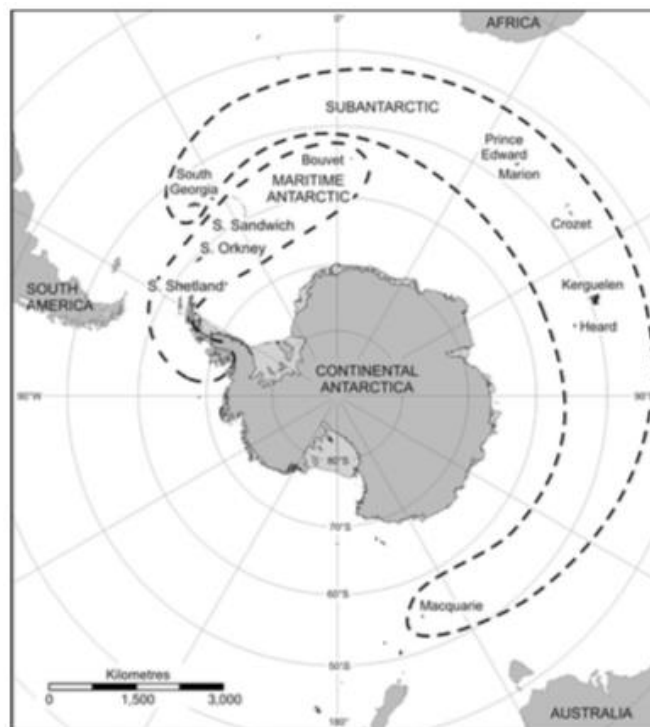


Figura 1.1: Zonas biogeográficas terrestres de la Antártida (Antártida Marítima, Antártida Continental y Subantártica), tomado de (Convey, 2010).

1.1.3-Antártida Marítima: Isla Rey Jorge

De acuerdo con la clasificación tradicional, el archipiélago South Shetland está ubicado en la Antártida Marítima. Dentro de ese archipiélago, la isla Rey Jorge es la más extensa y visitada del continente Antártico. Posee diez bases destinadas a la investigación. Durante el verano austral, esta isla recibe la visita de más de 500 personas, donde se llevan a cabo numerosos proyectos de investigación en diversas áreas (Figura 1.2). La isla se encuentra cubierta por el Glaciar Bellingshausen (Collins), aunque existen varios sitios sin cobertura permanente de hielo (Figura 1.2), ubicados en las zonas costeras. Dentro de estos sitios se destaca la península Fildes, el área sin cobertura permanente más extensa de la isla, en donde se alojan cuatro bases antárticas, incluyendo la Base Científica Antártica Artigas (BCAA) de Uruguay. Durante la estación estival, el paisaje de esta península adquiere un aspecto particular que contrasta con las áreas adyacentes cubiertas por el glaciar. El paisaje rocoso, con algunos lagos, se cubre de pequeños arroyos y cañadas de deshielo que se forman con agua proveniente del glaciar y nieve fundida. A medida que avanza el verano se van formando verdaderas alfombras de briofitas, líquenes y tapetes microbianos, que generan un paisaje singular.

La base BCAA está ubicada al lado del lago Uruguay, cerca del Glaciar Collins. La base trabaja de forma permanente, manteniendo una dotación básica de unas 8 personas. Durante los meses más cálidos, las instalaciones alojan además a investigadores, funcionarios y otras personas que participan en las campañas de verano.

Las actividades que se desarrollan en la Antártida están regidas por el Protocolo al Tratado Antártico sobre Protección del Medio Ambiente (1998) firmado en Madrid en 1991. El Tratado establece normas claras que regulan la actividad humana en el continente, intentando reducir al mínimo el impacto humano en la flora y fauna nativas para la preservación del ecosistema antártico. Las Zonas Antárticas Especialmente Protegidas (ZAEP) (Antarctic Specially Protected Area; ASPA) son áreas del continente con cierto valor científico, estético, histórico (fósiles) o natural destacable, que amerita ser administradas de forma particular. En la actualidad se han definido 72 ZAEPs, lo que implica la necesidad de acceder

a un permiso especial para ingresar, el cual es otorgado por alguno de los países que participan en el Tratado Antártico.

Por otro lado, las mayores tasas de calentamiento global se han registrado en la Antártida Marítima y el glaciar Bellingshausen es un claro testimonio de esa tendencia (Turner et al., 2005; Turner et al., 2007). Al retraerse, queda expuesto sustrato que es rápidamente colonizado por organismos nativos y exóticos.

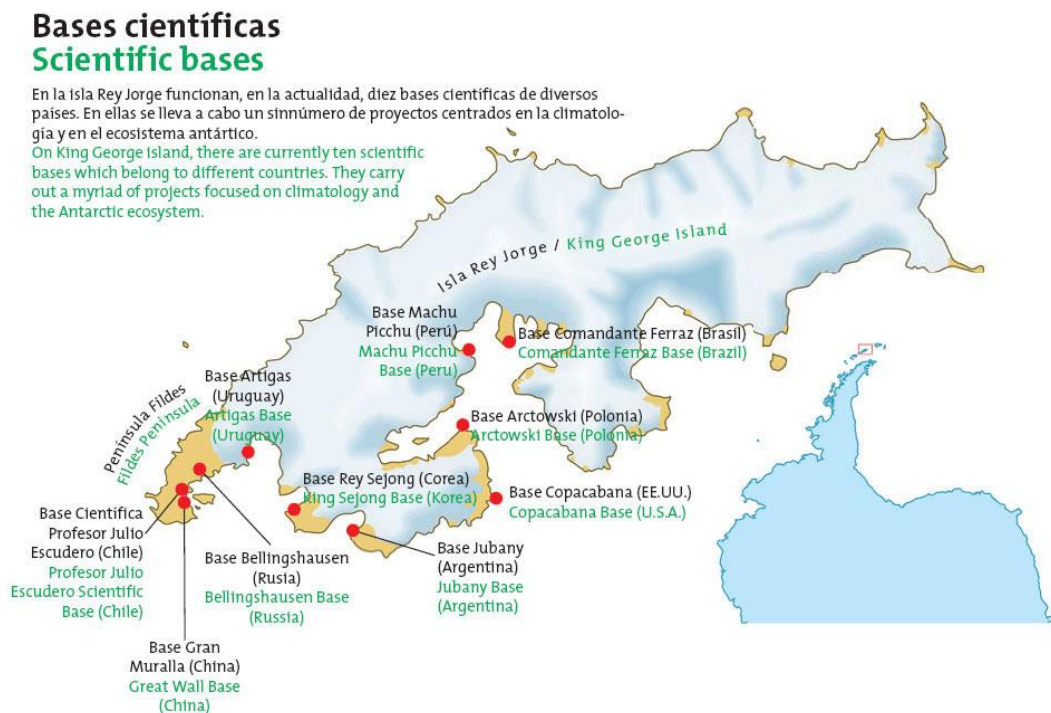


Figura 1.2: Mapa de la isla Rey Jorge, en donde se muestra el glaciar Collins y de forma aproximada, la ubicación de las bases científicas. (Extraído de <http://www.icarito.cl/2010/07/bases-cientificas-en-la-isla-rey-jorge.shtml/>).

1.1.4-Tapetes microbianos

Dentro del paisaje de la Península Fildes se destacan los tapetes microbianos, comunidades bentónicas muy diversas que se establecen en lagunas y cañadas de deshielo (Callejas et al., 2011; Callejas, et al., 2018; J. Komárek & Komárek, 2013; Komarek & Komarek, 1999). Los tapetes son capaces de incorporar de la atmósfera el C y N necesario para generar biomasa y sostener así las redes tróficas a partir de la interacción de sus diversos componentes microbianos (Alcántara-Hernández et al., 2014; Valdespino-Castillo et al., 2018). Sin embargo, se conoce relativamente poco la diversidad, biogeografía y patrones funcionales de tapetes microbianos antárticos. Su estudio podría aportar información sobre su funcionamiento en el presente y predecir su evolución frente a un aumento de la antropogenización y cambio ambiental.

1.1.5-Suelos antárticos

Las propiedades de los suelos antárticos varían con las distintas topografías, condiciones climáticas y biota asociada. En las áreas costeras de la Antártida Marítima, los suelos suelen ser ricos en nutrientes por la presencia de aves y mamíferos marinos. En las zonas montañosas o sitios especialmente secos como McMurdo Dry-Valleys, predominan los suelos pobres en nutrientes u oligotrofos (Cowan 2014).

En la Isla Rey Jorge, el retroceso de los glaciares ha permitido la colonización de las zonas expuestas por especies vegetales y la subsecuente formación de suelo. En las zonas costeras se puede observar el desarrollo de plantas vasculares como *Deschampsia antarctica* y *Colobantus quitensis* y algunas especies de líquenes y algas, claros indicadores de un suelo en incipiente desarrollo (Bölter, 2011). Las áreas libres de hielo próximas a la costa son colonizadas por vertebrados marinos. Las heces de aves y mamíferos enriquecidas en carbono (C) orgánico, nitrógeno (N) y fósforo (P), contribuyen a definir las propiedades fisicoquímicas del suelo y la composición de las comunidades bacterianas.

Diversos estudios dan cuenta de cómo las características del suelo y la estructura de las comunidades microbianas asociadas están íntimamente relacionadas. Ganzert y col., (2011) observaron que los suelos asociados a musgos (Isla Livingston, Islas Shetland del Sur) suelen estar enriquecidos en organismos del

grupo Bacteroidetes, mientras que los suelos oligotróficos son dominados por integrantes del filo Acidobacteria (Ganzert et al., 2011).

Asimismo, en ese estudio establecieron que los niveles de C y N total son altos en suelos con vegetación, a diferencia de los suelos mineralizados que tienen niveles de C y N bajos. Respecto al pH, su valor es neutro o alcalino en suelos mineralizados y ácido en suelos recubiertos por musgos (Ganzert et al., 2011).

Un estudio reciente desarrollado en Cape Shirreff (Isla Livingston), designada como zona ZAEP 149, determinó que la diversidad de las comunidades de microorganismos en suelos ubicados en asentamientos de aves y mamíferos marinos, viene determinada por la concentración de especies de vertebrados que los habitan. Las diferencias en la dieta y hábitos de vida se ve reflejada en la composición química de sus heces, las cuales impactan directamente en la estructura del suelo (Ramírez-Fernández et al., 2019). En suelos enriquecidos con fecas animales son abundantes los grupos Bacteroidetes, Actinobacteria y Chloroflexi, mientras que Gemmatimonadetes, en particular el género *Gemmatimonas*, está pobremente representado en suelos asociados con vertebrados marinos. Este género es abundante en suelos antárticos donde el agua y los nutrientes son escasos (Ramírez-Fernández et al., 2019).

Asimismo, el género *Rhodanobacter* es el más abundante en suelos asociados a vertebrados marinos (ricos en nutrientes). La correlación positiva entre *Rhodanobacter*, Bacteroidetes, Actinobacteria y Chloroflexi sugiere que sus miembros establecen asociaciones en suelos ricos en nitratos y que las especies de *Rhodanobacter* podrían tener un papel importante en la desnitrificación (Ramírez-Fernández et al., 2019). El mayor impacto en las propiedades fisicoquímicas del suelo se observó principalmente en áreas de asentamiento de colonias de pingüinos. El alto contenido en nutrientes inorgánicos como nitrato, fosfato y amonio afectan el desarrollo y crecimiento de la vegetación y por lo tanto disminuyen la diversidad de las comunidades microbianas (Ramírez-Fernández et al., 2019).

1.1.6-Estrategias para el estudio de comunidades microbianas antárticas

Los microorganismos son esenciales para el mantenimiento de diversos procesos, incluida la formación del suelo, la descomposición de los xenobióticos, el metabolismo de la materia orgánica y el apoyo de los ciclos biogeoquímicos de elementos específicos como C, N y P. Además, el estudio de microorganismos y sus interacciones con factores bióticos y abióticos podría utilizarse para el desarrollo de estrategias de biorremediación, procesos biotecnológicos, evaluación del manejo del suelo, etc.

Los primeros estudios microbiológicos en la Antártida se iniciaron durante el siglo XX (Darling & Siple, 1941; Straka & Stokes, 1960). Estos estudios incluyeron el cultivo y aislamiento de microorganismos y los resultados obtenidos respaldaron la idea existente de que la diversidad y abundancia de especies en los ensamblajes microbianos era relativamente baja.

La microscopía también se ha utilizado para evaluar la abundancia, sucesión y distribución de microorganismos en el suelo antártico, tapetes microbianos, etc. Algunos estudios microscópicos notables fueron desarrollados por (J. Komárek & Komárek, 2013; O. Komárek & Komárek, 1999), dirigidos a identificar cianobacterias en tapetes. Más tarde, desarrollando una estrategia polifásica, se integraron otros abordajes como criterios taxonómicos.

La incorporación de métodos moleculares permitió la detección de una mayor riqueza microbiana en las comunidades antárticas de lo que se anticipó con los métodos tradicionales de microscopía (Cowan, 2014). El uso de técnicas independientes del cultivo se centró en la descripción de la composición de la comunidad mediante análisis del gen 16S rRNA y de procesos fisiológicos potenciales, evaluando la presencia de genes funcionales, por ejemplo, genes de desnitrificación, nitrogenasa, etc. (Alcántara-Hernández et al., 2014; C. Callejas et al., 2011; Cecilia Callejas et al., 2018; Jungblut & Neilan, 2010). Se desarrollaron varios estudios moleculares relacionados, orientados a describir la estructura de la comunidad microbiana expuesta a diferentes ambientes. Estos métodos demostraron que las comunidades bacterianas de suelos antárticos eran relativamente diversas, pero algunas filas estaban generalmente presentes en cada comunidad, incluyendo Actinobacteria, Proteobacteria, Bacteroidetes, Acidobacteria, Gemmatimonadetes, Deinococcus-Thermus y

Cyanobacteria (Cowan, 2014). Además, estas comunidades contenían una alta proporción de microorganismos que no se habían descrito anteriormente.

(Jungblut & Neilan, 2010) determinaron la presencia de microorganismos fijadores de N₂ en tapetes microbianos bénticos de Orange Pond (McMurdo) utilizando métodos moleculares. Analizaron las secuencias del gen *nifH* a partir de bibliotecas de clones de transcriptos del gen, complementando también el estudio con la medición de actividades nitrogenasa, utilizando el ensayo de reducción de acetileno (ARA). En su estudio, se identificaron 18 filotipos de cianobacterias, así como Firmicutes, Beta, Gamma y Deltaproteobacteria, Espiroquetas y Verrumicrobia. Los transcriptos *nifH* pertenecían a *Nostoc* spp., lo que indica la importancia de estos organismos en la fijación de N₂ en estos sitios.

El cultivo de microorganismos antárticos continúa hoy en día como una herramienta importante para comprender los procesos fisiológicos, aplicaciones biotecnológicas e identificación de nuevos organismos, etc. (Peeters, Ertz, & Willems, 2011), obtuvieron una colección de más de 2.000 bacterias heterótrofas aisladas de muestras de tapetes microbianos acuáticos, recolectadas de diferentes sitios en la Antártida Continental y la Península Antártica. El análisis de las secuencias de 16S rRNA les permitió identificar cinco fila: Actinobacteria, Bacteroidetes, Proteobacteria, Firmicutes y Deinococcus-Thermus. Una proporción de estas secuencias (36.9%) fueron similares a otras previamente identificadas en el continente antártico.

1.2-Transferencia Horizontal Genética

Las nuevas herramientas moleculares para el estudio de la evolución han dejado en evidencia que la visión clásica que representa la relación entre los organismos y sus características del pasado con los organismos y caracteres del presente a través de ramas de un árbol, no siempre resulta la más adecuada para interpretar los procesos evolutivos (Boto, 2010). La secuenciación de genomas completos a partir de mitad de la década del 90 ha demostrado que la transferencia horizontal genética (THG) entre especies, fila e incluso dominios es un fenómeno más frecuente de lo pensado (Blakely, 2014).

La THG se define como el movimiento de material genético (genes completos, operones enteros o hasta fragmentos con secuencias no codificantes) entre organismos filogenéticamente diferentes y su integración estable en el genoma receptor (Doolittle, 1999a).

Este fenómeno tiende a ser más exitoso entre organismos emparentados. El genoma de organismos cercanos filogenéticamente contiene en general genes similares en secuencia, composición de bases y utilización de codones, lo que facilita su incorporación luego de la transferencia (Kloesges et al., 2011). Una vez incorporado, la mayor parte del material transferido permanecerá silenciado, salvo que sea replicado y expresado de forma exitosa, permitiendo la expresión de nuevas funciones en la célula receptora.

Desde los orígenes de la vida, la THG ha estado involucrada en el moldeado del genoma bacteriano. Se estima que entre un 1,6% y un 32,6% de los genes procariontes han sido adquiridos por este mecanismo (Koonin et al., 2002).

En bacterias, la THG es uno de los principales mecanismos de adaptación al ambiente, permitiendo la formación de genomas extraordinariamente heterogéneos y dinámicos (de la Cruz & Davies, 2000). La adquisición de nuevos genes permite formar genomas versátiles que facilitan una rápida respuesta a cambios del ambiente (Gillings, 2017).

La importancia de la THG ha influido significativamente en la ecología y patogenicidad de especies de bacterias patógenas, aunque también ha sido relevante en la evolución de bacterias no patógenas que habitan diversos ambientes (Ochman et al., 2000).

1.2.1-Mecanismos de transferencia horizontal genética

Clásicamente se conocen tres mecanismos de THG: conjugación, transformación y transducción (Figura 1.3).

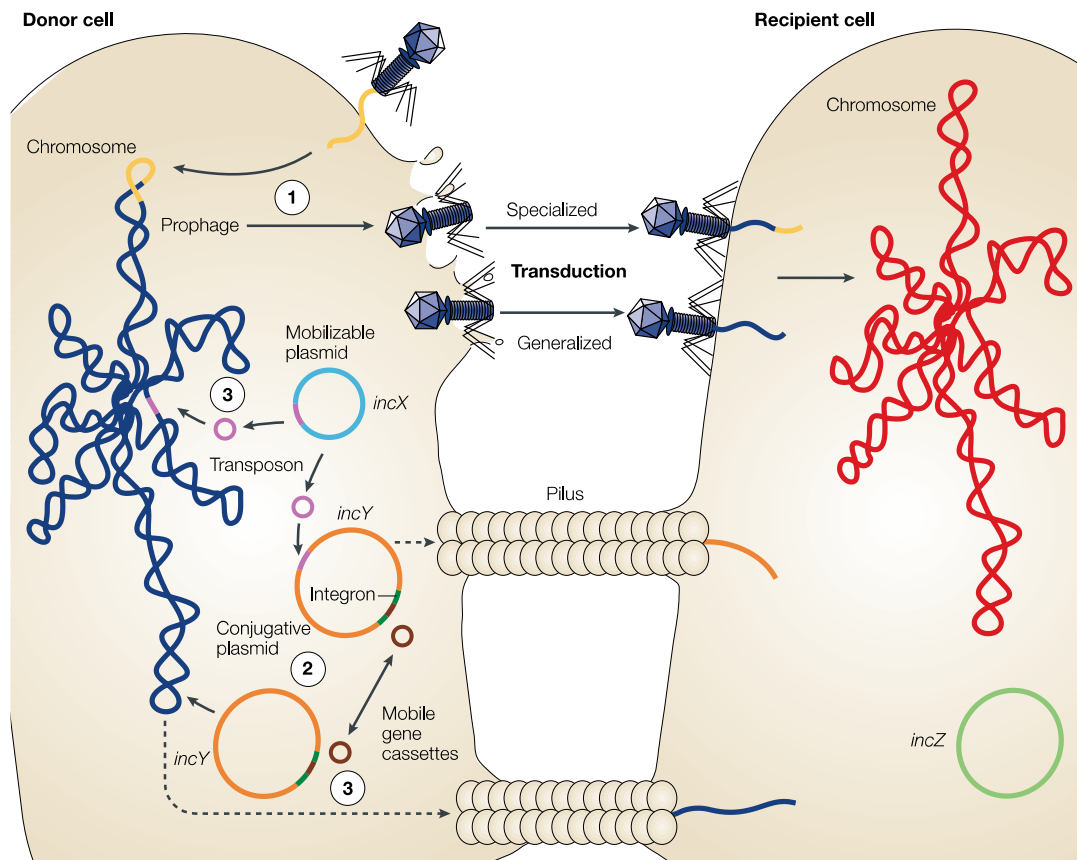


Figura 1.3: Mecanismos de transferencia de ADN entre bacterias. Tomado de (Frost, 2005).

La transformación fue el primer mecanismo de THG identificado en procariontes y refiere a la incorporación de ADN libre del ambiente, mediado por proteínas codificadas en el cromosoma bacteriano (Frost et al., 2005). Generalmente, la incorporación de fragmentos de ADN es independiente de su secuencia. Muchas especies de bacterias Gram negativas y Gram positivas pueden adquirir ADN ambiental. Algunas especies de bacterias son naturalmente competentes como *Neisseria gonorrhoeae* y *Haemophilus influenzae*, mientras que otras, como *Bacillus subtilis* y *Streptococcus pneumoniae*, sólo se vuelven competentes en determinadas etapas del ciclo de vida (Ochman et al., 2000).

La transducción es un proceso mediado por bacteriófagos (fago) autoreplicativos. Tras la infección de la célula, el ADN del fago se integra al cromosoma bacteriano (profago) en un proceso denominado conversión

lisogénica. El profago se transmite a la progenie de la bacteria infectada hasta que se escinde del cromosoma en el ciclo lítico y el ADN es liberado. Los fagos pueden empaquetar accidentalmente fragmentos de ADN del hospedero e inyectarlo en otro organismo receptor, donde por recombinación puede ser incorporado al cromosoma y por lo tanto ser heredado por la descendencia. La transducción posibilita de esta forma la diseminación de genes de resistencia a antibióticos de origen cromosómico.

La conjugación bacteriana es la transferencia de material genético desde una célula donadora a otra receptora mediante el contacto célula-célula. Este mecanismo fue descrito por primera vez por Lederberg y Tatum en 1946, quienes identificaron al factor-F (Fertilidad), un plásmido auto-transmisible de 100 kb (Lederberg y Tatum, 1946). Los plásmidos relacionados con el factor F serían los responsables de la emergencia de la resistencia a antimicrobianos. Estos elementos, que pueden ser transferidos a otras células y conferir fenotipos resistentes, se denominaron factores R (Resistencia) (Blakely, 2014).

Recientemente fue identificado un nuevo mecanismo de transferencia de ADN similar al empleado por los virus. Este mecanismo está mediado por un plásmido (pR1SE) aislado de una especie de haloarquea antártica, que codifica para proteínas que forman vesículas. El ADN plasmídico es empaquetado en estas vesículas, las cuales infectan otras células mediante la fusión con la membrana de la célula receptora. pR1SE se puede integrar y replicar como parte del genoma del huésped que así lo permita. Este mecanismo podría representar una estrategia alternativa empleada por los virus para transferir su propio ADN, sugiriendo que existe un nexo evolutivo entre plásmidos y virus (Erdmann et al., 2017).

1.2.2-Elementos genéticos móviles del genoma bacteriano

Los Elementos Genéticos Móviles (EGMs) son fragmentos de ADN que codifican para proteínas que median el movimiento de ADN en el genoma (intracelular) o entre células (intercelular) (Frost et al., 2005).

El movimiento intracelular de ADN mediado por transposones, involucra un proceso en el que recombinan al azar o saltan entre replicones y pueden ser transferidos junto con los fagos o plásmidos (Frost et al., 2005).

Los Elementos Integrativos Conjugativos (ICEs) también median la transferencia de ADN involucrando el contacto directo entre células. Los ICEs generalmente son auto-transmisibles, pero no auto-replicativos y sólo pueden duplicarse cuando son integrados al cromosoma o cuando se incorporan a otra célula durante la conjugación (Blakely, 2014).

Los ICEs incluyen transposones conjugativos como Tn916 (que requiere baja homología de secuencia) u otros elementos que necesitan reconocer secuencias específicas para su integración como el ICE SXT de *Vibrio cholerae* (Blakely, 2014).

Particularmente en bacterias Gram-negativas, ciertos elementos tales como secuencias de inserción ISs, el sistema integrón/gen cassette y transposones de la familia Tn3 y Tn5053 participan en la captura y movilización de genes de resistencia a antibióticos. Estos elementos aunque poseen características en común, difieren en los mecanismos utilizados para el movimiento de genes (Partridge, 2011).

1.2.3-Conjugación

La conjugación requiere elementos genéticos replicativos independientes que codifican para proteínas que facilitan su propia transferencia, como los plásmidos conjugativos y los elementos conjugativos integrados (ICEs) cromosómicos, que incluye los transposones conjugativos entre otros (Frost et al., 2005).

Los plásmidos son moléculas de ADN circular covalentemente cerrado y super-enrollado, auto-replicativos y extra-cromosomales, aunque también pueden existir plásmidos lineales. Los plásmidos se organizan en módulos como forma de asegurar su replicación, estabilidad y propagación a otras células. La forma de replicación varía considerablemente entre los distintos plásmidos. Generalmente portan genes que confieren ventajas de adaptación a la célula huésped tales como resistencia a drogas, metales pesados o genes que participan de rutas biodegradativas de xenobióticos (de la Cruz & Davies, 2000).

Según su capacidad de transferencia pueden ser conjugativos o movilizables. Los plásmidos conjugativos son aquellos que poseen la maquinaria necesaria para su auto-transferencia, mientras que los plásmidos movilizables requieren de proteínas codificadas en plásmidos conjugativos para ser transferidos. La

maquinaria necesaria para la movilización (denominada módulo MOB) incluye el origen de transferencia (*oriT*), una relaxasa y una proteína acopladora (T4CP). Todos los plásmidos transmisibles poseen una región MOB necesaria para la movilización, mientras que los auto-transmisibles tienen además el sistema de secreción tipo IV (T4SS). T4SS está formado por un grupo de genes que forman el módulo MPF (*mating pair formation*) que permite el ensamblaje y la formación del canal transmembrana que conecta ambas células (Smillie et al., 2010). Las relaxasas reconocen y se unen covalentemente al origen de transferencia *oriT*, catalizan el clivaje inicial y la recircularización de la cadena de ADN reconstituyendo el plásmido completo (Smillie et al., 2010).

1.3-Integrones

1.3.1-Introducción

Desde el descubrimiento de la penicilina por Alexander Fleming en 1928, los antibióticos han sido la principal herramienta para el tratamiento de enfermedades infecciosas. Sin embargo, su eficacia se ha visto disminuida por la emergencia de la resistencia bacteriana y particularmente el desarrollo de la multi-resistencia en los centros hospitalarios (Loot et al., 2017).

Poco tiempo después de la introducción de la terapia antibiótica, surge a mediados de 1950 el primer brote de multi-resistencia en el ámbito clínico. Sin embargo, recién en la década del 70 se pudo determinar que la multi-resistencia detectada estaba asociada a elementos genéticos transponibles como plásmidos, transposones y secuencias de inserción (Liebert et al., 1999).

Posteriormente, a fines de la década del 80, se pudo determinar que, en muchos casos, los responsables de la diseminación de la multi-resistencia eran estructuras genéticas “nuevas”, a las que se les denominó integrones (Stokes and Hall 1989). Estos integrones estaban embebidos en elementos genéticos móviles, el transposón Tn21 y se dispersaban mediante el plásmido NR1 (R100) (Liebert et al., 1999).

Los integrones son considerados verdaderos sistemas de ingeniería genética que poseen la maquinaria necesaria para la adquisición y expresión de nuevos genes. Mediante eventos de recombinación sitio específica, integran marcos de lectura

abiertos (*open reading frames*, ORFs) y los “convierten en genes funcionales”, asegurando su correcta expresión (Mazel, 2006).

1.3.2-Estructura de un integrón

Todos los integrones comparten una estructura funcional común compuesta por una plataforma estable asociada a una región variable de genes *cassette* (GCs) que codifican para funciones accesorias (Cambray et al., 2010) (Figura 1.4).

La región estable contiene: i) un gen integrasa *intI* que codifica para una tirosin recombinasa sitio específica (IntI). La enzima cataliza la inserción y escisión de los GCs (Stokes & Hall, 1989).

ii) Un sitio de recombinación primario (*attI*) adyacente a *intI* donde se insertan los GCs (Collis et al.,1993).

iii) Un promotor *Pc* que regula la expresión de los genes insertados. *Pc* se encuentra en la región codificante del gen *intI* o entre el gen y el sitio *attI*, orientado hacia el punto de integración (Collis & Hall, 1995);(Jové et al., 2010).

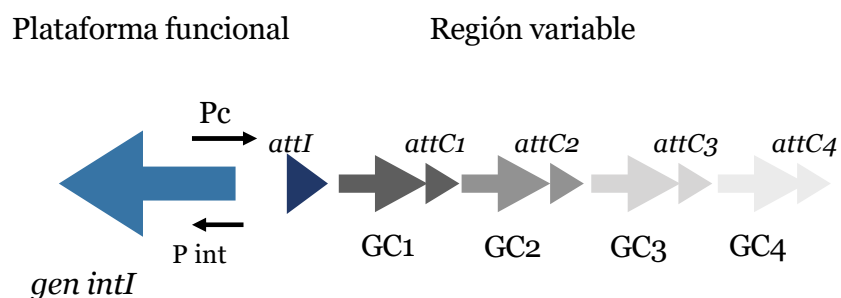


Figura 1.4: Organización del integrón. La plataforma estable del integrón está compuesta por el gen *intI*(azul), el sitio de recombinación primario *attI*(azul oscuro), el promotor *Pc* que dirige la expresión de los GCs y el promotor de *intI*, *Pint*. Las flechas indican la dirección de la transcripción de los ORFs. La región variable la constituye el arreglo de GCs, el nivel de expresión de los GCs está reflejado por la intensidad del color de las flechas. Los GCs próximos al promotor serán los más expresados, los genes más alejados

constituyen el reservorio, que está disponible para ser expresados cuando la célula lo requiera.

1.3.3-Clasificación

Los integrones son estructuras ancestrales y estables que han jugado un rol esencial en la evolución del genoma bacteriano (Rowe-Magnus et al., 2001). Si tomamos como base la presencia del gen *intI*, se encuentran en aproximadamente entre el 9-17% de los genomas disponibles en las bases de datos (Boucher et al., 2007);(Cambray, et al., 2010);(Cury et al., 2016). Están distribuidos a lo largo de todos los fila de bacterias Gram-negativas, excepto en α -proteobacteria, siendo más prevalentes entre γ - proteobacteria (Cury et al., 2016).

Históricamente, los integrones han sido clasificados en dos grupos: los integrones cromosómicos (ICs) que se localizan en el cromosoma bacteriano y los integrones móviles (IMs), que están físicamente asociados a elementos genéticos móviles como transposones y que pueden estar contenidos en plásmidos conjugativos (Mazel, 2006).

Actualmente los ICs se denominan integrones cromosómicos sedentarios (ICSs), debido a que los IMs también pueden encontrarse en el cromosoma y el término puede resultar confuso (Cury et al., 2016).

Los ICSs comparten una historia evolutiva, y su filogenia refleja un patrón de herencia vertical. Por este motivo, estos integrones no son clasificados en clases como los IMs, sino acorde a la especie que los contiene (Cambray et al., 2011).

El estudio filogenético de los respectivos genes *intI*, permite distinguir tres grandes grupos: i) el grupo de Proteobacterias de suelo y agua dulce, que incluye los integrones de clase 1 (*intI1*) y clase 3 (*intI3*); ii) el grupo de las γ -proteobacterias de ambientes marinos, que incluye los integrones clase 2 (*intI2*) e integrones localizados en elementos conjugativos integrativos SXT y en el plásmido pRSV1 de *Vibrio* y iii) los integrones que tienen el gen *intI* en sentido inverso. Este grupo incluye miembros de Cyanobacteria, Planctomycetes, Spirochaetes y Chlorobi, provenientes de ambientes diversos (Boucher et al., 2007);(Cambray, et al., 2010).

El hecho de que los dos primeros grupos formen taxones ecológicamente relevantes sumado a algunas incongruencias observadas en los estudios filogenéticos de las integrasas IntI, refuerza la idea de la existencia de eventos de THG entre bacterias que comparten ambientes similares (Mazel, 2006).

1.3.4-Sitio *attI*

El sitio *attI* contiene dos dominios de unión a la enzima integrasa denominados L y R, que forman el sitio *core* de recombinación. Cuando los GCs son integrados (*attI* x *attC*), el punto de recombinación ocurre en el triplete 5'-GTT-3' del dominio R, entre las bases A y C de la hebra complementaria (Hall et al., 1991). El sitio de unión L es degenerado respecto a R y la región central (entre L y R) es muy variable. Por este motivo, los sitios *attI* son difíciles de identificar entre los distintos integrones. El grado de especificidad existente entre las integrasas y sus sitios *attI* sugiere que éstos han evolucionado de forma simultánea (Rowe-Magnus & Mazel, 2001).

Las secuencias *attI* pueden llegar a ser reconocidas por proteínas de clases diferentes (ver luego descripción de clases). Se ha demostrado experimentalmente que integrasas móviles de clase 1 pueden integrar GCs en sitios *attI2* y *attI3*, de integrones clase 2 y clase 3 respectivamente, aunque la eficiencia de recombinación para estos casos es mucho menor. Sin embargo IntI1 no es capaz de reconocer el sitio *attI* del super-integrón de *Vibrio cholerae* (Biskri et al., 2005).

Los sitios *attI* de los integrones clase 1 tienen dos sitios adicionales de unión a la integrasa, que consisten en repetidos directos imperfectos denominados DR1 y DR2 situados *upstream* del sitio *core*. Asimismo, estas secuencias adicionales no parecen ser frecuentes en los sitios *attI* de integrones de origen ambiental (Elsaied et al., 2011).

1.3.5-Sitios *attC*

El rasgo más distintivo de los sitios *attC* es que presentan una gran diversidad de secuencia pero a su vez forman estructuras secundarias muy conservadas.

Los sitios *attC* poseen dos dominios homólogos invertidos, R''-L'' y L'-R' separados por una región central variable en largo (20-104 pb) y secuencia (Mazel, 2006). Sus extremos presentan secuencias palindrómicas conservadas que les permiten formar estructuras secundarias por apareamiento intracadena (Escudero et al., 2018). (Figura 5).

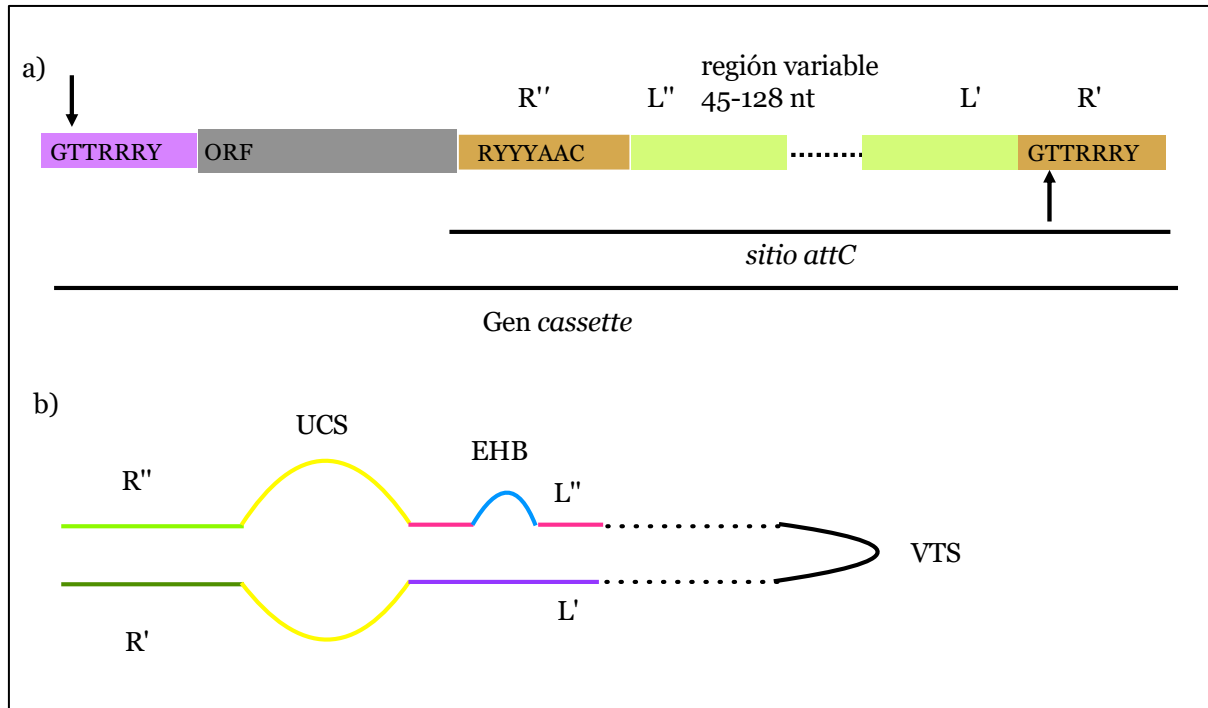


Figura 1.5: Estructura de un GC integrado (forma lineal) y los sitios de recombinación conservados. Las flechas verticales indican el punto de recombinación. El sitio *attC* contiene los dominios de unión a la enzima IntI (R'',R', L'' y L'). Estos sitios tienen secuencias palindrómicas parciales que se aparean y forman una estructura cruciforme estable que es reconocida por la enzima. Entre los dominios L'' y L' se encuentra la región variable que forma una estructura secundaria estable también reconocida por la enzima IntI. b) Esquema de la estructura secundaria de los sitios *attC* durante el plegamiento. UCS (región central impar), VTS (estructura variable terminal), EHB (bases extra-helicoidales).

La región conservada consiste en dos triplete invertidos 5'-AAC-3' y 5'-GTT-3' localizados en los dominios R'' y R' respectivamente, que pueden extenderse 7 pb y se denominan *core* y *core* inverso. El *core* inverso (RYYAAC, donde R es una

purina e Y una pirimidina) se localiza en el extremo 5' del sitio *attC* y el *core* complementario (GTTRRRY) en el extremo 3'. El punto de recombinación se encuentra en el triplete 5'-GTT-3' del dominio R' y el punto de corte ocurre entre las bases A y C de la hebra complementaria (Escudero et al., 2018).

Estos dominios se pliegan formando una estructura cruciforme entre R'' y R' y L'' y L' que contienen tres estructuras particulares: las bases extra-helicoidales (EHBs), el separador central impar (UCS) y la estructura variable terminal (VTS) (Cury et al., 2016).

Los integrones presentan un sistema de recombinación único, debido en parte a las características de los sitios *attC*. La información necesaria para la recombinación no está enteramente contenida en su secuencia primaria sino que también viene determinada por la estructura secundaria. Este rasgo atípico explica porqué diferentes sitios *attC* pueden ser reconocidos por la misma enzima (Bouvier et al., 2009).

1.3.6-La enzima integrasa IntI

La integrasa pertenece a la familia de las tirosin (Y)-recombinasas (Nunes-Düby et al., 1998) pero forma una subfamilia específica cercanamente emparentada a la proteína XerCD, un hetero-tetrámero formado por dos subunidades de la proteína XerC y dos de la proteína XerD (Boyd et al., 2009). Las Y-recombinasas están ampliamente distribuidas entre los procariotas, donde participan en la resolución de multímeros circulares de ADN que se forman durante la replicación, en la integración y escisión de genomas virales y también en la alternancia de la expresión de genes (Escudero et al., 2016).

Las Y-recombinasas reconocen secuencias específicas de ADN doble cadena. La recombinación incluye la formación de un complejo sináptico entre dos moléculas de ADN y cuatro monómeros de la recombinasa. Una cadena de cada sustrato es cortada y transferida a la molécula opuesta formando una estructura en H (modelo de Holliday). Esta estructura es eliminada mediante el corte y transferencia de las cadenas restantes (Grindley et al., 2006) (Figura 1.6).

La reacción de recombinación en los integrones es una excepción a la reacción de recombinación que sigue el modelo de resolución de Holliday. Las integrasas

pueden variar las vías de recombinación en función del tipo de sustrato que procesan. Son sistemas de recombinación híbridos que reconocen secuencias de ADN de doble cadena (sitio *attI*), o de cadena única no conservada, que forma una estructura secundaria específica (sitio *attC*) (Escudero et al., 2016). Esta singularidad se debe a la presencia de un dominio adicional en el sitio catalítico de la enzima que no se encuentra en ninguna otra proteína (Nunes-Düby et al., 1998).

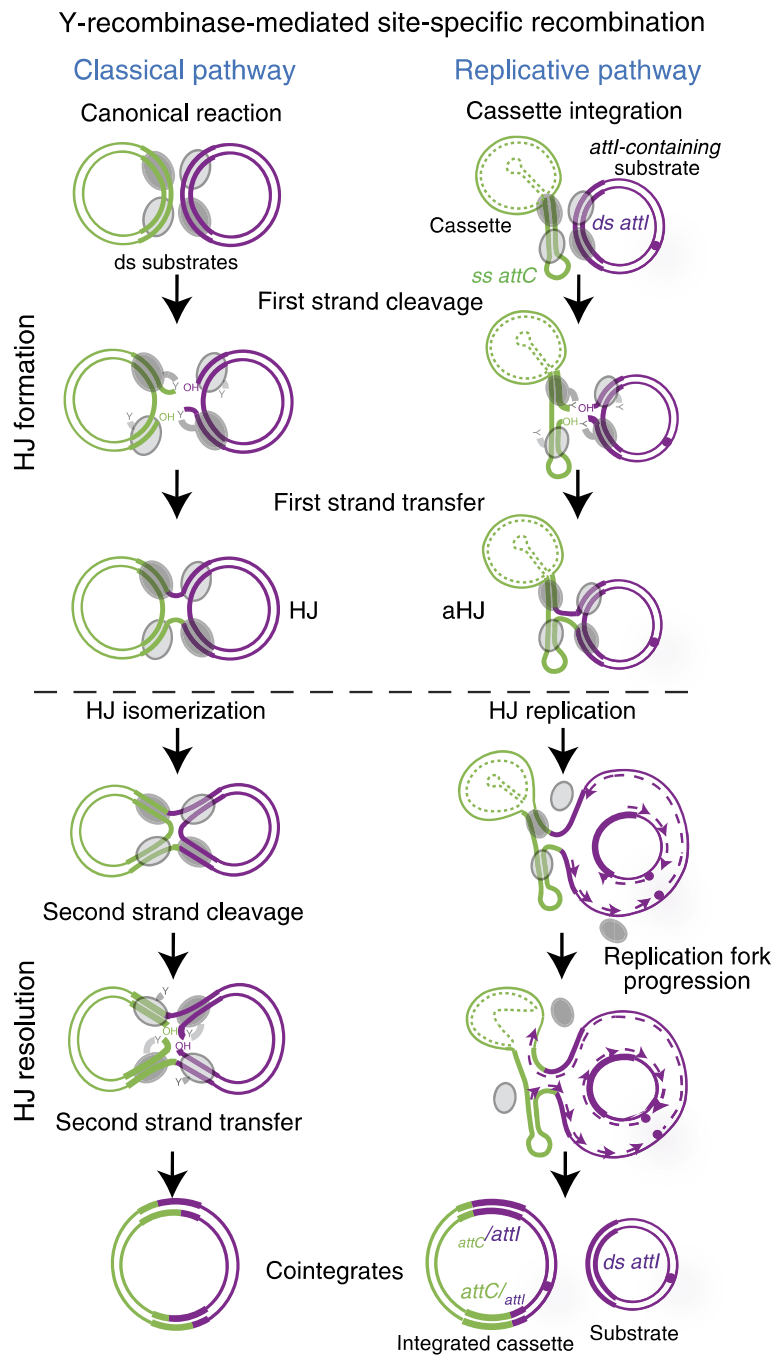


Figura 1.6 : El diagrama muestra los dos modelos de recombinación entre las tirosín-recombinasas. A la derecha se muestra el modelo de recombinación de la mayoría de las Y-recombinasas, a la izquierda el modelo de recombinación exclusivo de las integrón-integrasas. Figura extraída de (Escudero et al., 2016).

Las integrasas presentan el típico plegamiento tridimensional de las Y-recombinasas, los dominios característicos *patch* I, II y III, *box* I y II y los residuos que conforman el sitio activo de la enzima RKHRHY (Grindley et al., 2006). Además poseen un dominio adicional de 19 aminoácidos de largo, único en este tipo de recombinasas, que contiene una α -hélice denominada I2, esencial para la actividad de la enzima (Messier & Roy, 2001). Este dominio participa específicamente en el ensamblaje del complejo que permite acomodar y estabilizar las bases extra helicoidales que sobresalen del sitio *attC* cuando se encuentran en su forma activa (MacDonald et al., 2006). La interacción entre el dominio I2 de la enzima y las bases extra-helicoidales de los sitios *attC* sugiere que hubo un proceso evolutivo en conjunto hacia el sistema de recombinación actual. La singularidad de las IntI radica en la capacidad de reconocer y unirse a moléculas de ADN estructuralmente diferentes. Si bien no se sabe en profundidad las bases estructurales de este mecanismo ambiguo, se ha observado que existe un balance entre el reconocimiento de uno u otro sitio (Escudero et al., 2018).

1.3.7-Reacciones de recombinación de la enzima IntI

Las enzimas integrón integrasas pueden catalizar tres tipos de reacciones con diferentes niveles de eficiencia (Figura 1.7):

i) *attI* x *attC*:

Es la reacción más eficiente y como resultado integra los GCs en el integrón.

La inserción preferencial de los GCs en el sitio *attI* tiene significación biológica desde el punto de vista de la ventaja adaptativa que confiere al hospedero, ya que asegura la expresión de los GCs situados próximos al promotor.

La sucesiva escisión y pérdida de *cassettes* es un proceso dinámico que genera variabilidad en el arreglo de GCs y produce cambios en su patrón de expresión (Collis & Hall, 1995); (Baharoglu et al., 2010).

ii) *attC* x *attC*:

La recombinación entre dos sitios *attC* produce la escisión de los GCs en forma de moléculas circulares covalentemente cerradas (Collis & Hall, 1992).

Estas estructuras pueden contener varios GCs, promoviendo la eliminación o reordenamiento de bloques completos de genes.

iii) *attI* x *attI*:

Se ha propuesto que la recombinación entre dos sitios *attI* representa la reacción ancestral de la actividad de la enzima, dado que es muy similar a la de otras Y-recombinasas (Escudero et al., 2018). Este tipo de reacción puede generar híbridos de sitios *attI* que fusionan integrones diferentes en nuevos arreglos de genes. También puede mover al integrón en distintos sitios dentro del cromosoma mediante la inserción en un sitio de recombinación secundario (Gillings, 2014). Desde un punto de vista evolutivo y considerando el origen cromosómico de los integrones, esta reacción (*attI* x *attI*) es poco relevante. Si los integrones son idénticos no se generarán cambios genotípicos significativos. Sin embargo, sí puede tener implicancias biológicas en el ambiente clínico.

Los integrones móviles de tipo clínico son tan prevalentes que se encuentran en más de una copia. Estos integrones pueden contener arreglos variables bajo el control de distintos promotores Pc. La recombinación entre dos sitios *attI* x *attI* puede resultar en un movimiento en bloque de los GCs de integrones diferentes. De esta forma, un único evento de recombinación puede provocar la expresión de múltiples determinantes de resistencia (Escudero et al., 2018).

positiva tanto en integrones móviles como en integrones cromosómicos (Cambray et al., 2011).

En *E. coli* y en la mayor parte de las β -Proteobacteria y γ -Proteobacteria, los sitios de unión de LexA tienen un motivo palindrómico de 16 pb (CTGTatataatACAG), generalmente denominado caja LexA. El mecanismo de acción involucra una interferencia alostérica de la ARN polimerasa. La respuesta SOS es inducida por la presencia de fragmentos de ADN de cadena simple, que aparecen cuando existe daño celular. También la entrada de ADN foráneo durante los fenómenos de transformación natural o conjugación disparan la respuesta SOS. Estos fragmentos de ADN se unen de forma inespecífica a la proteína RecA promoviendo su polimerización. Esto induce la degradación de LexA por ruptura autocatalítica y se libera la represión de los genes SOS (Escudero, 2015).

La respuesta SOS ha sido asociada a fenotipos clínicamente relevantes como la activación y diseminación de factores de virulencia de bacteriófagos, transposones, islas de patogenicidad y elementos integrativos conjugativos que contienen genes de resistencia a antibióticos (Cambray et al., 2011).

El mecanismo de regulación de expresión de la integrasa estaría íntimamente conectado a la fisiología bacteriana y al ambiente, permitiendo la evolución y adaptación cuando es necesario. Cuando se dispara una señal de alarma, la actividad de la IntI promueve la recombinación y el reordenamiento del arreglo de GCs que favorecen la adaptación y supervivencia de la célula. Del mismo modo, la expresión de la enzima es reprimida cuando la configuración del arreglo de GCs es favorable para la bacteria y un cambio en el orden de los genes podría resultar deletéreo para la célula (Escudero, 2015).

Esta conexión con la respuesta SOS se explica claramente con el mecanismo de captura y diseminación de genes de resistencia antibiótica en integrones móviles. Algunos antibióticos tienen como blanco de acción sistemas relacionados con el ADN bacteriano, las ADN topoisomerasas en el caso de las quinolonas y la síntesis de nucleótidos en el caso de trimetoprim (Escudero, 2015).

Este mecanismo también fue demostrado para antibióticos del grupo de los aminoglucósidos y β -lactamasas que actúan sobre el ribosoma y la pared celular

respectivamente. Se ha propuesto que la generación de especies intermediarias reactivas del oxígeno causan daño en el ADN mediante la oxidación de nucleótidos que finalmente inducen la respuesta SOS, aún en concentraciones sub-inhedorias (Baharoglu et al., 2013).

El ADN simple cadena es a la vez sustrato de la enzima y la señal que dispara la recombinación. Este ADN ingresa a la célula cuando se produce la transformación natural o la conjugación, entonces la conexión con la respuesta SOS estimula la adquisición de nuevos genes con funciones que pueden ser ventajosas para la célula. Asimismo, en ambientes estables, la desregulación de la enzima puede ser riesgosa cuando el orden y contenido de los GCs son óptimos (Gillings, 2014).

El reciente hallazgo de que la expresión de la integrasa clase 2 no está bajo el control del regulón SOS (Jové et al., 2017) y la presencia del codón stop prematuro en los genes de integrones clase 2 demostraría el efecto deletéreo que tiene la enzima cuando no existe un buen sistema de represión. Una explicación posible sería que la IntI2 esté bajo el control de otro mecanismo regulatorio o que hubiera perdido la caja LexA (Escudero et al., 2018).

Un caso interesante que demuestra la conexión entre las condiciones ambientales y la actividad de los integrones es el de la integrasa de *Vibrio cholerae* (IntIA) que está bajo el control de un segundo regulador, a través de la proteína receptora del AMPc (CPR). La caja de unión de CPR se encuentra en la región promotora de IntIA, entre los promotores Pc y Pint. CPR regula la expresión de la enzima independientemente de la respuesta SOS. CPR es uno de los principales reguladores del catabolismo del carbono, adaptando el metabolismo de la célula en función de la fuente de carbono disponible (Escudero et al., 2018).

En conjunto, estas observaciones indican que existe una íntima conexión entre la expresión de la enzima, la fisiología de la célula y el ambiente. Esto indica que los integrones funcionan como elementos que promueven la adaptación del hospedero cuando las condiciones externas lo requieren.

1.3.9- Genes cassette (GCs)

i) Estructura

Los GCs constituyen la parte variable y móvil del integrón. Poseen una estructura simple que consta de una secuencia codificante (CDS) y un sitio de recombinación secundario en su extremo 3', denominado *attC* (anteriormente conocido como elemento de 59 bases), que es reconocido por la integrasa (Recchia & Hall, 1995).

Los sitios *attC* son sitios de recombinación peculiares y diferentes a los sitios clásicos de las Y-recombinasas. Poseen características estructurales únicas que proveen a los integrones de un mecanismo de recombinación distintivo e innovador (Escudero et al., 2018). El tamaño de los *cassettes* varía entre 500 y 1000 pb (Stokes et al., 1997). Pueden estar libres en la célula como elementos móviles no replicativos en forma circular o en forma lineal cuando están insertados en el integrón. Una vez que son integrados se vuelven funcionales y el promotor *Pc* dirige su expresión. La sucesiva integración de los GCs en la plataforma estable del integrón genera un arreglo de genes que se disponen en tándem (Recchia & Hall, 1995).

En los ICSs, el número de GCs puede variar desde cero hasta más de 200, como en el superintegrón de *Vibrio vulnificus*, que posee 217 GCs. El alto grado de identidad de secuencia de los sitios *attC* en *Vibrio cholerae* (>80%) sugiere que estos sitios son especie-específicos. Por el contrario, los GCs de IMs muestran diferencias tanto en la secuencia como en el largo de sus sitios *attC*, así como también inconsistencias en el uso de codones de los ORFs que codifican. El arreglo de GCs más largo encontrado en IMs posee ocho genes (Rowe-Magnus & Mazel, 2001).

ii) Expresión de los GCs

La mayoría de los GCs carecen de promotor propio y su expresión está condicionada a su inserción en el sitio *attI* del integrón y a la proximidad del promotor *Pc*. La distancia de los GCs al promotor afecta su nivel de expresión generando un gradiente que irá disminuyendo hasta un punto en el cual los genes

no serán expresados. El gradiente de expresión dependerá de la fuerza del promotor, de la posición en el arreglo y naturaleza del GC (Collis & Hall, 1995).

La correcta orientación de los GCs respecto al promotor Pc viene determinada por los sitios *attC* (Nivina et al., 2016).

El modelo experimental utilizado para estudiar la expresión de los GCs es el de los integrones de clase 1. En estos integrones, el promotor Pc se localiza en la región codificante de *intI1*, aunque en algunos casos puede haber un segundo promotor Pc2 en el sitio *attI1*. Las variantes de Pc, Pc2 y combinaciones de Pc-Pc2 exhiben diferencias en la fuerza de expresión que producen distintos fenotipos para un mismo arreglo de GCs (Collis & Hall, 1995); (Jové et al., 2010).

Existe una interferencia en la transcripción entre el promotor de los GCs (Pc) y el promotor de la integrasa (Pint). Las variantes más fuertes de Pc disminuyen el nivel de actividad de Pint, incluso aunque Pint esté inducido por la respuesta SOS. Este fenómeno puede resultar ventajoso para el caso de las IntI de IMs que se transfieren a especies que carecen de LexA y la desregulación de la integrasa es costosa. La interferencia en la transcripción entre Pc y Pint implica un balance entre la expresión de los GCs que codifican para funciones que confieren una ventaja adaptativa y la expresión de la integrasa que cataliza la recombinación y permite la captura de nuevos genes. Estudios realizados con variantes de promotores de GCs en aislamientos de cepas clínicas y ambientales de *E.coli* demostraron en ambas la prevalencia de promotores Pc más débiles. Estos resultados sugieren que la balanza se inclina a favor de promover la incorporación de nuevos genes (mediada por la expresión de Pint) más que a la expresión de los GCs existentes (Escudero et al., 2018).

En algunas ocasiones se identificaron GCs con promotor propio en integrones de clase 1. Tal es el caso del gen *cmlA1* y de *qnrVC1* que confiere resistencia a cloramfenicol y a quinolonas respectivamente. Asimismo, en integrones de clase 2 se reportó un gen funcional de resistencia a eritromicina *ere(A)* localizado en el plásmido pIP1100 (Rowe-magnus, et al., 2003) y dos genes del sistema toxina-antitoxina en el super-integrón de *Vibrio cholerae* (Guérout et al., 2013).

La naturaleza del GC influye en el fenotipo de la célula. Algunos GCs tienen un motivo al cual se unen los ribosomas (sitio de unión al ribosoma, RBS). En el caso

de los GC que carecen del RBS (*ribosome binding site*), la traducción comienza corriente arriba y continúa con el siguiente gen como si fuera un operón. El sitio *attI* de integrones clase 1, codifica un ORF (*orf11*) que contiene un RBS funcional (presente en todos los transcriptos generados a partir de *Pc*) y dirige gran parte de la traducción de los GCs que carecen de RBS (Escudero et al., 2018).

Los arreglos de GCs de gran tamaño proveen a la célula de un variado *pool* de genes que son mantenidos a bajo costo. Estos *cassettes*, aunque estén silenciados, permanecerán disponibles para ser expresados cuando la célula lo requiera (Loot et al., 2017).

iii) Diversidad y funciones de los GCs

Los GCs están ampliamente distribuidos en la naturaleza, han sido hallados en todos los ambientes estudiados (suelos, sedimentos marinos y lacustres, biofilms, superficie de plantas y en simbioses de eucariotas) (Stokes et al., 2001);(Elsaied et al. 2007);(Elsaied et al., 2011);(Koenig et al., 2009);(Gillings et al. 2009);(Holmes et al., 2003).

Existen diferencias entre las funciones para las cuales codifican los GCs de IMs y de ICSs. Los IMs contienen casi exclusivamente genes de resistencia a antibióticos mientras que los ICSs portan en su mayoría genes cuya función se desconoce (Rowe-Magnus et al., 2001). En integrones de clase 1 se han descrito más de 130 genes de resistencia a casi todas las familias de antibióticos. Estos incluyen genes que confieren resistencia a todos los aminoglicósidos, β -lactamasas, trimetoprim, cloramfenicol, estreptomicina, rifampicina, fosfomicina, macrólidos, quinolonas y antisépticos (Partridge et al., 2009);(Macdonald et al., 2006).

Estudios en especies del género *Vibrio* revelaron que más del 65% de los GCs codifican para proteínas que no tienen homólogos en las bases de datos. Un 13% de los GCs tienen homólogos cuya función se desconoce y sólo un 20% codifican para proteínas de función conocida (Boucher et al., 2007). Estas funciones están relacionadas con una variedad de procesos celulares, metabolismo y almacenamiento de información. Se han identificado genes que median la interacción con el ambiente externo como la adhesión y formación de biofilms

(señales peptídicas o múltiples dominios transmembrana) y genes de resistencia a bacteriófagos (Rapa et al., 2013).

Los GCs de los CSs presentan como un rasgo distintivo genes del sistema toxina-antitoxina (TA). Este sistema estabiliza segmentos de ADN mediante un proceso conocido como muerte post-segregacional. Está compuesto por una toxina estable y una lábil que neutraliza la primera. La permanencia de este sistema en el genoma es asegurada por la diferencia en la vida media de cada una. Si la célula pierde los genes del sistema TA, la antitoxina se degrada primero y la toxina provoca la muerte celular (Van Melderen & De Bast, 2009). Se cree que los sistemas TA tienen la función de estabilizar los arreglos de GCs de gran tamaño. A diferencia de la mayoría de los GCs, los sistemas TA contienen su propio promotor que les permite funcionar independientemente de su posición en el arreglo y por lo tanto de su distancia al promotor *Pc*.

Los sistemas TA se clasifican en tres tipos según la naturaleza y modo de acción de la anti-toxina. En los tipos I y III, las antitoxinas son pequeños ARNs que impiden la traducción de la toxina. En el tipo II, la antitoxina es una proteína que inhibe la toxina mediante la formación de un complejo proteico (Van Melderen & De Bast, 2009).

En el super-integrón de *V. cholerae* N16961, algunos de los GCs del sistema TA tipo II son integrados en la orientación opuesta, asegurando su expresión mediante su propio promotor (Iqbal et al., 2015).

iv) ¿Cómo se originan los GCs?

Esta es la pregunta para la cual aún no hay respuestas. Se han planteado varias hipótesis sobre el mecanismo de creación de *novo* de GCs. Si bien se conoce la estructura de los sitios *attC*, el plegamiento y su conexión con los procesos celulares del hospedero, no existe un modelo que demuestre el vínculo inicial de un gen con el sitio *attC* (Escudero et al., 2018).

Uno de los modelos propone que los GCs tendrían origen en los ARNs mediante la participación de intrones del grupo IIC-*attC* (Léon & Roy, 2009). Los intrones del grupo II son ARN catalíticos capaces de escindirse de su molécula precursora. Constan de dos componentes: un ARN denominado ribozima y una proteína con

actividad madurasa y reverso transcriptasa. Los intrones del grupo II muestran afinidad por secuencias palindrómicas incluyendo los sitios *attC* (Léon & Roy, 2009).

El mecanismo propuesto es que dos intrones IIC-*attC* de grupos idénticos se integran al sitio *attC* y al transcripto terminal del gen. Mediante recombinación homóloga se fusionan y se forma un intermediario entre el gen, el intrón y el sitio *attC*. Este intermediario es transcripto a ARN, el intrón se escinde y retro-transcribe el gen-*attC*-ARN mensajero formando un GC (Escudero et al., 2018).

Sin embargo, existen argumentos en contra de esta posición que sostienen que sería poco probable que un sistema que ha evolucionado durante millones de años junto al genoma del hospedero “dejara en manos” de un elemento independiente una función tan importante como la creación de GCs. Si los GCs tuvieran su origen en intrones, sería lógico que éstos formen parte de la plataforma estable del integrón, principalmente en el caso de los super-integrones, donde el alto grado de identidad entre sus sitios *attC* sugiere que existe un nexo entre la generación de los GCs y el hospedero. Por otra parte, tampoco se explica el mecanismo por el cual se generan los GCs que poseen promotor propio como los genes del sistema toxina-antitoxina (Escudero et al., 2018).

Esta teoría tampoco podría explicarse en el caso del super-integrón de *Vibrio cholerae* que posee 179 GCs y donde no se ha reportado la presencia de intrones del grupo IIC-*attC* (Escudero et al., 2018). Asimismo, Cury y col., 2016, confirmaron mediante el modelo de covarianza, que existe una correlación positiva entre la homología de los sitios *attC* y el número de GCs en el integrón: a mayor número de GCs habría mayor homología entre sus sitios *attC*. La homología que presentan los sitios *attC* dentro de un mismo integrón sugiere que los GCs son generados por la misma enzima integrasa que forma parte del integrón (Cury et al., 2016).

1.3.10-Integrones cromosómicos sedentarios (ICSs)

Los ICSs se encuentran en el genoma de γ -proteobacterias, Vibrionaceae y sus parientes más cercanos, la familia Xanthomonadaceae y algunos miembros de la familia Pseudomonadaceae. También se describieron en Chlorobi, Cyanobacteria, Spirochaetes y Planctomycetes (Mazel, 2006)(Gillings, 2014).

Los ICS han sido extensamente estudiados en el género *Vibrio*. Todas las especies conocidas hasta el momento contienen integrones en su cromosoma. Poseen como característica distintiva cientos de GCs (36-219) en su región variable, que representa entre 0,7-3% del genoma y por este motivo también se denominan super-integrones (SI). En *Vibrio parahemolyticus* se identificaron 72 GCs mientras que en *Vibrio vulnificus* más de 200 (Mazel, 2006).

Una característica particular de los super-integrones es que las secuencias de sus sitios *attC* difieren menos del 10% entre sí. El alto grado de identidad de secuencia que comparten sugiere que los GCs son ensamblados mediante su asociación con el sitio *attC* de un SI y un fragmento de ADN entrante (Mazel, 2006).

El SI mejor estudiado es el de *Vibrio cholerae* que se describió por primera vez en 1998 por el grupo de Mazel y col., (1998). Investigando la relación entre los GCs de integrones móviles y un *cluster* de secuencias repetidas VCRs (*Vibrio cholerae repeats*), observaron que tenían una organización genética similar a la del cromosoma 2 de *V. cholerae*. Esta observación llevó a la identificación de una nueva enzima integrasa denominada VchIntIA (Mazel et al., 1998).

Este integrón tiene dos características que lo diferencian de los integrones móviles: i) la región variable contiene un gran número de GCs y sus sitios *attC* presentan más de un 80% de identidad entre sí; ii) la plataforma localizada en el cromosoma no está asociada a ningún elemento móvil (Mazel et al., 1998)

El SI de *V. cholerae* representa el 3% del genoma de la bacteria y contiene 172 GCs. Al igual que los GCs del resto de las especies de *Vibrio*, la mayoría codifican para funciones desconocidas. Los pocos GCs que han sido caracterizados o cuya función ha sido inferida por homología con otras proteínas, tienen una diversidad de funciones que refleja la plasticidad genómica y fenotípica que los integrones le confieren al hospedero (Boucher et al., 2007).

En especies de *Xanthomonas*, la región variable contiene entre 1-22 GCs. En cuanto a las funciones que cumplen, los GCs son específicos de cada cepa porque cada patovar se especializa como patógeno de una especie de planta en particular (Gillings, 2014).

Los ICS también están presentes en el género *Pseudomonas*, distribuidos entre distintas especies, indicando múltiples eventos de THG. El número de *cassettes* varía en número (10-32) y contenido (Gillings, 2014).

También se identificaron integrones en el genoma de otras especies que no comparten los rasgos típicos de los ICSs. En *Nitrosomonas europaea*, el integrón no posee GCs y en *Shewanella*, el ICS contiene pocos GCs con sitios *attC* estructuralmente diferentes, más parecido a los IMs (Mazel, 2006).

En el filo *Spirochaetes* y particularmente en el género *Treponema* existen integrones que tienen una organización particular donde el gen *intIA* está orientado en sentido inverso al habitual. En estos integrones, el sitio *attI* se ubica en el extremo 3' y la transcripción de *intIA* se realiza en el mismo sentido que los GCs. En *Treponema denticola*, la región variable contiene un arreglo con 45 GCs con sitios *attC* muy similares entre sí (Mazel, 2006).

También se han encontrado ICSs separados en fragmentos y localizados en cromosomas distintos como en las cepas de *Photobacterium profundum* SS9 y *V. fischeri* ES114 (Mazel, 2006).

1.3.11-Integrones móviles (IM)

Como ya fue mencionado anteriormente, los integrones fueron inicialmente identificados en aislamientos clínicos resistentes a múltiples antibióticos. Debido a que estaban localizados en elementos transponibles, se pensó que eran móviles en sí mismos. Posteriormente se determinó que estos integrones surgieron a partir de la asociación de integrones cromosómicos con transposones y plásmidos conjugativos. Esta asociación facilitó su diseminación entre bacterias patógenas intra o inter-especie. A su vez, la presión selectiva impuesta por el uso de antibióticos en humanos, animales de producción y el ambiente, favoreció la selección de este tipo de integrones (Cambray, et al., 2010)(Gillings 2014).

Los IMs se agrupan en cinco clases cuyas IntI exhiben entre 40-58% de identidad entre sí. Todos ellos están asociados a elementos móviles como transposones, plásmidos conjugativos y secuencias de inserción (ISs) (Mazel, 2006). Los integrones de clase 1 y clase 3 están asociados al transposón Tn402 (Collis, et al., 2002);(Xu et al., 2007), mientras que los integrones de clase 2 se asocian exclusivamente con el transposón Tn7 y sus derivados (Ramírez et al., 2010). Si bien las cinco clases están asociadas a determinantes de resistencia a antibióticos, solamente las clases 1, 2 y 3 han estado históricamente involucradas en la diseminación de fenotipos multi-resistentes y por lo tanto son consideradas las más relevantes para la salud humana (Mazel, 2006).

Los IMs son los principales responsables del ensamblaje y diseminación de genes de resistencia a antibióticos en bacterias Gram-negativas (Fluit & Schmitz, 2004). Se encuentran ampliamente distribuidos en los géneros *Pseudomonas*, *Vibrio* y *Acinetobacter* y en la familia Enterobacteriaceae (Mazel, 2006). Generalmente presentan *cassettes* de resistencia a antibióticos agrupados en arreglos relativamente pequeños (Mazel, 2006). Al presente se han identificado más de 130 GCs que confieren resistencia a casi todas las familias de antibióticos, excepto a tetraciclina (Rowe-Magnus et al.,2002); (Fluit & Schmitz, 2004b); (Partridge et al., 2009).

La clase 1 es la más extendida y mejor caracterizada por ser la más prevalente en bacterias Gram-negativas de origen clínico (22-59%) (Labbate., et al., 2009). Por esta razón representan el modelo experimental de integrones por excelencia (Escudero, 2015).

La estructura típica de un integrón clase 1 (Figura 1.8) consta de dos segmentos conservados que flanquean una región central (región variable), donde se insertan los *cassettes*. El extremo 5' conservado (5'CS) contiene el gen *intI1* y el sitio de recombinación primario *attI1*. Como ya fue mencionado, dentro del segmento conservado se encuentran uno o más promotores que regulan la expresión de los *cassettes*. El extremo 3' (3'CS) no siempre está presente en los integrones clase 1. Esta región incluye el gen *qacEΔ1* (una forma truncada del gen *qacE*) que confiere resistencia a compuestos de amonio cuaternario, el gen *sul1* que confiere resistencia a sulfonamidas y *ORF5* de función desconocida (Stokes et al., 2006).

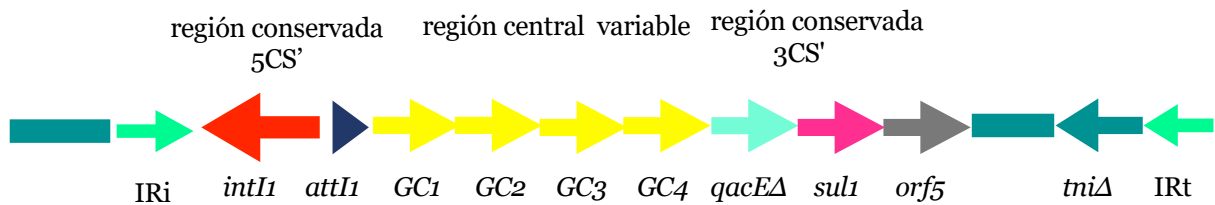


Figura 1.8: Integrón de clase 1 de tipo clínico. Consta de dos regiones conservadas 5CS' y 3CS'. El segmento 5CS' contiene el gen *intI1* y el sitio *attI1*. El segmento 3'Cs contiene los genes *qacEA* y *sul1* que confieren resistencia a desinfectantes y sulfonamidas respectivamente, el *orf5* de función desconocida y los fragmentos del gen que codifica para funciones de transposición (*tniΔ*). La región variable porta los GCs de resistencia a antibióticos. *IRi* e *IRt* (invertidos repetidos directos).

Los integrones de clase 1 pueden estar inmersos en transposones defectivos o funcionales. La mayor parte están contenidos en el transposón *Tn402*, probablemente uno de los pocos transposones activos, que a su vez está contenido en un transposón más grande, el transposón *Tn21* (Brown et al., 1996). *Tn402* contiene el integrón y un módulo de transposición completo flanqueado por las secuencias repetidas invertidas *IRi* e *IRt*. Se especula que tras varios eventos de inserción y deleción, este transposón dio origen a los integrones clase 1 “clínicos” (Gillings et al., 2008).

Los integrones de clase 2 y clase 3 también portan genes de resistencia a antibióticos, sin embargo la diversidad de estos genes es mucho menor comparada con los integrones de clase 1.

Los integrones clase 2 están asociados exclusivamente con el transposón *Tn7* y sus derivados. La mayoría de los integrones clase 2 presentan una mutación que resulta en un codón *stop* prematuro en el aminoácido 179, que da como resultado una proteína no funcional (Barlow et al., 2006). El hecho de que todos los integrones clase 2 clínicos tengan la misma mutación sugiere que se habrían originado de un ancestro común (Gillings, 2014).

Los integrones clase 2 generalmente contienen el mismo arreglo de *cassettes* con tres genes, *dfra1*, *sta2* y *aadA1* (Barlow et al., 2006). Sin embargo, se encontraron nuevos arreglos de genes de resistencia a antibióticos que hasta el momento no

estaban descritos en esta clase de integrones (Biskri et al., 2003) (Ramírez et al., 2005). No está claro si estas variantes se generaron por la supresión momentánea de la mutación del codón *stop* o por la recombinación mediada por la IntI1 actuando en *trans*, ya que esta enzima puede reconocer y recombinar sitios *attI2* (Mazel, 2006).

Algunos trabajos también dan cuenta de integrones clase 2 con integrasas funcionales, en cepas de *E.coli* (Marquez et al., 2008) y *Providencia stuartii* (Barlow et al., 2006). En este último caso los GCs codifican para funciones desconocidas. Asimismo, estudios independientes del cultivo en ambientes con diverso grado de impacto antropogénico, dan cuenta de integrasas ambientales clase 2 posiblemente funcionales (Rodríguez-Minguela et al., 2009).

Los integrones clase 3 fueron inicialmente identificados en aislamientos clínicos asociados a GCs de resistencia a antibióticos (Arakawa et al., 1995). También se encontraron integrasas clase 3 no “clínicas” en el cromosoma de dos especies de *Delfia* asociadas a GCs que codificaban para proteínas de función desconocida (Xu et al., 2007).

Los integrones clase 3, al igual que los de clase 1 y clase 2, tienen su origen en bacterias ambientales. Su organización genética sugiere que fueron capturados a partir de un transposón Tn402 pero con la particularidad de que el gen *intI3* está orientado en dirección inversa a *intI1* (Collis et al., 2002).

Los integrones de clase 4 y clase 5 están involucrados en la resistencia a trimetoprim en especies de *Vibrio*. Los integrones clase 4 forman parte de un elemento integrativo y conjugativo SXT, encontrado en *Vibrio cholerae* (Hochhut et al., 2001). La clase 5 está embebida en un transposón compuesto, insertado en un plásmido de *Alivibrio salmonicida* (Sorum et al., 1992).

Los IMs no sólo contienen genes de resistencia a drogas, se identificaron integrones clase 1 y clase 3 no “clínicos”, asociados a GCs de función desconocida (Cambray et al. 2011). Su función en ambientes naturales todavía no está bien comprendida. Desde el inicio de su caracterización y hasta el presente se continúa estudiando y tratando de comprender en profundidad su origen y evolución.

La prevalencia de las distintas clases de integrones en aislamientos clínicos se debe a diferencias en la penetración, persistencia y promiscuidad de las plataformas a las cuales están asociados. También las diferencias en las características de los integrones influyen en el costo energético y la permanencia

en el hospedero, en la actividad de la integrasa y en las diferencias en el reconocimiento de los sitios *attC*. Las integrasas de clase 1 son las que tienen un rango de sustrato más amplio, quizás por eso sea la clase más extendida (Escudero et al., 2018).

Anteriormente se creía que los integrones móviles eran mantenidos en el genoma por la presión antibiótica, en la actualidad existe una visión más integradora que relaciona la estabilidad de los integrones con la fisiología del hospedero. Esta estabilidad se ve favorecida por las ventajas adaptativas que confiere el integrón en situaciones de estrés sin representar demasiado costo energético para la célula (Escudero et al., 2018).

1.3.12-Origen de los integrones clase 1 clínicos

En Betaproteobacterias no patógenas se identificaron integrones cromosomales de clase 1. Estos incluyen miembros de los géneros *Hydrogenophaga*, *Aquabacterium*, *Acidovorax*, *Imtechium*, *Azoarcus*, y *Thauera*. Estos integrones exhiben diversidad de secuencia del gen *intI1* y contienen GCs que codifican para funciones desconocidas. Asimismo, todos los integrones cromosomales clase 1 ambientales tienen en común secuencias conservadas en ambos extremos con puntos de corte específicos. Estas secuencias se han observado en integrones cromosómicos clase 1 de un amplio rango de especies, sugiriendo que estos integrones pueden moverse, aunque se desconoce el mecanismo involucrado (Gillings, 2014).

Los integrones clase 1 de bacterias ambientales tienen acceso a un vasto *pool* de GCs que les permite intercambiar genes con integrones de otras clases. Muchos de estos genes codifican para bombas de eflujo que confieren resistencia a compuestos tóxicos que se encuentran en los ecosistemas naturales. Algunas de ellas pertenecen a la familia de genes *qac* que codifican para la resistencia a compuestos de amonio cuaternario (Gillings et al., 2009).

La presión ejercida por el hombre con el uso de antibióticos en salud humana, en la agricultura y en animales de producción favoreció el reclutamiento de determinantes de resistencia en bacterias presentes en el ambiente.

La mayoría de las clases de integrones que se encuentran en bacterias patógenas resistentes a antibióticos comparten una historia evolutiva reciente.

Particularmente, los integrones clase 1 clínicos comparten ciertas características en común. Todos están localizados en elementos móviles como plásmidos conjugativos, transposones o secuencias de inserción y la región variable contiene pocos GCs que codifican para determinantes de resistencia a antibióticos (Mazel, 2006). Asimismo todos los genes *intI1* clínicos son idénticos entre sí a nivel de secuencia, a diferencia de los genes *intI1* ambientales, que exhiben menos del 95% de identidad nucleotídica. Se considera que los integrones clínicos actuales representan una de las variantes de los genes *intI1* ambientales (Gillings et al., 2008).

Se ha propuesto que los integrones móviles de clase 1 surgieron a partir de un ICS contenido en una Betaproteobacteria no patógena de agua dulce (Gillings et al., 2008). El integrón localizado en el cromosoma fue capturado por un transposón Tn402. Este integrón con sus GCs formó un híbrido junto con Tn402 que contenía además los extremos repetidos invertidos y la maquinaria necesaria para la transposición (Gillings et al., 2008). La mayoría de los integrones clase 1 clínicos actuales no poseen las funciones de transposición de Tn402, a diferencia de integrones clase 1 ambientales que todavía conservan la maquinaria de transposición.

La diseminación de los integrones clase 1 asociados a Tn402 se vio favorecida porque Tn402 puede reconocer y recombinar sitios *res* que están presentes en plásmidos. Por ejemplo, el transposón Tn21, (ampliamente distribuido entre plásmidos), surgió de la asociación de Tn402 con un elemento móvil que contenía el operón *mer*, que codifica para la resistencia a mercurio (Gillings, 2014).

Los compuestos de amonio cuaternario comenzaron a ser empleados como desinfectantes en la era pre-antibiótica y posiblemente por eso el gen *qacE* haya sido el primer gen en ser reclutado en la plataforma de integrones clínicos (Gillings, 2014). El fenómeno de la resistencia antibiótica comenzó a mediados de 1930 con la introducción de las sulfonamidas. Por lo tanto, no es de extrañar que la resistencia a sulfonamidas haya sido el siguiente evento en ser fijado durante la evolución de los integrones clínicos (Gillings, 2014).

El gen *sulI* codifica para la enzima dihidropteroato sintasa. Su inserción en el integrón provocó la delección de parte del gen *qacE* y su sitio *attC*, generando el extremo conservado 3'CS, característico de los integrones clase 1 clínicos.

Posteriores deleciones provocaron la pérdida de los módulos de transposición de Tn402 (Gillings et al., 2008).

La unión del ICS con Tn402 facilitó así su dispersión entre bacterias patógenas (Cambray, et al., 2010). Su rápida evolución permitió la adaptación a nuevos ambientes y el surgimiento de la multi-resistencia antibiótica entre bacterias de relevancia clínica (Escudero et al., 2015).

1.3.13- Origen viral de los integrones?

Escudero y col., (2018) han especulado que los integrones podrían tener un origen viral. Por un lado proponen que los sitios *attC* y el dominio I2 de las integrasas representan un claro ejemplo de co-evolución que cambió el sistema de recombinación ancestral al sistema actual de un híbrido de ADN cadena doble-cadena simple. El dominio I2 conserva su función original pero a su vez funciona como la adaptación estructural que permite el reconocimiento de nuevos sustratos de ADN de cadena simple.

Por otra parte, los fagos comparten algunas similitudes con los integrones en cuanto a la conexión con la fisiología del hospedero. Al igual que la actividad de la integrasa, el nexa con la respuesta SOS está ampliamente extendida en los fagos.

Como ejemplo citan al fago CTX ϕ de *V. cholerae*. El genoma del fago es integrado como ADN de cadena simple y adopta una forma en horquilla similar a la adoptada por los sitios *attC*. Si bien el fago no codifica para una integrasa, secuestra la recombinasa XerCD del hospedero. Si bien las proteínas XerCD, carecen del dominio I2 y el fago no posee bases extra-helicoidales como los sitios *attC*, existen algunas características en común que alimenta la especulación de un posible origen viral de los integrones. La similitud en el mecanismo de recombinación de los *cassettes* y el fago CTX ϕ , la relación filogenética entre las XerCD y las integrasas así como el hecho de que co-existen en algunas especies bacterianas, como *Vibrio cholerae*, son algunas de ellas.

CAPÍTULO 2

2-Estrategias dependientes de cultivo

- I) Integrones en una colección de aislamientos heterótrofos obtenidos en medio de cultivo para oligotrofos (R2A).**

- II) Elementos móviles que portan integrones clase 1 en enterobacterias antárticas.**

2.1- Antecedentes e Introducción

Desde hace varios años trabajo en la línea de investigación que se desarrolla en el grupo Microbiología Molecular del Depto. BIOGEM (IIBCE), orientada al estudio de elementos de transferencia horizontal de genes (THG) en bacterias del ambiente antártico. En el marco de esa línea de trabajo, participé en el proyecto dirigido a la identificación y caracterización de integrones en bacterias aisladas en distintos sitios de la Península Fildes (Maestría). Los sitios analizados incluyeron zonas expuestas a influencia antropogénica (próximas a las bases de operaciones), zonas de asentamiento de aves y mamíferos marinos y otras zonas que forman parte de áreas especialmente protegidas, denominadas zonas ZAEP (Zonas Antárticas Especialmente Protegidas, Antarctic Special Protected Area) donde se necesita un permiso especial para tener acceso.

En este estudio obtuve una colección de aislamientos en medio de cultivo rico LB (Luria-Bertani) (Sambrook et al., 1989), que crecieron a diferentes temperaturas (5°C, 25°C y 30°C). Luego seleccioné los aislamientos que exhibieran resistencia a algún tipo de antibiótico. Obtuve así una colección de bacterias heterótrofas psicrotolerantes resistentes al menos a alguno de los siguientes antibióticos: ampicilina (Amp), kanamicina (Km), gentamicina (Gm), estreptomicina (Strp), espectinomicina (Spc), trimetoprim (Trim), tetraciclina (Tc), cloramfenicol (Cm) y ácido nalidíxico (Nal).

A partir de esta colección y empleando un par de cebadores dirigidos a amplificar parte del gen integrasa clase 1 (*intI1*) (Mazel et al., 2000), se seleccionaron mediante PCR los aislamientos que generaran productos de amplificación del tamaño esperado. Mediante este protocolo, se obtuvieron cinco aislamientos psicrotrofos resistentes a Trim, Amp y Strp y positivos para la amplificación del fragmento *intI1* (GC4, IA12, CN11, CDTR5 y HTP19). El análisis de secuencia del fragmento *intI1* confirmó la presencia de genes integrasa clase 1 en los aislamientos analizados. La identidad de los aislamientos, establecida mediante análisis de secuencia parcial del gen ARNr 16S, indicó que los mismos pertenecían al género *Enterobacter* sp. (Antelo et al. 2018) (Ver Tabla 2.1).

Table 1. Description of *TrpA* isolates from which a fragment of the expected size was amplified by PCR using primers int_14-int_15¹.

Isolate	Sampling location (sample)	NR RNA	int ¹	int ²	Antibiotic resistance						
					Hydrolytic proteinase ³	Amp ³	Trp ³	Str ³	Cm ³	Nd ³	Spc ³
CD73	63°17'N, 31°53'W (December 200)	Enterovirus	+	+	+	R	R	R	S	S	S
	(strain sent to desalination)										
EP19	63°13'N, 31°53'W (March 200)	Enterovirus	+	+	+	R	R	R	R	S	S
	(brown microbe not with hydrolytic located next to Hall Three Point)										
1412	63°12'N, 31°54'W (March 200)	Enterovirus	+	+	+	R	R	R	R	S	S
	(col with region lines from Halley Island)										
GC4	63°16'N, 31°53'W (December 200)	Enterovirus	+	+	+	R	R	R	R	S	S
	(strain sent to Orlund Glacier)										
CN1	63°17'N, 31°54'W (December 200)	Enterovirus	+	+	+	R	R	R	R	S	S
	(isolated from Nenu Bay)										

R = resistant, S = sensitive

³Gene detected by DNA sequence analysis

Amp = ampicillin, Cm = chloramphenicol, Nd = nalidixic acid, Spc = spectinomycin, Str = streptomycin, Trp = trimethoprim

Tabla 2.1: Descripción de las enterobacterias psicrotolerantes Trim^R que contienen el integrón de clase 1. Tomado de (Antelo et al. 2018).

El estudio preliminar del integrón contenido en estos aislamientos se efectuó mediante análisis de secuencia nucleotídica de algunas regiones, diseñando los cebadores correspondientes. Este análisis indicó que los mismos presentaban un perfil idéntico entre ellos y muy similar al del integrón contenido en el plásmido pKOX105 de *Klebsiella oxytoca* (Carattoli et al., 2010) y p12 de *Klebsiella pneumoniae* (Gootz et al., 2009), alojados en cepas de origen clínico. Próximo al sitio *attI* del integrón, se encontró el gen *cassette dfrA14* conservado, que codifica para la enzima dihidrofolato reductasa y que confiere resistencia a trimetoprim (ver Figura 2.1).

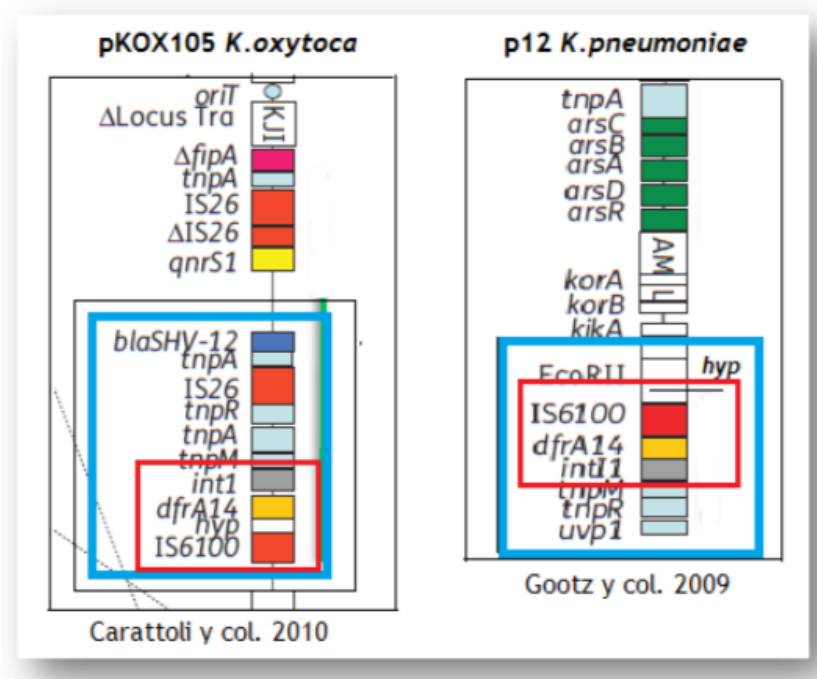


Figura 2.1: Región parcial del mapa físico de los plásmidos pKOX105 y p12. Dentro del recuadro rojo se indican: el gen integrasa clase 1 (*intI1*), el gen *cassette dfrA14*, la proteína hipotética (*hyp*) y la secuencia de inserción *IS6100*. Adaptado de (Carattoli et al., 2010)

Clásicamente, los integrones de clase 1 suelen estar asociados a elementos móviles como plásmidos. Además, la similitud de secuencia de los integrones de nuestros aislamientos con los de pKOX105 y p12, sugería que los mismos podrían estar alojados a nivel plasmídico. El análisis de los perfiles electroforéticos de las miniprep plasmídicas preparadas a partir de los aislamientos *intI1*⁺ indicaron que estas enterobacterias contenían plásmidos de gran tamaño.

Es así que, en el marco de la tesis de Doctorado, nos planteamos analizar en detalle la ubicación de los integrones identificados en el genoma de los aislamientos obtenidos. Mediante ensayos de conjugación usando *Escherichia coli* DH5α como organismo recipiente, logramos establecer que los integrones están contenidos a nivel plasmídico, dado que los transconjugantes adquirieron resistencia a Trim. Además, los transconjugantes exhibieron el fenotipo Amp^r Strp^r. El trabajo se complementó entonces con el análisis de secuencia de tres plásmidos.

Los plásmidos auto-transmisibles y movilizables actúan como vectores que median el intercambio de genes entre bacterias de distintas especies (Frost et al., 2005). En particular, los plásmidos pertenecientes al grupo de incompatibilidad IncN, como pKOX105 y p12, son considerados de relevancia para la salud pública, dado que facilitan la diseminación de una gran variedad de determinantes de resistencia mediante la incorporación de transposones, secuencias de inserción y GCs (Eikmeyer et al., 2012). pKOX105 y p12 poseen regiones IncN muy conservadas y un 100% de identidad a nivel de aminoácidos entre las proteínas que componen su genoma *core* (Carattoli et al., 2010). Las regiones involucradas en la replicación, transferencia horizontal, funciones de mantenimiento y estabilidad forman parte del genoma *core*, cuyo tamaño es de 35 kb aproximadamente para estos dos plásmidos. En la región variable se encuentran genes de resistencia, así como secuencias de inserción y transposones. Las enzimas Dhfr14, *IntI1* y la proteína hipotética son idénticas (100%). Esta analogía sugiere que tanto pKOX105 como p12 se habrían originado a partir de un ancestro común que posteriormente, mediante eventos de adquisición, eliminación e inserción, fueron generando las distintas variantes que existen actualmente (Carattoli et al., 2010).

Considerando estos antecedentes y tomando en cuenta que los integrones identificados en este estudio provienen de una región aislada geográficamente y

distante a centros hospitalarios, resulta de particular interés profundizar en el análisis y caracterización de los plásmidos que portan el integrón clase 1 con el fin de determinar el posible origen y evolución de estos elementos genéticos.

Objetivo General

Identificar y analizar elementos de transferencia horizontal genética (THG) en comunidades microbianas que se desarrollan en distintos sitios de la Antártida marítima y continental.

Objetivo Específico

- 1) Establecer la presencia de integrones en una colección de aislamientos obtenidos en medio de cultivo para oligotrofos (R2A) empleando cebadores dirigidos a amplificar genes integrasa de clase 1, 2 y 3 (Mazel et al., 2000).
- 2) Analizar el contexto genómico de los integrones clase 1 identificados previamente en el marco de la Maestría. En caso de estar ubicados a nivel plasmídico, analizar la secuencia completa de los mismos en algunos aislamientos.

2-Integrones en una colección de aislamientos heterótrofos obtenidos en medio de cultivo para oligotrofos (R2A).

2.2-Materiales y Métodos

2.2.1-Procesamiento de muestras y aislamientos de microorganismos heterótrofos para la búsqueda de integrones clase 1 (*intI1*), clase 2 (*intI2*) y clase 3 (*intI3*).

Se tomaron muestras de sedimento, tapetes microbianos, agua de lagos y cañadas y agua de mar, en distintos sitios de la Península Fildes (Isla Rey Jorge, Antártida marítima) durante las campañas estivales organizadas por el Instituto Antártico Uruguayo entre los años 2008 y 2017. Las características de cada muestra y coordenadas geográficas se describen en la Tabla S1 (Material Suplementario).

Se prepararon suspensiones a partir de cada una de las muestras (sedimento/suelo y tapetes microbianos), mezclando aproximadamente 1 gramo de muestra con 1 mL de solución salina PBS (*Buffer* fosfato salino) 1X.

En el caso de las muestras de agua, se filtraron previamente a través de membranas de nitrato de celulosa (50 mm de diámetro) de 0,45 µm de tamaño de poro (Whatman[®], Reino Unido). Luego, las membranas se sumergieron en matraces de 250 mL en medio R2A (50 mL) y en medio marino (50 mL) en el caso de agua de mar (ver material suplementario para composición del medio). Se incubaron con agitación (200 rpm) a 10°C, entre 5 y 7 días.

Se realizaron diluciones seriadas (1/10) en PBS 1X a partir de cada suspensión y se sembraron 200 µl de cada una por rastrillado en placas de Petri conteniendo medio R2A sólido (Reasoner et al., 1985). Las placas fueron incubadas a 5°C y 25°C. Se seleccionaron las colonias formadas de modo de generar una colección de aislamientos que intentara representar todas las morfologías presentes. Las suspensiones de células fueron almacenadas a -80°C en el medio correspondiente y glicerol 20%.

2.2.2-Búsqueda de genes *intI1*, *intI2* e *intI3* en organismos cultivables.

Se prepararon lisados con grupos de tres y cinco clones de los aislamientos seleccionados para el análisis de la presencia de genes integrasa de integrones. Estos lisados se emplearon como moldes en reacciones de PCR utilizando cebadores para amplificar parte del gen integrasa de clase 1 (*intI1*), clase 2 (*intI2*), clase 3 (*intI3*) (Mazel et al., 2000). En la Tabla S5 (Material suplementario) se resumen los aislamientos utilizados.

2.3-Resultados

2.3.1- Búsqueda de integrones clase 1, 2 y 3:

Mediante reacciones de PCR se analizaron 206 aislamientos provenientes de 29 muestras diferentes (Tabla S5), empleando cebadores para amplificar parte del gen *intI* clase 1, 2 y 3 (Mazel et al., 2000).

Se obtuvo un fragmento del tamaño esperado (500 pb) para el gen integrasa clase 1 (*intI1*) en un aislamiento cultivado en medio R2A a 5°C (LU22_37) proveniente de una muestra de agua dulce (LU6). El análisis de secuencia nucleotídica (Blastn) del gen *intI* de LU22_37 exhibió 99% de identidad con la secuencia del gen integrasa clase 1 (*intI1*), contenido en el plásmido (posición 162041-162483) pF214D_HI2-274.883 pb de la cepa F2_14D de *E. coli* (MK461931.1). La cepa F2_14D de *E. coli* fue aislada del tracto rectal de un cerdo (Australia, 2007). Además, el análisis de secuencia aminoacídica deducida (Blastx) del gen *intI* de LU22_37 mostró 99% de identidad y 96% de cobertura con una secuencia integrón integrasa de la cepa ABNIH6 de *Acinetobacter baumannii* (EMT91539.1) de origen clínico.

Asimismo, se obtuvieron resultados positivos para genes *intI* clase 2, en cuatro aislamientos provenientes de la misma muestra de la cual se aisló LU22_37 (LU6, Lago Uruguay). El análisis de secuencia confirmó la presencia de integrones clase 2 (*intI2*) en LU22G, LU23S, LU25G y LU26G. Las secuencias aminoacídicas

deducidas del gen *intI2* de los cuatro aislamientos resultaron ser idénticas entre sí.

El análisis de secuencia nucleotídica (Blastn) de LU22G, LU23S, LU25G y LU26G exhibió 100% de identidad y 100% de cobertura con la secuencia del gen *intI2* de la cepa FC1653 de *Sigella sonnei* (CP037997.1) de origen clínico. También, las secuencias aminoacídicas deducidas de LU22G, LU23S, LU25 y LU26G, (Blastx) mostraron identidad con la misma secuencia integrasa de clase 2 alojada en el plásmido R721 de *E. coli* (BAB120601.1) (Kim & Komano, 1992).

La búsqueda de genes *intI3* en los aislamientos analizados no dio resultados positivos. Este resultado es esperable, si se considera que esta clase de integrones ha sido reportada en pocas ocasiones y la mayoría se limitan al continente asiático.

Para determinar la identidad de los aislamientos se amplificó parte del gen ARNr 16S empleando cebadores universales 27f y 1492r (Lane et al., 1991).

El análisis de secuencia de LU22_37 (Blastn) exhibió 99% de identidad con la correspondiente de la cepa S13 de *E. hormaechei* (CP031565.) de origen clínico. Por su parte, LU26G mostró 98% de identidad (Blastn) con la cepa SF-3-08 de *Pseudomonas* sp. Esta cepa fue identificada en una especie de esponja de agua dulce (*Ephydatia fluviatilis*) (HE794905.1). Las secuencias ARNr 16S de LU22G, LU23S y LU25G exhibieron niveles de identidad menores a 97% con especies del género *Pseudomonas* sp., por lo que no es posible confirmar su identidad.

De un total de 206 aislamientos analizados solamente se logró recuperar un integrón de clase 1 y cuatro integrones de clase 2. Los cinco aislamientos fueron recuperados de muestras de agua dulce del Lago Uruguay (próximo a la BCAA). El bajo número de genes *intI* identificados empleando técnicas dependientes de cultivo es esperable cuando se realiza la búsqueda de integrones en organismos cultivables sin una selección previa de fenotipo resistente. También podría explicarse por una baja frecuencia de genes *intI* en el ambiente antártico.

2.3.2.-Conclusiones y Perspectivas

Se analizó una colección de 206 aislamientos obtenidos en un medio de cultivo para oligotrofos crecidos a 5°C y 25°C para la búsqueda de integrones de clase 1, 2 y 3 (Mazel et al. 2000).

Fue posible recuperar únicamente cinco aislamientos positivos para el gen integrasa *intI* que correspondieron a un integrón de clase 1 (LU22_37) y cuatro integrones de clase 2 (LU22G, LU23S, LU25G y LU26G). Todos los aislamientos provenían de la misma muestra de agua dulce (LU6) del Lago Uruguay, próximo a la BCAA.

El análisis de secuencia (Blastn y Blastx) permitió determinar que los cinco integrones están relacionados con integrasas de bacterias mesófilas de origen clínico.

La identidad de los aislamientos solo pudo ser confirmada en LU22_37 que exhibió identidad con una *Enterobacteria* de origen clínico y en LU26G que mostró identidad con una *Pseudomonas* de origen ambiental.

Para continuar este trabajo se propone:

Caracterizar de forma completa los integrones de clase 2 identificados en los aislamientos del Lago Uruguay. La caracterización incluye la identificación de los GCs que forman parte de la región variable del integrón.

II) Elementos móviles que portan integrones clase 1 en enterobacterias antárticas.

2.4 -Materiales y Métodos

2.4.1-Ensayos de conjugación bacteriana

Se empleó *Escherichia coli* DH5 α (Nal^r) como cepa receptora y cinco aislamientos Trim^r Amp^r Strp^r como cepas donadoras (GC4,IA12,TR5,CN11,HP19) (Tabla 2.1). Para los cultivos iniciales, las células receptoras y donadoras se resuspendieron en tubos con 5 ml de caldo LB (Luria Bertani) y con ácido nalidíxico (Nal) y estreptomicina (Strp), respectivamente, a una concentración final de 50 $\mu\text{g}/\text{mL}^1$. Los tubos se incubaron a 37°C con agitación (200rpm) durante toda la noche. Al día siguiente se transfirieron 50 μl de las suspensiones a tubos con 5 ml de caldo LB fresco sin antibiótico y se incubaron a 37°C hasta alcanzar una $\text{DO}_{620\text{nm}} = 0.5$ aproximadamente. Los cultivos se centrifugaron, se resuspendieron en 5 ml de caldo LB y se mezclaron diferentes volúmenes en una proporción de 1:1 y 10:1 (cél. receptora: célula donadora). La conjugación se realizó sobre filtros de acetato de celulosa (Sartorius, Alemania)(0,45 μm) en placas con medio LB sólido y se incubaron *overnight* a tres temperaturas (5°C, 25°C y 30°C). Al día siguiente, los filtros fueron transferidos a tubos con 5 ml de caldo LB y agitados con vórtex a máxima velocidad para desprender las células adheridas a los filtros. Se prepararon diluciones seriadas de las suspensiones y se sembraron por rastrillado volúmenes de 100 μl y 200 μl en placas con medio LB agar Strp (50 $\mu\text{g}/\text{mL}^1$) Nal (50 $\mu\text{g}/\text{mL}^1$). Las placas se incubaron a 37°C para la selección de transconjugantes.

2.4.2- Secuenciación plasmídica en plataforma Ion Torrent

Para obtener el ADN plasmídico, las cepas CN11, TR5 y HP19 fueron cultivadas en caldo Mueller-Hinton suplementado con Trim (50 $\mu\text{g}/\text{mL}^1$) a 25°C, *overnight*. El ADN fue purificado empleando el kit de extracción Gene Plasmid Midprep Elute® (Sigma-Aldrich, EE.UU.), siguiendo las instrucciones del fabricante. La secuenciación se realizó en un equipo Ion 316 Chip 197 en Ion PGMTM System (Thermo Fisher Scientific) en el IIBCE (Montevideo, Uruguay).

2.4.3- Secuenciación genómica en plataforma Illumina

Se purificó el ADN genómico total de una cepa de *E. coli* DH5 α transconjugante que contenía el plásmido pHP19, empleando el kit Genomic DNA DNeasy Blood & Tissue Kit (Qiagen, Alemania). La secuenciación completa del genoma fue realizada en un equipo Illumina HiSeq 2500 en Macrogen (Seúl, Corea del Sur).

2.4.4-Procesamiento de datos, anotación y análisis de secuencias

Los datos obtenidos a partir de la plataforma Ion Torrent fueron procesados empleando el programa CLC Genomics Workbench v9.5.1 (Qiagen, Alemania). Previo al ensamblaje se realizaron los controles de calidad correspondientes. Las lecturas pertenecientes a *E. coli* DH5 α obtenidos mediante la secuenciación Illumina fueron mapeadas contra el genoma de *E. coli* DH5 α (PRJNA205928), a efectos de sustraer las secuencias correspondientes a pHP19. Estas secuencias se combinaron con las lecturas de calidad obtenidas con la secuenciación en la plataforma Ion Torrent, de forma de generar un ensamblado consenso. Esto se llevó a cabo con el programa CLC Genomics Workbench De Novo Sequencing, empleando los parámetros predeterminados.

Todos los contigs obtenidos fueron analizados a efectos de eliminar aquellos que contenían secuencias de ADN provenientes de la bacteria receptora *E. coli* DH5 α , obteniéndose 12 contigs con un total de 248.262 pb.

La anotación de las secuencias de los plásmidos pTR5, pCN11 y pHP19 se hizo de forma automática en la plataforma online Rapid Annotation using Subsystems Technology (RAST) (<http://rast.nmpdr.org/rast.cgi>) (Aziz et al., 2008)

Las secuencias plasmídicas obtenidas fueron comparadas contra la base de datos del NCBI Genbank (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) empleando los programas BLASTn y BLASTx (Zhang et al., 2000).

Asimismo, para la anotación y visualización de pHP19 se empleó el programa Snap Gene view (http://www.snapgene.com/products/snapgene_viewer/).

2.5-Resultados

2.5.1-Ensayos de conjugación

Con el fin de determinar si los plásmidos que portaban el integrón clase 1 (*intI1*) eran de tipo conjugativo y podían ser transferidos entre bacterias y potencialmente entre enterobacterias, se realizaron ensayos de conjugación empleando los aislamientos GC4, IA12, TR5, CN11 y HP19, como cepas donadoras (Tabla 2.1) y células de *E. coli* DH5 α como cepa receptora.

Se obtuvieron transconjugantes de los aislamientos CN11 y HP19. Si bien se logró un número reducido de transconjugantes, se obtuvieron resultados a las tres temperaturas ensayadas (5°C, 25°C y 30°C).

Los transconjugantes también expresaron resistencia a Amp y Trim, sugiriendo que los genes que confieren esta resistencia están localizados en plásmidos. La presencia de genes *intI* en los transconjugantes fue confirmada por PCR y la identidad de las cepas fue determinada mediante análisis de secuencia del gen ARNr 16S y su incapacidad de crecer a 5°C.

2.5.2- Secuenciación plasmídica

Mediante secuenciación *shotgun* en una plataforma Ion Torrent (IIBCE, Montevideo), se obtuvo la secuencia completa de los plásmidos pTR5, pCN11 y pHP19, purificados a partir de tres enterobacterias antárticas (TR5, CN11 y HP19 respectivamente). Estos aislamientos pueden albergar más de un plásmido. La denominación pTR5, pCN11 y pHP19 refiere a los replicones que puedan estar contenidos en los aislamientos parentales correspondientes.

Para los plásmido(s) de TR5 (pTR5), la reacción de secuenciación generó 86175 lecturas con un largo promedio de 204 pb. Para los plásmido(s) de CN11 (pCN11), se generaron 97875 lecturas de 206 pb de largo y para los de HP19 (pHP19) 89140 lecturas de 202 pb de largo promedio, con un total de 17, 20 y 18 millones de bases respectivamente.

El resultado obtenido de la secuenciación permitió determinar que los plásmidos pTR5, pCN11 y pHP19 eran parcialmente idénticos entre sí en la región correspondiente al integrón.

En la Figura 2.2 se muestran los ORFs anotados por el servidor RAST en los contigs que contiene el integrón para cada uno de los plásmidos, pTR5, pCN11 y pHP19.

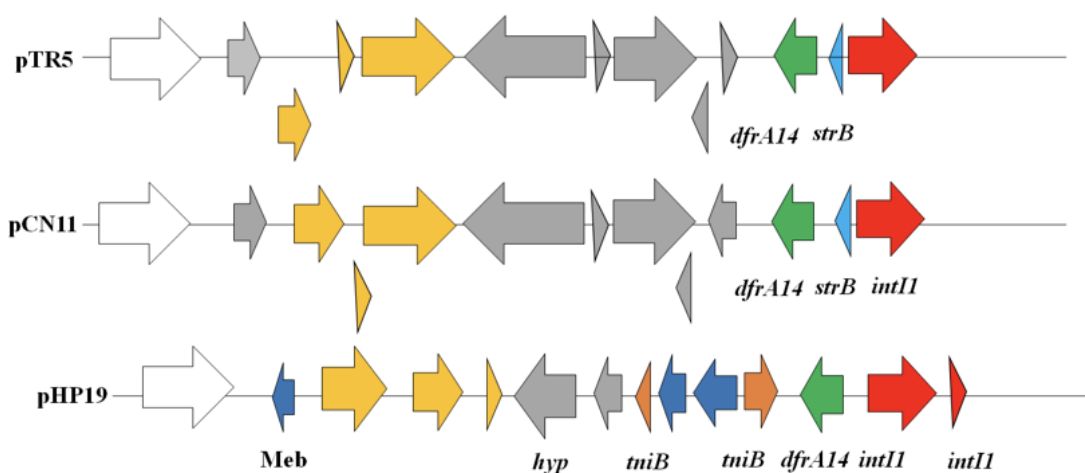


Figura 2.2: Organización parcial de los ORFs identificados en los contigs 2 de pTR5, 15 de pCN11 y 26 de pHP19.

Gen integrasa *intI1* (rojo), gen de resistencia a trimetoprim *dfrA14* (verde) gen de resistencia a estreptomicina *strB* (celeste). Proteínas de unión a membrana (naranja), DNA citosina metil-transferasa (amarillo), proteínas hipotéticas (gris) y una probable proteína de membrana (blanco).

Además de las secuencias plasmídicas esperadas, algunos de los contigs obtenidos contenían secuencias de origen presumiblemente cromosómico. A efectos de obtener secuencias plasmídicas sin contaminación cromosomal, se seleccionó uno de los transconjugantes obtenidos previamente. En este caso, se dispone del genoma completo de la cepa hospedera (*E. coli* DH5 α). El clon *E.coli* DH5 α (pHP19), fue secuenciado en una plataforma Illumina HiSeq2500 en Macrogen (Seúl, Corea del Sur).

La reacción de secuenciación generó 1.015.5134 lecturas en un total de 1.0G pb con un inserto promedio de 420 pb de largo.

De esta forma, la secuencia del plásmido pHP19 se recuperó a partir de los datos de secuenciación del genoma completo de la cepa transconjugante. En este caso, presumimos que el evento de conjugación más probable es el que permite la transferencia de un solo plásmido, aunque no es posible afirmarlo. Se sustrajeron los contigs cromosomales y se compararon los resultados obtenidos en las secuenciaciones con las plataformas Ion Torrent e Illumina. Finalmente, se obtuvieron doce contigs que sumaron un total de 248.262 pb a partir de pHP19.

La mayoría de las secuencias obtenidas en nuestro estudio exhibieron un alto grado de identidad (88-100%) a nivel nucleotídico (Blastn) con las de p34998-210.894kb-D de *E. hormaechei* subsp. *steigerwaltii* 34998 (CP012169.1) (Chavda et al., 2016). Estas secuencias se distribuyen en seis contigs (Nº1 al Nº6) que suman 200.012 pb y representan el 94,84% de p34998-210.894kb-D.

Por otra parte, los contigs Nº 7 al Nº 12, exhibieron 75-100% identidad (Blastn) con el plásmido pENT-d4a-60338 pb (CP008900.1) de la cepa ECNIH3 de *Enterobacter cloacae* KPC⁺, aislada de un centro hospitalario (año 2011, EUA) (Conlan et al., 2014). Estos seis contigs suman un total de 48250 pb y representan el 79,96% de pENT-d4a-60338 pb (Ver Tabla 2.2).

Tabla 2.2: Se muestran las posiciones en el plásmido pENT-d4a que exhibieron identidad con los contigs de pHP19.

Contigs de pHP19	Posiciones en pENT-d4a	identidad
Contig N° 7	49570-50022, 50619-50882, 52744-53740, 53819-54063 y 54089-55010.	84-100 %
Contig N°8	5275 - 7870 y 7913 - 14629.	90-100 %
Contig N°9	21246 - 33671	97-100 %
Contig N°10	33852-39926	93-99 %
Contig N°11	40665-45317 y 46090-46706	85-95 %
Contig N°12	1-5309 y 58679-60388	90-98 %

Asimismo, el contig N°6 que contiene el integrón, es 100% idéntico (desde la posición 51362 a la 54641) a la región originalmente identificada en pKOX105 de *K. oxytoca* de origen clínico (Carattoli et al., 2010).

En total se identificaron 317 ORFs que codifican para el gen integrasa de clase 1 (*intI1*), genes de resistencia a aminoglucósidos, β -lactámicos, trimetoprim, cloramfenicol y genes de resistencia a metales (arsénico, cobre, telurito, mercurio y níquel).

Además, pHP19 contiene varios ORFs que codifican para proteínas necesarias para la conjugación, enzimas involucradas en el ensamblaje del sistema de secreción tipo IV y otros ORFs que codifican para proteínas de replicación, partición y estabilidad plasmídica, entre otras.

A continuación se presentan los doce contigs de pHP19 ordenados en base al tamaño:

El contig de mayor tamaño, contig 2 (115.734 pb) (Figura 2.3) contiene 145 ORFs, de los cuales un 44% codifican para proteínas hipotéticas, un 15% codifica para resistencia a metales y antibióticos, otro 15% para proteínas involucradas en el proceso de conjugación y estabilidad plasmídica y un 10% para proteínas de movilidad.

Este contig exhibió 99-100% de identidad con dos regiones que se extienden desde la posición 72.628 a la 112.768 y desde la 114.724 a la 190339, en el plásmido p34998-210.894kb de *Enterobacter hormaechei* subsp *steirgerwaltii* cepa 34998 (CP012169.1) que representan el 99,94% del contig 2. Además, incluye varios ORFs que codifican para proteínas que componen la maquinaria de conjugación de plásmidos del grupo de incompatibilidad IncF (TraD, TraE, TraF, TraG, TraH, TraI, TraK, TraN, TraT, TraU, TraV, TraW, TraX, TrbB, TrbC). Los plásmidos del grupo F poseen un operón *tra* que contiene todos los genes necesarios para el ensamblaje del poro conjugativo y la transferencia del ADN plasmídico (Wong et al., 2012). Las proteínas Tra forman el transferosoma (poro transmembrana por el cual la molécula de ADN es transferida a la célula receptora) y el relaxosoma, complejo de unión al ADN en el *oriT* (Wong et al., 2012). El relaxosoma incluye una relaxasa-helicasa (TraI) que se une específicamente a la secuencia *oriT*, que corta y desenrolla el ADN, una proteína de unión al ADN (TraM), que media la señalización para el inicio de la transferencia y una proteína de transferencia con actividad ATPasa (TraD).

En el contig 2 de pHP19 se identificó un ORF que codifica para la proteína multifuncional TraI (1807 aa) y la proteína de transferencia TraD (851 aa). Sin embargo, no se identificó la proteína de unión al ADN, TraM.

Asimismo, cabe mencionar que en el contig 12 (7346 pb) se identificaron dos ORFs que codifican para una relaxasa-helicasa TrwC, que es homóloga a TraI. TrwC forma parte del relaxosoma del plásmido R388, IncW (César et al., 2006). Las proteínas TrwC de pHP19, de 327 aa (ATZ71634.1) y 331 aa (ATZ71637.1), exhibieron 95% y 98% de identidad (a nivel nucleotídico, Blastn) con la región que se extiende desde la posición 3025 a 4009 y desde 1115 a 2110,

respectivamente, en el plásmido pENT_d4a de *Enterobacter cloacae* cepa ECNIH3 (CP008900.1).

pHP19 codifica también para proteínas involucradas en el ensamblaje del pilus del sistema conjugativo de replicones IncF (TraH, TraF, TraU, TraW, TraC, TraV, TraB, TraK y TraE).



Figura 2.3: Contexto genómico del contig 2 de pHP19 de *Enterobacter* sp. HP19. Se muestran ORFs involucrados en el proceso de conjugación (verde) y estabilidad plasmídica (rosado). ORFs que codifican para resistencia a metales y antibióticos (naranja), para proteínas de movilidad (lila), proteínas hipotéticas (amarillo) y otros ORFs (gris). Anotación RAST-Server (Aziz et al., 2008). Visualización SnapGene® Viewer 4.2.11.

Se anotaron los ORFs que codifican para proteínas de resistencia a metales. Se identificaron tres proteínas TerD (579pb, 576pb, 582pb), una proteína TerA

(1158pb) que confieren resistencia a telurio, una proteína TerB (456pb) de resistencia a telurito y una proteína integral de membrana, anotada como TerC (1041pb).

El operón de resistencia a arsénico incluye ORFs que codifican para el represor del operón, ArsD (363pb), una proteína de resistencia, ArsH (714pb), una bomba de eflujo (1290pb) y una ATPasa transportadora de arsenito (1752pb). También se anotaron genes que codificarían para dos proteínas de resistencia a compuestos de cromo, una proteína de transporte, ChrA (1377pb) y una proteína de resistencia, ChrB (978pb).

El contig 3 de 47601 pb (Figura 2.4) posee 53 ORFs, de los cuales un 40% codifican para proteínas involucradas en la resistencia a metales, factores de virulencia y respuesta al estrés, un 34% codifican para proteínas hipotéticas y un 9% para proteínas de partición y replicación plasmídica.

Este contig exhibió un 99% de identidad con p34998-210.894kb entre la posición 204.300 y 210.894 y desde la posición uno hasta 41.006. En este contig se encuentran dos ORFs (*repA*) que codifican para proteínas de inicio de la replicación (RepFIIA y RepFIIB), así como también proteínas de partición (*parA/parB* y *sopA/sopB*). Además, contiene un ORF que codifica para una relaxasa acetilasa, necesaria para el ensamblaje del pili conjugativo (TraX).

También se encuentran ORFs que codificarían para proteínas involucradas en resistencia a cobre (CusC, CusF, CopC, CopD y CopG) y proteínas de unión al metal (PcoE). Además, se anotaron dos ORFs que codifican para un sistema de dos componentes, sensor de metales pesados: una histidin-quinasa (1401 pb) y la proteína reguladora de respuesta (681pb).

estreptomycin fue confirmada experimentalmente en todos los trasconjugantes. Se identificaron también ORFs que codificarían para una dihidropteroato sintetasa (EC 2.5.1.15) (816 pb) que confiere resistencia a sulfonamidas y para una β -lactamasa (861 pb), que se sitúa contiguo a tres ORFs que codifican para proteínas de movilidad (de 153 pb, 402 pb y 3165 pb). Estas tres proteínas de movilidad junto con la β -lactamasa constituirían el módulo del transposón Tn3 (Grindley 2002).

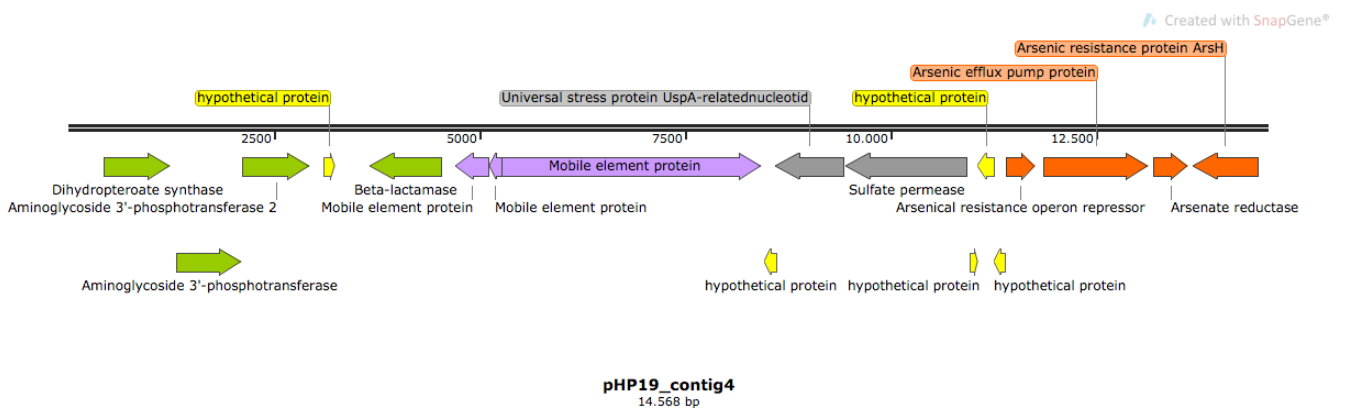


Figura 2.5: Representación esquemática del contig 4 de pHP19. Este contig incluye cuatro ORFs que codifican para resistencia a antibióticos (verde). Tres ORFs que codifican para las proteínas de movilidad (violeta), cuatro ORFs que integran el subsistema de resistencia a arsénico (naranja) y cinco proteínas hipotéticas (amarillo). Los ORF en color gris codifican para una sulfato permeasa y una proteína UspA involucrada en respuesta a factores de estrés. Anotación RAST-Server (Aziz et al., 2008). Visualización SnapGene® Viewer 4.2.11.

El contig 9 (9289 pb) (Figura 2.6) contiene ORFs que codifican para proteínas involucradas en el sistema de secreción tipo IV (T4SS) y ocho proteínas hipotéticas. Posee dos ORFs que codificarían para ATPasas citoplasmáticas encargadas de proporcionar la energía necesaria para la transferencia y ensamblaje del sistema de secreción tipo IV (VirB4 y VirB11). Las proteínas de la familia VirB4 están muy conservadas y a diferencia de otros componentes del sistema, están presentes en todos los T4SSs descritos hasta el momento. VirB4 es importante para la estabilización del resto de los componentes del T4SS. Además,

tiene un rol esencial en la formación del pilus, transporte del sustrato y en la virulencia.

La familia de proteínas VirB11 se encuentra ampliamente distribuida en sistemas conjugativos y bacterias patógenas Gram-negativas. Estas proteínas funcionan como reguladores en la formación del pilus y también en la translocación del sustrato. Asimismo, están presentes los ORFs que codificarían para dos proteínas que forman el canal periplásmico (VirB9 y VirB10), una proteína que forma la base del canal en la membrana interna (VirB8), otra proteína que forma el pilus (VirB5) y una lipoproteína de transferencia (PilN).

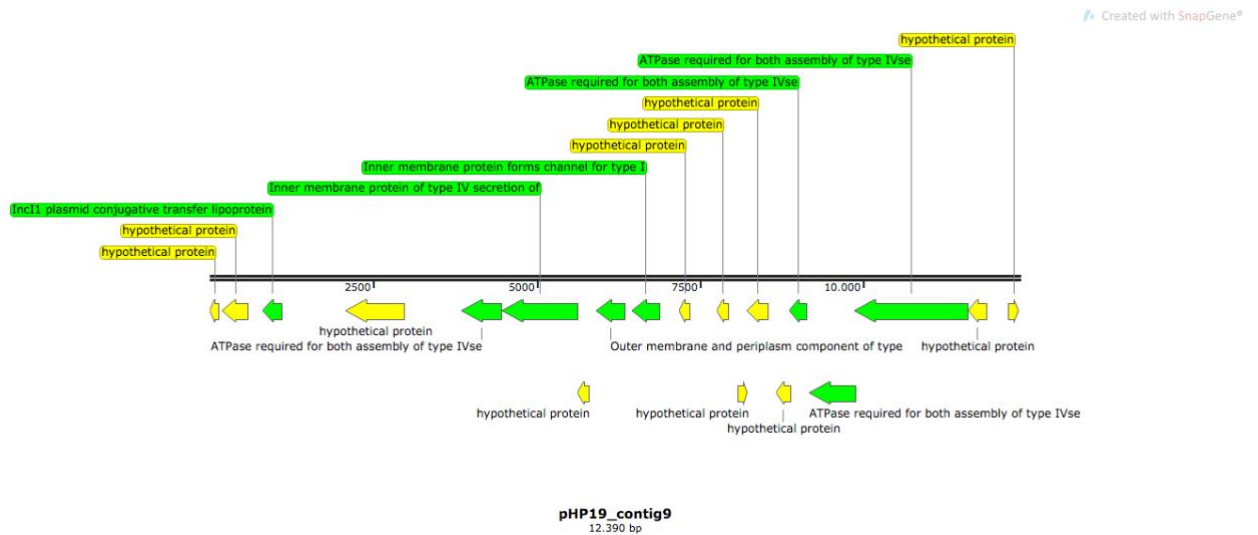


Figura 2.6 : Contig 9 de pHP19 que contiene ORFs que codifican para proteínas del sistema de secreción tipo IV (verde) y proteínas hipotéticas (amarillo). Anotación RAST-Server(Aziz et al., 2008). Visualización SnapGene® Viewer 4.2.11.

El contig 6 (Figura 2.10) de pHP19 (7239 pb) exhibió 100% de identidad con la región que contiene un integrón clase 1, desde la posición 46126 a la posición 53364, de p34998-D. Esta región contiene el gen integrasa *intI1* y otros ORFs que codifican para una proteína de movilidad, dos probables proteínas de unión a ATP (TniB) (Pfam 05621) y una DNA citosina metiltransferasa (EC 2.1.1.37).

Adyacente al integrón se encuentra el GC que codifica para la enzima dihidrofolato reductasa DfrA14 (EC 1.5.1.3). que confiere resistencia a Trim y un ORF que codificaría para una proteína hipotética (1209 pb), que exhibe un 100% de identidad (Blastp) con una endonucleasa de restricción *EcoRII* (WP_001749969.1). Este contig incluye también un ORF que codificaría para una probable proteína de membrana STY4561 (510 pb).

Además este contig también es muy similar a la región originalmente identificada en pKOX105 de *K. oxytoca* de origen clínico (Carattoli et al., 2010). La región 100% idéntica (desde la posición 51362 a la 54641) incluye los genes *intI1*, *dfrA14*, dos ORFs que codifican para proteínas de movilidad, una DNA citosina metiltransferasa, una probable proteína de membrana STY4561 y dos proteínas hipotéticas.

Asimismo, empleando el programa Integron Finder (Cury et al. 2016), fue posible identificar el promotor de *intI1* (Pint), el sitio de recombinación primario *attI*, el promotor Pc y el sitio de recombinación secundario *attC* correspondiente al GC *dfrA14*.

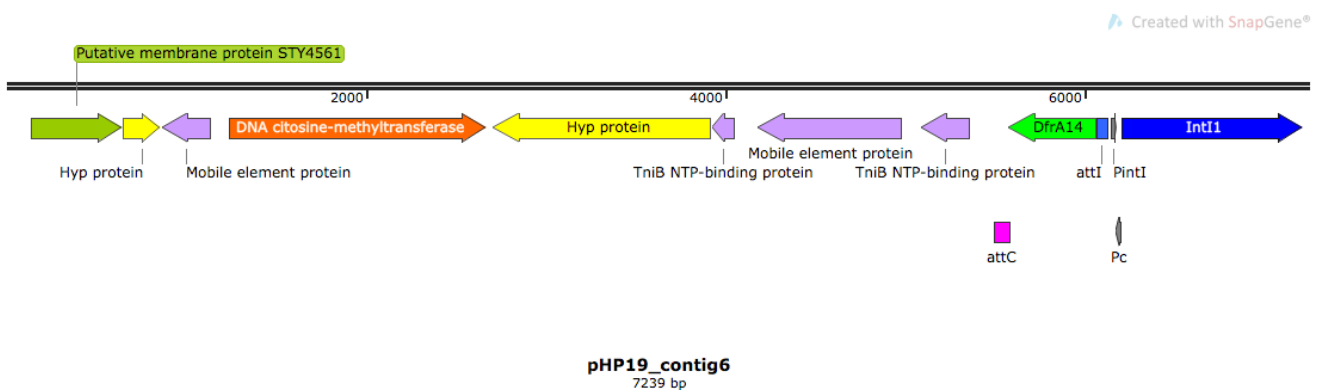


Figura 2.10: Representación esquemática de la región de pHP19 que contiene el integrón. Se muestra el gen integrasa clase 1, *intI1* (azul) contiguo al promotor que dirige su expresión, Pint (gris), el sitio *attI* (celeste), el promotor Pc (gris) que controla la expresión del gen *cassette dfrA14* (verde) y el sitio *attC* (rosado). Se indican también los ORFs que codifican para proteínas de movilidad y dos proteínas TniB (lila), dos proteínas hipotéticas (amarillo), una DNA citosina metiltransferasa (naranja) y una probable proteína de unión a membrana (verde claro). Anotación con Integron Finder (Cury et al. 2016) y RAST-Server (Aziz et al., 2008). Visualización Snap Gene © Viewer v 4.2.11.

El contig 10 (6076 pb) contiene ORFs que codifican para proteínas necesarias para el sistema de secreción tipo II (PulF y PilQ) y una ATPasa involucrada en la conjugación.

En el contig 11 (5702 pb) hay ORFs responsables de la expresión de dos proteínas hipotéticas, una DNA-recombinasa, una serin acetiltransferasa y una proteína del sistema de transferencia conjugativo del grupo de incompatibilidad plasmídica IncQ (TraQ).

El contig 12 (7346 pb) contiene ORFs que codifican para dos resolvasas, una proteína del sistema de conjugación del grupo de incompatibilidad IncF (TraD) y dos relaxasas del sistema de conjugación del grupo de incompatibilidad IncW (TrwC) y siete proteínas hipotéticas.

En la figura 2.11 se muestra el mapa que representa el alineamiento de los contigs 2 y 6 de pHP19 comparados con p34998. Las regiones que exhibieron identidad alta (99%-100%) incluyen genes involucrados en el proceso de conjugación, estabilidad plasmídica, resistencia a metales y antibióticos y también el integrón de clase 1 con sus GC asociados.

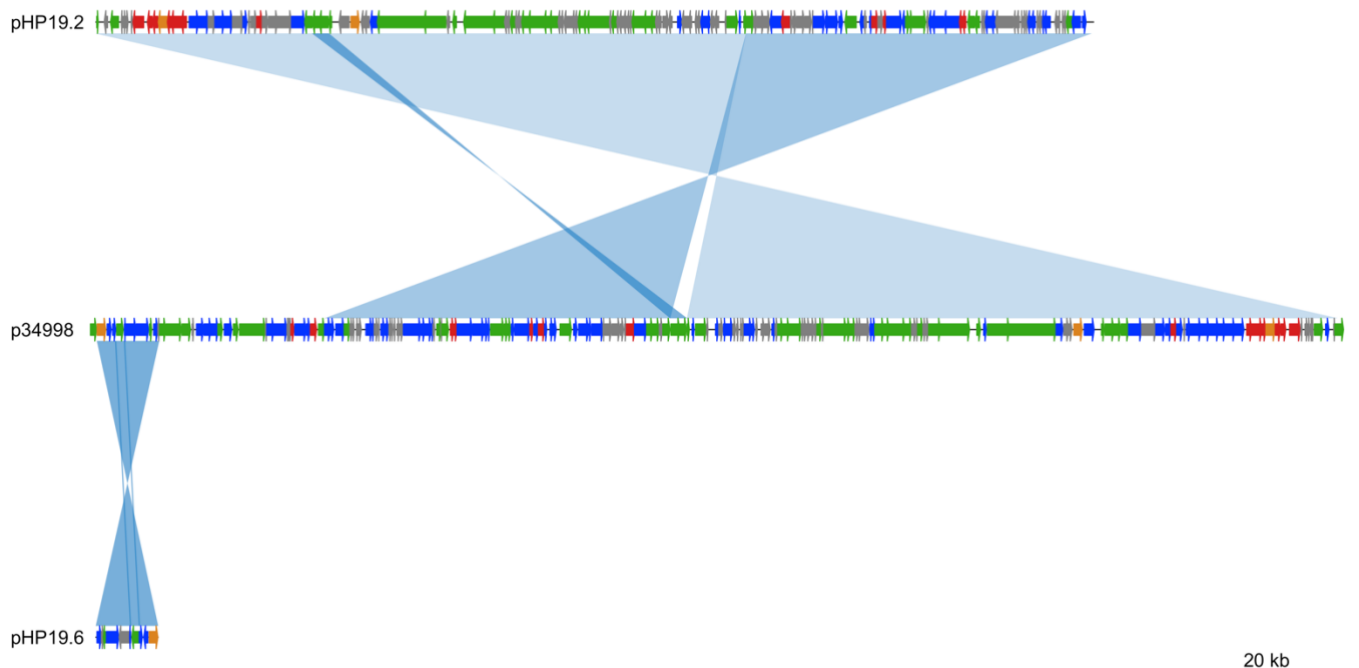


Figura 2.11: Representación esquemática del alineamiento pHP19 (contigs 2 y 6) contra p34998. ORFs involucrados en movilidad, conjugación y estabilidad (verde). ORFs de resistencia (rojo). ORFs que codifican para proteínas hipotéticas (gris). ORFs de funciones diversas (azul).

2.6-Discusión

En este capítulo se describieron ensayos de conjugación realizados con aislamientos de enterobacterias antárticas psicrotolerantes que portan integrones clase 1 de tipo “clínico” (TR5, CN11 y HP19). Se obtuvieron transconjugantes a partir de ensayos diseñados con las cepas CN11 y HP19 a tres temperaturas (5°, 25° y 30°), usando *E. coli* como cepa recipiente. Estas enterobacterias fueron aisladas a partir de muestras colectadas en distintos sitios de la Península Fildes (Isla Rey Jorge). Los sitios de colecta incluyeron zonas bajo diferentes grados de influencia antropogénica y de presencia de animales, incluyendo áreas con acceso restringido como las zonas especialmente protegidas (ZAEP). TR5 fue colectada en un área próxima a los viejos depósitos de almacenamiento de combustible pertenecientes a Rusia, usados por la Base Artigas hasta hace poco años. CN11 (Caleta Norma) está ubicada relativamente cerca de la BCAA, en una zona elevada y próxima a la costa. HP19 proviene del Halfthree-Point, una zona que forma parte de un área especialmente protegida (ZAEP 125d), donde se requiere un permiso especial para su ingreso, que se encuentra distante a las Bases de operaciones (Antelo, et al., 2015).

El integrón contenido en estas enterobacterias está ampliamente conservado en los aislamientos analizados. Tanto el gen integrasa *intI1* como el GC *dfxA14* resultaron ser idénticos en todos los aislamientos (Antelo et al. 2018). El análisis de tipificación bacteriana rep-PCR mostró un perfil electroforético muy similar para todos los aislamientos, sugiriendo que podría tratarse de un mismo organismo distribuido en la isla y que un componente externo podría funcionar como factor de dispersión.

La estrategia diseñada para la identificación de integrones clase 1 en organismos cultivables, introdujo un doble sesgo. Por un lado el cultivo en un medio enriquecido diseñado para la selección de un determinado grupo de organismos, favoreció la selección de enterobacterias. Por otro lado, también se introdujo un sesgo adicional con el antibiótico elegido para la selección (Trim), dado que los GCs de resistencia a Trim se encuentran comúnmente asociados a integrones clase 1.

En un estudio paralelo desarrollado con muestras colectadas en Halfthree-Point (2008) y Caleta Norma (2013) (ver Capítulo 3 Sección 3.2) se construyeron

bibliotecas de amplicones empleando el mismo par de cebadores dirigidos a amplificar el gen *intI1* (Mazel et al., 2000). En este caso se eliminó el sesgo de las técnicas de cultivo y fue posible identificar una gran variedad de integrasas ambientales, además de las integrasas clase 1 de tipo “clínico” (Antelo et al., 2015). El análisis filogenético de las integrasas permitió determinar que este par de cebadores no es específico para amplificar integrasas clase 1 de origen clínico. En ese estudio se identificó una gran diversidad de secuencias que se distribuyeron en varios grupos claramente separados de las integrasas móviles de clase 1, 2 y 3 (clínicas) (Antelo et al., 2015).

Complementando los ensayos de conjugación, el plásmido pHP19 fue secuenciado empleando la técnica *shotgun* en una plataforma Ion Torrent e Illumina y ensamblado de forma parcial.

Nuestro estudio permitió determinar que los plásmidos pTR5, pCN11 y pHP19, contenidos en enterobacterias antárticas psicrotolerantes eran idénticos entre sí y muy similares a los plásmidos descritos en enterobacterias mesófilas aisladas del ambiente hospitalario.

La mayoría de las secuencias obtenidas exhibieron alta similitud con las del plásmido p34998-D-210.894kb of *E. hormaechei* subsp. *steigerwaltii* de origen clínico (CP012169.1)(Chavda et al., 2016).

Estas secuencias incluyeron genes requeridos para la replicación, estabilidad y conjugación. Estos resultados sugieren que pHP19, al igual que p34998-210.894kb (CP012169.1), posee dos orígenes de replicación y dos sistemas involucrados en la estabilidad plasmídica. Las regiones son muy similares (99% por BLASTn) a las correspondientes de p34998-210.894kb. Los genes involucrados en el proceso de conjugación posiblemente sean funcionales, dado que fue posible transferir pHP19 desde su hospedero (HP19) a células de *E. coli* DH5 α a diferentes temperaturas.

Por último, pHP19 también contiene un fragmento conservado (~7200 bp) casi idéntico a la región de p34998-210.894kb y de pKOX105 (54641bp) (HM126016) que incluye el integrón de clase 1 y el GC *dfrA14* de resistencia a Trim.

Asimismo, se identificaron *clusters* de genes relacionados con la resistencia a metales como mercurio, arsénico, telurio, níquel y cobre, 100% idénticos a los genes encontrados en p34998-210.894kb. Otras regiones, sin embargo, fueron similares a las del plásmido pENT-d4a de la cepa ECNIH3 de *E. cloacae*, aislada

del tracto respiratorio de pacientes internados en un centro hospitalario (Conlan et al., 2014).

Estos resultados sugieren que pHP19 está conformado por un mosaico de secuencias de otros plásmidos que han sido previamente caracterizados en enterobacterias mesófilas de origen clínico. El origen y evolución de pHP19 probablemente sea difícil de determinar, sin embargo resulta evidente que la THG entre bacterias mesófilas y bacterias nativas es un hecho frecuente en el ecosistema antártico.

Al igual que p34998-D, pHP19 pertenece al grupo de incompatibilidad IncF. Los plásmidos IncF multi-resistentes poseen un rango de huésped limitado a los géneros de la familia *Enterobacteriaceae*. Han sido descritos en especies de enterobacterias que causan severas infecciones a los humanos. Los determinantes de resistencia identificados en esta familia incluyen genes de resistencia a aminoglucósidos, quinolonas y β -lactamasas (Carattoli, 2009).

Los plásmidos IncF dotan a la célula del mecanismo denominado exclusión de superficie, mediante el cual una célula que contiene un plásmido del grupo IncF no podrá incorporar otro plásmido del mismo grupo en un siguiente proceso de conjugación. Esta propiedad viene dada por una proteína específica de plásmidos denominada TraT. En células de *E. coli*, se observó además que la proteína TraT confiere resistencia a la célula hospedera, disminuyendo su susceptibilidad a la fagocitosis (Carattoli, 2009).

Los plásmidos IncF están ampliamente extendidos en células de *E. coli* productoras de β -lactamasas TEM-1. El ORF que codifica para la β -lactamasa identificado en pHP19 exhibió un 100% de identidad (Blastn) con el gen *bla*_{TEM-1} de *E. coli* plásmido pVOG1(FJ183463) y con el gen *bla*_{TEM-1} de p34998-D de *E. hormaechei* subsp. *steigerwaltii* (CP012169.1). Este ORF está contiguo a tres ORFs que codifican para una recombinasa, una resolvasa y una transposasa que juntos forman el módulo de transposición de Tn3.

Los genes *bla*_{TEM-1} son movilizados por el transposón Tn3 y generalmente coexisten con genes *bla*_{CTX-M-15} en el mismo plásmido (Grindley 2002).

En un trabajo desarrollado en la Península Antártica, Hernández y col., (2012) identificaron genes *bla*_{CTX-M-1} y *bla*_{CTX-M-15} en cepas de *E. coli* aisladas de muestras de agua en sitios próximos a las Bases científicas en la Península Antártica (Hernández et al., 2012).

La rápida emergencia de β -lactamasas CTX-M y su diseminación a nivel global entre organismos patógenos es un claro ejemplo de la función que cumplen los elementos genéticos móviles en la THG (Rossolini et al. , 2008).

Las ESBL (β -lactamasas de espectro extendido) han sido bien estudiadas en el ambiente hospitalario pero su caracterización en bacterias ambientales es aún escasa. Si bien las enterobacterias descritas en este trabajo no poseen β -lactamasas CTX-M, contienen la plataforma genética (transposón Tn3) que comúnmente alberga estos genes y podrían ser por lo tanto incorporados al mismo. Además, su asociación a elementos móviles, particularmente un plásmido conjugativo (pHP19) que porta un integrón clase 1 facilitaría su dispersión mediante THG entre bacterias que se encuentran en el ambiente antártico.

Recientemente Phan y col., 2015, determinaron la secuencia completa del plásmido multi-resistente pEC958 (135,6kb) (IncF) de *E.coli* ST131, una cepa asociada a infecciones del tracto urinario e infecciones sanguíneas. Al igual que pHP19, pEC958 tiene dos replicones, RepFIA y RepFII. En pEC958, RepFII es necesario para la replicación mientras que RepFIA contiene genes esenciales para la partición. Los autores sugieren que ambas regiones RepFIA y RepFII actúan de forma cooperativa para asegurar su estabilidad y transmisión a la descendencia. La asociación de la replicación (RepFII) y la partición (RepFIA) podría representar una adaptación co-evolutiva en plásmidos del grupo F (Phan et al., 2015).

En un estudio previo desarrollado en Bahía Fildes (Isla Rey Jorge), Power y col. (2016) identificaron integrones clase 1 en 62 cepas de *E. coli* antárticas aisladas de sedimento, agua de mar y del intestino de invertebrados marinos (*Legionella elliptica*). Estos integrones tenían las características típicas de integrones clínicos. La región variable contenía GCs de resistencia a sulfonamidas (*dfrA5* y *dfrA17*), aminoglucósidos (*aadA1*, *aadA2*, *aadA4* y *aadA5*) y en el extremo 3' el segmento conservado *qacED* (Power et al., 2016).

Al igual que los integrones de clase 1 identificados en este estudio, los genes *intI1* caracterizados por Power y col., exhibieron 100% de identidad a nivel nucleotídico con aislamientos clínicos de origen humano (Power et al., 2016)

2.7-Conclusiones y Perspectivas

En este trabajo se identificaron integrones clase 1 de tipo “clínico” en enterobacterias psicotolerantes resistentes a antibióticos, aisladas de sitios expuestos a diferentes condiciones por la presencia humana y animal, en la Península Fildes (Isla Rey Jorge, Antártida marítima).

Se determinó la presencia del gen *intI1* en tres plásmidos, pCTR5, pCN11 y pHP19, contenidos en cepas de *Enterobacter* sp. CTR5, CN11 y HP19, respectivamente. Estos plásmidos fueron transferidos a células de *E.coli* DH5 α mediante conjugación a tres temperaturas (5°C, 25°C y 30°C).

Mediante secuenciación *shotgun* se determinó la secuencia de pCN11, pTR5 y pHP19. El plásmido pHP19 fue secuenciado de forma completa y ensamblado en 12 contigs. El análisis de secuencia a nivel nucleotídico exhibió alta similitud (99-100%) con el plásmido p34998-210.894kb de *E.hormaechei* subsp. *steigerwaltii* de origen clínico (CP012169.1).

Asimismo, seis contigs que representan el 19,43% de pHP19, fueron muy similares al plásmido pENT-d4a-60.338kb originalmente aislado de *E. cloacae* ECNIH3 de origen clínico (CP008900.1).

Finalmente, se debe considerar que la estrategia empleada para la identificación de integrones clase 1 introdujo un doble sesgo al seleccionar en un medio de cultivo enriquecido con antibióticos que favoreció la selección de enterobacterias. Por esta razón, no es posible especular sobre la real abundancia de genes *intI1* de tipo “clínico” en el ambiente antártico.

Tampoco es posible afirmar si las enterobacterias psicotolerantes que portan los integrones son organismos autóctonos que adquirieron estos elementos genéticos mediante THG o son enterobacterias exógenas adaptadas al frío, que fueron introducidas por agentes de dispersión como aves migratorias, el viento o por el hombre.

Para describir de forma completa el plásmido que porta el integrón, este trabajo debería continuar con la caracterización funcional de pHP19, verificando la resistencia a todos los metales que se describen en su secuencia.

Asimismo, se propone determinar la abundancia del gen *intI1* de tipo “clínico” empleando un par de cebadores específicos para la variante clínica de integrones de clase 1.

2.8-Conclusiones generales

Se evidenció la presencia de integrones clase 1 de tipo “clínico” en enterobacterias psicotolerantes resistentes a antibióticos, en el ecosistema Antártico. Estos integrones están localizados en plásmidos conjugativos de gran tamaño (>200 kb), muy similares a plásmidos descritos en enterobacterias mesófilas de origen clínico. La conjugación fue verificada a bajas temperaturas (5°C), sugiriendo la habilidad de estas enterobacterias de transferir el fenotipo de resistencia antibiótica entre comunidades bacterianas de un ambiente extremo y distante geográficamente a centros hospitalarios.

CAPÍTULO 3

3. Estrategias independientes de cultivo

I-Construcción de bibliotecas de amplicones *intI*

II-Análisis taxonómico y funcional *in silico* de metagenomas de tapetes microbianos y sedimento-suelo. Identificación de genes integrasa *intI* y GCs asociados.

3.1 Introducción

3.1.1- Integrones en ecosistemas naturales

En un comienzo, los estudios de integrones se focalizaron en los ambientes clínicos y los mismos se basaban en métodos dependientes de cultivo. Sin embargo, la incorporación de nuevos métodos independientes de cultivo permitió generar nueva información. Los estudios metagenómicos han permitido recuperar nuevas secuencias de integrasas, demostrando que existe una gran diversidad genética, la cual en su mayoría es aún desconocida (Nielsen et al. 2004; Nemergut et al., 2004; Gillings et al., 2008; Rodríguez-Minguela et al., 2009; Elsaied et al. 2011 ; Elsaied et al., 2014).

El análisis de *dataset* metagenómicos ha revelado que existe un extenso *pool* de GCs ambientales que codifican para funciones de adaptación al ambiente y otros, en su mayoría, de función desconocida (Holmes et al., 2003).

El obstáculo en la identificación y caracterización de las funciones de adaptación de los integrones radica en que existen pocas herramientas que ayuden al reconocimiento de los sitios *attC* de los GCs (Cury et al., 2016). Si bien existen algunas herramientas que identifican secuencias conservadas, éstas fueron diseñadas para reconocer los sitios *attC* de integrones móviles de clase 1, 2 y 3. Estas herramientas resultan inadecuadas para identificar y alinear sitios *attC* con secuencias muy divergentes (Cury et al. 2016).

3.1.2- Descripción de las herramientas empleadas para el análisis in silico de metagenomas

i) Servidor MG-RAST

MG-RAST v4.0.3 (*Metagenomics Rapid Annotation using Subsystem Technology*) es un servidor de acceso libre creado con el fin de proporcionar un sistema de alto rendimiento para el análisis filogenético y la anotación funcional de metagenomas (Glass & Meyer, 2011).

MG-RAST procesa y lleva adelante la asignación automática de secuencias de ADN ambiental mediante la comparación con bases de datos de nucleótidos y

proteínas. La anotación está organizada en módulos, de modo que permite incorporar nuevos pasos en el análisis o comparar datos en cualquiera de las etapas. Por último, los resultados del análisis pueden ser descargados y analizados en variedad de formatos de uso frecuente. En la Figura 3.1 se describen los distintos pasos que sigue un análisis clásico en MG-RAST.

El servidor presenta por defecto valores pre-establecidos para los distintos parámetros que pueden ser cambiados por el usuario.

La primera etapa consiste en el procesamiento de las lecturas que pasan por un control que elimina las secuencias que no cumplan con niveles de calidad adecuados. También se remueven los artefactos que se generan durante el proceso de secuenciación, se eliminan las secuencias duplicadas y se asigna a cada lectura un número de identificación único (ID).

La segunda etapa abarca la búsqueda de genes, empleando la aproximación *Frag Gene Scan* que puede identificar regiones codificantes en secuencias de ADN de 75 pb o más. Esta nueva aproximación permite el análisis en datos crudos y también en *contigs* ya ensamblados. Cabe mencionar que el uso de *FragGeneScan* sólo es adecuado para análisis de organismos procariontes.

La identificación de proteínas y de ARNr, potencialmente codificadas por las regiones identificadas, se realiza en paralelo mediante la construcción de *clusters* usando como criterio un 90% y 97% de identidad respectivamente y seleccionando la más representativa de cada grupo.

El servidor MG-RAST emplea la base de datos M5NR (Wilke et al., 2012) para determinar la afiliación filogenética de las secuencias. La base de datos no redundante M5NR contiene secuencias y anotaciones de diversas fuentes. Posteriormente, la búsqueda de similitud se realiza con el algoritmo BLAT (Kent, 2002) contra una base de datos no redundante derivada de M5NR que integra otras bases de datos como GenBank, SEED, IMG, UniProt, KEGG, eggNOG y bases de datos ribosomales que incluyen, SILVA (Pruesse et al., 2007), Greengenes (De Santis et al., 2006) y RDP (Garrity et al., 2006).

El servidor provee de varios métodos para acceder a los distintos tipos de datos, incluyendo la reconstrucción filogenética, que junto con la clasificación funcional permite la descripción de rutas metabólicas para los metagenomas.

La visualización de los resultados ya sea de anotación de categorías funcionales o análisis filogenéticos mediante distintos tipos de gráficos o *heatmap*, resulta muy amigable para el usuario.

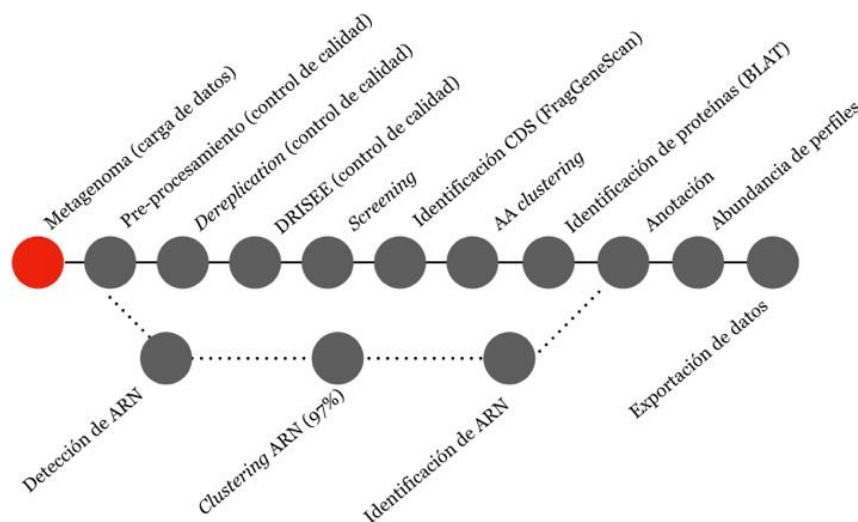


Figura 3.1: Detalle del proceso que sigue un análisis empleando MG-RAST.

ii) Programa Integron Finder para la identificación y anotación de integrones

Integron Finder es un programa desarrollado por Cury y colaboradores (Cury et al. 2016) para la identificación de integrones y sus elementos más característicos: el gen que codifica para la enzima integrón integrasa (IntI) y los GCs con sus sitios *attC* asociados. Este programa permite la identificación, caracterización y estudio de diversidad de cualquier clase de integrones en genomas bacterianos.

La detección de los genes *intI* se realiza empleando los perfiles del modelo oculto de Markov (*Hidden Markov Model*). Para la detección de los sitios *attC* se emplea un modelo probabilístico, el modelo de covarianza que resulta adecuado para modelar elementos que tienen una estructura muy conservada pero son variables a nivel de secuencia. Además Integron Finder identifica los sitios *attI*, los promotores que dirigen la expresión de la integrasa (Pint) y el/los promotor(es) de los GCs (Pc.)

Este programa es de acceso público y puede ser usado a través del servidor del usuario (http://mobyli.pasteur.fr/cgi-bin/portal.py#forms::integron_finder) o como aplicación independiente, para proyectos genómicos o metagenómicos a gran escala. Para desarrollar este programa los autores construyeron un perfil de proteínas a partir de 402 IntI homólogos (Cambray, Sanchez-Alberola, et al., 2011a), con la región específica de las integrón tirosin recombinasas, al que denominaron intICterm (Cury et al. 2016). Por otro lado, el perfil genérico de las tirosin recombinasas fue extraído de la base de datos Pfam (<http://pfam.xfam.org/>).

En la Figura 3.2 se muestra el diagrama con las distintas etapas en el proceso de identificación y anotación de integrones por el programa Integron Finder.

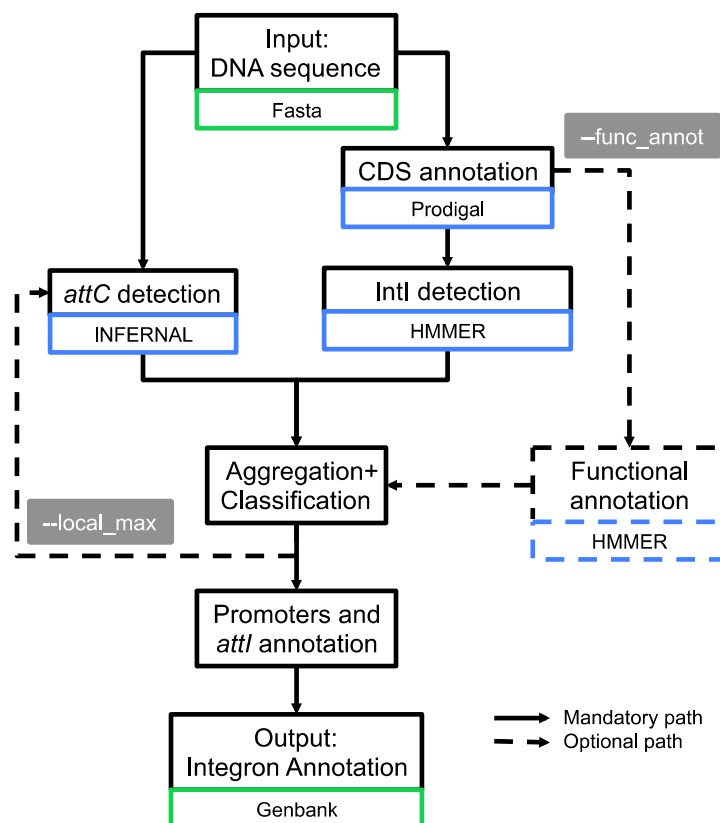


Figura 3.2: Diagrama que describe los pasos empleados por Integron-Finder para identificar y anotar integrones. (Las líneas negras continuas indican los pasos pre-determinados, las líneas punteadas son pasos opcionales). Tomado de Cury et al., 2016.

El archivo de entrada de Integron Finder es una secuencia de ADN en formato fasta. La anotación de la secuencia se realiza con Prodigal v.2.6.2. Se destaca la ventaja que tiene la etapa de anotación especialmente cuando se trata de secuencias nuevas o que están pobremente anotadas. A continuación, el programa busca los perfiles de las integrón integrasas con *hmm search* (HMMER version 3.1bi) y los sitios *attC* con el modo predeterminado de *cmsearch* de INFERNAL v 1.1. (Nawrocki et al., 2013).

Los datos de salida son clasificados en tres categorías:

i) cuando un elemento contiene el gen *intI* y al menos un sitio *attC* se denomina integrón completo. En este caso, la búsqueda se realiza entre la secuencia de *IntI* y a una distancia de 4 kb desde el sitio *attC* más distante o hasta que no encuentra más sitios *attC*.

ii) si el elemento posee el gen *intI* pero no tiene sitios *attC* asociados se denomina Ino. Nótese que Ino no refiere necesariamente a la definición original de integrón, la cual incluye el sitio *attI* además del gen *intI*. La secuencia *attI* es desconocida para la mayoría de las clases de integrones, por lo que resultaría muy difícil realizar su búsqueda. En este caso, el programa busca los sitios *attC* en las dos hebras y hasta 4 kb de distancia desde la *IntI*. Luego continúa buscando los promotores y los sitios *attI* próximos a la integrasa.

iii) cuando un elemento contiene al menos dos sitios *attC* pero carece de la enzima integrón integrasa asociada se denomina CALIN (CAssette Lacking INtegrase). En este caso, la búsqueda de los sitios *attC* se realiza en la misma hebra pero en ambas direcciones.

Por último, existe la opción de “anotar función” para las secuencias codificantes que se encuentran entre el gen *intI* y 200 pb después del último sitio *attC*. En el caso de los CALIN, la búsqueda comprende 200 pb antes del primer sitio *attC* y 200 pb después del último *attC*. En este caso se usa la base de datos de perfiles de proteínas. El usuario puede elegir cualquier otra base de datos compatible con perfiles HMMs.

iii) Anotación funcional de genes con eggNOG mapper

Para lograr una anotación funcional precisa y consistente, es imprescindible generar una interpretación correcta de los datos genómicos obtenidos. La asignación de funciones a genes nuevos presenta la dificultad de distinguir entre genes ortólogos y genes parálogos (Huerta-Cepas et al., 2017).

Los genes ortólogos son el resultado de eventos de especiación luego de los cuales generalmente conservan su función original. Los genes parálogos, que se originan por eventos de duplicación, quedarán probablemente sometidos a las fuerzas de la presión selectiva. Esta presunción, conocida como “Conjetura Ortóloga”, representa la base de los métodos de anotación de funciones (Tatusov et al, 1997). Como los genes ortólogos poseen la misma función, es posible inferir las funciones entre miembros del grupo COG (*Cluster Orthologs Genes*), lo que facilita la predicción de funciones para genomas que están pobremente caracterizados (Tatusov et al., 1997).

La base de datos eggNOG (*evolutionary genealogy of genes Non-supervised Orthologs groups*) es una base de acceso público y provee anotaciones funcionales para grupos ortólogos (OG) de proteínas a diferentes niveles taxonómicos.

La aplicación eggNOG mapper (Huerta-Cepas et al., 2017) fue diseñada para generar una rápida anotación funcional de secuencias nuevas. La misma permite la anotación de grandes colecciones de secuencias codificantes en genomas, metagenomas o transcriptomas. En la Figura 3.3 se describe brevemente en forma esquemática el funcionamiento de la aplicación eggNOG mapper.

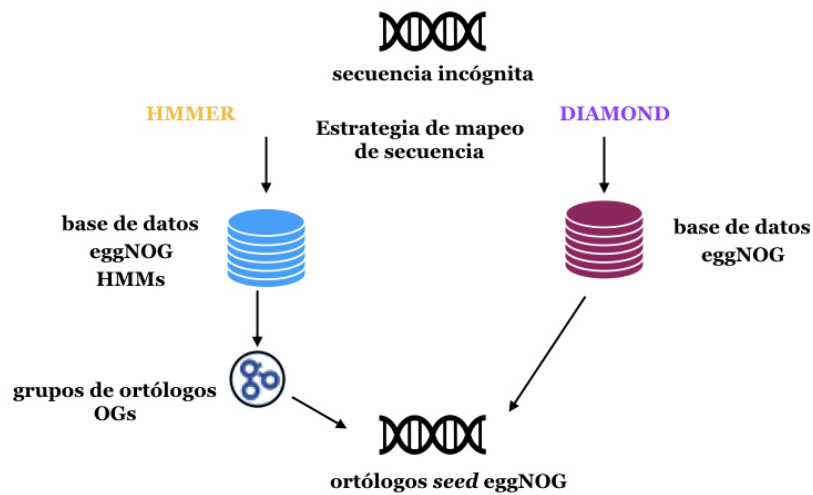


Figura 3.3: Representación gráfica del flujo de trabajo de eggNOG-mapper.

1) La etapa inicial abarca el mapeo de la secuencia con dos opciones de búsqueda, una basada en los modelos ocultos de Markov (HMMs) y la otra en DIAMOND (*Double Index Alignment of Next-generation sequencing Data*) (Buchfink et al., 2014). En primer lugar se compara la secuencia incógnita (*query*) contra la colección de HMMs de la base de datos eggNOG (Huerta-Cepas et al., 2016), para tener una aproximación inicial de la anotación funcional de los *hits*. Cada secuencia *query* se compara contra las proteínas eggNOG representadas por el mejor *hit* HMM. La mejor secuencia para cada *query* es guardada en la base *seed* de ortólogos y empleada para recuperar otros ortólogos. Cuando se maneja un gran set de datos como los metagenomas, es adecuado emplear herramientas más rápidas para el mapeo. En este caso se recomienda el uso de DIAMOND que busca el mejor *hit* (*seed* ortólogo) de cada secuencia directamente entre todas las proteínas de eggNOG. En ambos casos se llega al mejor *hit* de ortólogos *seed* en eggNOG.

2) En la etapa de asignación de ortólogos, se realiza una búsqueda más minuciosa como forma de descartar secuencias con baja homología. Para cada secuencia *query* se recupera el mejor *hit* que coincide con una proteína de la base de datos eggNOG. Éste a su vez es usado para rastrear dentro de una lista de ortólogos incluidos en una base de árboles filogenéticos de eggNOG previamente

analizados. Los ortólogos taxonómicamente distantes son automáticamente descartados.

3) Todos los descriptores funcionales disponibles para los ortólogos recuperados son transferidos a las correspondientes proteínas *query*. De forma predeterminada, la asignación funcional está automáticamente restringida a los ortólogos taxonómicamente más cercanos para cada secuencia incógnita, ya sea utilizando ortólogos individuales o todos los disponibles.

La descripción de las funciones se basa en la versión más reciente de eggNOG que actualmente incluye la base curada de Gene Ontology (GO) (Gene Ontology Consortium 2015), KEGG (rutas metabólicas) (Kanehisa et al., 2014) y las categorías funcionales COG (Galperin et al., 2015). Además también se predice para cada secuencia el nombre de la familia de genes.

Al presente, la colección HMM de eggNOG incluye 1.911.745 perfiles de secuencias OGs, con 1678 bacterias, 115 archaeae, 238 eucariotas y 352 virus (Huerta-Cepas et al., 2017).

Objetivo General

Identificar y analizar elementos de transferencia horizontal genética en comunidades bacterianas que se desarrollan en distintos sitios de la Antártida marítima y continental.

Objetivo Específico

Analizar la diversidad de genes integrasa (*intI*) y GCs asociados mediante técnicas independientes de cultivo.

3.2- Construcción de bibliotecas de amplicones *intI*

Al momento de comenzar este trabajo, no había en la literatura reportes de integrones en comunidades microbianas del ecosistema antártico, excepto un estudio de Stokes y col. (2001), en donde identificaron GCs de integrones a partir de ADN total de muestras de suelo ornitogénico, en Cabo Deninson (Antártida continental) (Stokes et al., 2001).

En nuestro estudio se empleó un par de cebadores diseñado por Mazel y col., (2000) para amplificar integrones clase 1 (*intI1F/intI2R*) en aislamientos “clínicos”. Sin embargo, en un estudio independiente de cultivo utilizando estos mismos cebadores, Nemergut y col., (2004) identificaron doce nuevos genes integrasa *intI* en muestras de relaves mineros contaminados con metales pesados, además de recuperar los clásicos integrones clase 1 (Nemergut et al., 2004). Estas secuencias exhibieron una gran diversidad en comparación con integrasas de clase 1 (*IntI1*), presentando menos del 90% de similitud a nivel aminoacídico. Los resultados obtenidos en este trabajo sugieren que, empleando los mismos cebadores, mediante un abordaje independiente del cultivo es posible detectar una diversidad de genes *intI* mayor que la recuperada a partir de organismos cultivables.

Con el propósito de determinar la presencia y diversidad de integrones clase 1 en un ambiente aislado geográficamente, donde el impacto por la actividad humana es mínimo, se construyeron bibliotecas de amplicones de fragmentos del gen *intI* a partir de ADN total aislado de muestras de sedimento/suelo y matas microbianas de la Isla Rey Jorge (Antártida marítima).

3.2.1-Materiales y Métodos

Las muestras fueron colectadas en tubos estériles de 50 mL y almacenadas a -20°C hasta su procesamiento en el laboratorio (IIBCE, Montevideo), excepto

durante su traslado, que fueron mantenidas refrigeradas en hielo. Las muestras consistieron en matas microbianas asociadas a briofitas (HP19), sedimento/suelo asociado a briofitas (CN11) y sedimento/suelo (IRJ, CP2 y PD2b) (Tabla 3.1).

Se hace notar que la biblioteca IRJ se construyó a partir de una muestra colectada en un sitio de la Isla Rey Jorge para el cual no se dispone de datos de coordenadas geográficas. Esta muestra está identificada en la Tabla suplementaria S1 como CH39.

Tabla 3.1: Muestras empleadas para el análisis metagenómico de genes *intI* mediante biblioteca de amplicones.

Muestra	Campaña	Coordenadas geográficas	Sitio	Tipo de muestra
HP	Mar/2008	S 62 13 29 W 58 57 9	Halfthree-Point	Mata marrón agua de deshielo estancada
CN	Ene/2013	S 62 11 13,6 W 58 54 48,8	Caleta Norma	Sedimento asociado a briofitas
PD2b	Abril/2015	S 62 09 03 W 58 56 27	Pasaje Drake	Sedimento
IRJ	2015	s/d	Isla Rey Jorge	Sedimento
CP2	Febrero/2016	S 62 14 35 W 58 39 99	Caleta Potter	Sedimento

i) Purificación de ácidos nucleicos totales

La extracción de ADN de las muestras seleccionadas para análisis se realizó con el kit ZR Soil Microbe DNA Microprep M (Zymo Research[®], EUA) siguiendo las instrucciones del fabricante.

El ADN purificado se utilizó como molde en reacciones de PCR para la amplificación de un fragmento del gen *intI1* (484 pb) (Mazel et al., 2000). El protocolo empleado para este ensayo es el mismo que está descrito en Antelo y col., 2015 (Antelo et al., 2015).

ii) Construcción de bibliotecas de amplicones *intI*

Para la construcción de bibliotecas se emplearon los productos de PCR del tamaño esperado. Las mezclas de reacción de PCR fueron analizadas mediante electroforesis en gel de agarosa para visualizar los fragmentos amplificados. Los productos se purificaron empleando el kit Gene JET Extraction Kit (Molecular Biology, Thermo Scientific[®], Lituania), siguiendo las instrucciones del fabricante. El ADN purificado fue clonado en el vector pCR[®]2.1 del kit de clonado TOPO-TA cloning kit (Invitrogen, EUA.). Los productos de clonado obtenidos fueron transformados en células químicamente competentes de *Escherichia coli* TOP10 (Invitrogen, EUA.). Las células transformadas fueron incubadas *overnight* a 37°C y sembradas en placas con LB Km 50 µg/mL y Xgal 20 µg/mL. Se seleccionaron las colonias blancas para la purificación plasmídica. Esta se realizó a partir de cultivos en 5 mL de LB Km 50 µg/mL (Sambrook et al., 1989) incubados *overnight*. Se verificó el tamaño correcto del inserto mediante electroforesis en gel de agarosa, previa digestión del plásmido con la enzima *EcoRI*. El ADN plasmídico fue enviado a MACROGEN (Seúl, Corea) para secuenciación usando el primer intIF (Mazel et al. 2000).

iii) Análisis de secuencias de ADN

Las secuencias obtenidas se analizaron con el programa VecScreen a efectos de identificar y remover las lecturas del vector (<http://www.ncbi.nlm.nih.gov/tools/vecsreen/>).

Las secuencias fueron analizadas con el programa BLAST en la base de datos de proteínas no redundantes del NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) (Zhang et al., 2000). Para el análisis también se construyó una base de datos con metagenomas de comunidades microbianas de sedimentos marinos contaminados de cuatro sitios de la Isla Rey Jorge (IDs 3300000129, 3300000132, 3300000119, and 3300000123) (<https://img.jgi.doe.gov/cgi-bin/m/main.cgi>). El valor de corte de e-value fue de 10^{-5} y 100 aminoácidos la cobertura mínima de nuestras secuencias IntI.

A efectos de identificar los motivos conservados entre las integrón-integrasas (IntI) (Nunes-Düby et al., 1998), las secuencias aminoacídicas de cada biblioteca fueron alineadas con secuencias de integrasas de referencia y tirosin recombinasas XerC y XerD (Nº de acceso: IntI, AAQ16665; IntI2, AAT72891; IntI3, AAO32355; XerC *E. coli*, PoA8P6; XerD *E. coli*, PoA8P8; XerC *Thiobacillus denitrificans*, 499632306; XerD *T. denitrificans*, 516743720).

Para el alineamiento se empleó el programa Mafft versión 3.717 (Katoh & Standley, 2013). La visualización y edición se realizó con el programa Aliview versión 1.20 (Larsson, 2014).

iv) Análisis filogenético

Se realizó un análisis filogenético basado en el alineamiento de las secuencias aminoacídicas deducidas del fragmento del gen *intI*, de las bibliotecas HP y CN y los mejores *hits* obtenidos con el programa BLAST seleccionados de la base de datos del NCBI y de IMG/G (Integrated Microbial Genomes and Metagenomes). Para el análisis se empleó el método de Máxima Verosimilitud en el servidor <http://phylogeny.fr> (Dereeper et al., 2008) con los parámetros predeterminados del programa PhyML.

También, empleando el mismo método, se construyó un árbol que incluyó las secuencias IntI identificadas en los genomas completos de la base de datos DOE JGI, además de las IntI antárticas y las IntI de referencia.

3.2.2 Resultados

Se construyeron dos bibliotecas de genes *intI* a partir de muestras de Halfthree-Point (HP) (2008) y Caleta Norma (CN) (2013). Se obtuvieron 98 secuencias de alta calidad, 62 de HP y 36 de CN. El análisis de secuencia de ADN (Blastn) indicó que 42 clones de la biblioteca HP eran muy similares (99-100% identidad) a integrones de clase 1 (gen *intI1*). Las secuencias de los 42 insertos eran idénticas al del gen integrón integrasa identificado en cinco aislamientos de *Enterobacter* sp. Trim^r, Amp^r, Strp^r recuperados de muestras de sedimento/suelo de

diferentes sitios de Península Fildes (Antelo et al. 2018). Los 42 fragmentos del gen *intI* mostraron 100% de identidad (Blastn) con un fragmento del gen *intI* de una cepa de *Pseudomonas putida* DLL- E4 (CP007620.1), *K. pneumoniae* Kp15 plásmido pENVA (HG918041) y *Xanthomonas oryzae* pv. *oryzae* cepa YNA12-2 (HQ662557) de origen clínico y ambiental.

Las restantes secuencias *intI* de HP y las 36 secuencias de CN exhibieron distintos niveles de similitud (Blastx) con secuencias IntI obtenidas en análisis metagenómicos de muestras ambientales de orígenes diversos (Nield et al. 2001, Elsaied et al. 2007, Elsaied et al. 2011, Rodríguez-Minguela et al. 2009).

i) Análisis filogenético de secuencias IntI de las bibliotecas HP y CN



Figura 3.4: Árbol filogenético de secuencias IntI obtenido de las bibliotecas HP19 y CN11 y los mejores *hits* seleccionados del Genbank y de la base de datos IMG/M. El árbol se construyó empleando el servidor (<http://www.atgc-montpellier.fr/phym/>) con el programa PhyML y las opciones predeterminadas. Las secuencias de HP se indican en azul y las de CN en rojo. Las secuencias de IntI conocidas (clase 1 a clase 8) se muestran en rosado y las secuencias IntI de aislamientos de enterobacterias previamente descritos se muestran en verde.

ii) Obtención de bibliotecas de amplicones *intI* de IRJ, PD2 y CP2

Posteriormente, se construyeron tres bibliotecas de IntI de sedimento/suelo de Pasaje Drake (PD2), Caleta Potter (CP2) e Isla Rey Jorge (IRJ). Se obtuvieron un total de 36 secuencias IntI de calidad, 8 IntI de PD2, 14 IntI de IRJ y 14 IntI de CP2.

Los resultados del análisis filogenético realizado con las secuencias IntI de las bibliotecas PD2, CP2 e IRJ-SD se describen en la Sección 3.3.2.-iii de este capítulo.

3.3-Análisis taxonómico y funcional *in silico* de metagenomas de tapetes microbianos y sedimento/suelo antárticos. Identificación de genes integrasas intI y GCs asociados.

En este trabajo empleamos metagenómica *shotgun* y herramientas bioinformáticas para analizar la presencia de integrones y GCs asociados en comunidades microbianas de la Antártida. El estudio incluyó el análisis de catorce metagenomas de tapetes microbianos y tres metagenomas de sedimento/suelo colectados en distintos sitios de la Antártida marítima y continental (Tabla S2, Material Suplementario).

3.3.1-Materiales y Métodos

i) Sitios de muestreo

En el marco del proyecto del Fondo Conjunto de Cooperación México-Uruguay: “Respuestas bióticas al cambio ambiental global: Evidencias de los ecosistemas terrestres Antárticos y sus servicios ambientales”, se obtuvieron catorce metagenomas totales de tapetes microbianos. Las muestras fueron colectadas de distintos sitios de la Antártida marítima y continental, representando un gradiente parcialmente latitudinal. Este trabajo fue realizado por dos integrantes del Instituto de Ecología de la UNAM (Universidad Nacional Autónoma de México), Patricia Valdespino y Luisa Falcón. Los metagenomas fueron secuenciados en el Centro Yale para Análisis Genómico (YCGA) empleando la tecnología Illumina HiSeq4000.

Asimismo el análisis incluyó metagenomas de sedimento/suelo colectados en Península Fildes (Isla Rey Jorge, Antártida marítima), en la campaña de abril de 2015. Los sitios de muestreo incluyeron Isla Ardley (IA6), Halfthree-Point (HTP3) y la cámara séptica de la BCAA (CS1).

La colecta y procesamiento de las muestras fue realizada por Matías Giménez, integrante de la Unidad de Microbiología Molecular (IIBCE), en el marco de su Tesis de Maestría. La secuenciación de estos metagenomas se realizó en MacroGen (Seúl, Corea del Sur), empleando la tecnología Illumina HiSeq 2500.

El conjunto de contigs obtenidos fue empleado para la búsqueda de genes integrasa *intI*.

ii) Búsqueda de genes integrón integrasa (*intI*) en metagenomas de tapetes microbianos y sedimento/suelo completamente secuenciados:

El análisis de metagenomas antárticos se realizó en varias etapas. El primer paso consistió en usar el conjunto de metagenomas como base de datos y las secuencias aminoacídicas deducidas de los genes integrasa de referencia como *query* para la búsqueda de integrones en cada metagenoma. Este análisis se llevó a cabo empleando los algoritmos Blastn y Blastx del programa BLAST (Zhang et al., 2000).

Los contigs mayores a 2000 pb que contenían genes *intI* fueron enviados al servidor RAST (Rapid Annotation using Subsystem Technology) para su anotación (Aziz et al., 2008).

En una segunda etapa, el análisis se realizó empleando el programa Integron Finder (<https://github.com/gem-pasteur/Integron-Finder>) (Cury et al. 2016) para la detección de integrones y sus componentes más distintivos: la enzima integrón-integrasa IntI, los sitios *attI*, *attC*, los promotores *P_{int}*, *P_c* y los GCs.

iii) Anotación de GCs de resistencia a antibióticos (Base de datos Resfam)

El programa Integron Finder emplea la base de datos Resfams para la anotación de GCs de resistencia a antibióticos asociados a integrones.

Resfams es una base de datos de referencia, específica de familia de proteínas y perfiles asociados a modelos ocultos de Markov (HMMs), organizados por ontología y validada para función de resistencia a antibióticos (Gibson et al., 2015)(<http://www.dantaslab.org/resfams>). Esta base recoge 166 perfiles HMMs que representan la mayoría de las clases de genes de resistencia a antibióticos, incluyendo genes de resistencia contra beta-lactamasas, aminoglucósidos, fluoroquinolonas, glicopéptidos, macrólidos, tetraciclina, bombas de eflujo y factores moduladores de la transcripción de resistencia antibiótica.

iv) Análisis filogenético de integrasas antárticas con secuencias IntI de referencia

El análisis filogenético incluyó 49 secuencias IntI identificadas por el programa Integron Finder en el *dataset* de los metagenomas antárticos analizados (Capítulo 3, Sección 3.3.1.ii) y 81 secuencias obtenidas a partir de las bibliotecas de amplicones *intI* (Capítulo 3, Sección 3.2.2). El análisis también incluyó 45 secuencias de integrasas de organismos conocidos y las secuencias XerC y XerD de *E. coli* y *Thiobacillus denitrificans*, que se emplearon como ramas externas para la construcción del árbol. Las secuencias de referencia fueron tomadas de la base de datos INTEGRALL The Integron Database (Moura et al., 2009) (<http://integrall.bio.ua.pt> y del NCBI GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>).

Para identificar los motivos conservados característicos de las integrasas, las secuencias aminoacídicas seleccionadas para este estudio fueron alineadas contra secuencias *intI* de referencia (clase 1, 2, 3, 4 y 5) y tirosín recombinasas XerC y XerD (Nº de acceso al GenBank: IntI1, AAQ16665; IntI2, AAT72891; IntI3, AAO32355; IntI4, AAC38424; IntI5, AF180939; XerC *E. coli*, POA8P6; XerD *E. coli*, POA8P8; WP011313040, XerC *T. denitrificans*; WP018078134; XerD *T. denitrificans* (Nunes-Düby et al., 1998).

Para el alineamiento se empleó el programa Mafft versión 3.717 (Katoh & Standley, 2013). Para la visualización y edición se usó el programa Aliview versión 1.20 (Larsson, 2014).

Se incluyó también como referencia Integrasas representativas de especies que portan integrones cromosómicos y super-integrones.

El análisis fue realizado con el método de Máxima Verosimilitud en base al mejor modelo de evolución de proteínas (LG+G) seleccionado por el servidor ATGC Montpellier Bioinformatic Plattform (<http://www.atgc-montpellier.fr/phyml/>) con las opciones predeterminadas en el programa PhyML v 3.0 (Guindon et al., 2010).

v) Contexto genómico de los integrones completos identificados por el programa Integron-Finder

El programa I-F provee los resultados como archivos en formato GenBank que describe todos los elementos genéticos asociados con un integrón. Estos archivos

pueden emplearse como entrada en determinados programas que permiten una visión interactiva de todos los *hits* encontrados. En este caso, los archivos GenBank que contenían integrones completos fueron anotados con el programa SnapGene® Viewer 4.2.11 para visualizar el contexto genómico del integrón.

vi) Anotación funcional genómica de GCs mediante la asignación de ortología con eggNOG-mapper

Las secuencias de proteínas obtenidas a partir del programa Integron Finder fueron anotadas contra la base de datos eggNOG-mapper (v.5.1) (Huerta-Cepas et al., 2016) <http://eggnogdb.embl.de/#/app/emapper>. La base de datos eggNOG 4.5 provee de alineamientos múltiples y anotación funcional de los genes candidatos. Las secuencias anotadas se agrupan en 20 categorías funcionales.

vii) Composición taxonómica de la comunidad en metagenomas antárticos

Para determinar la composición taxonómica de la comunidad, los metagenomas fueron anotados en el servidor MG-RAST (Meyer et al., 2008) empleando los parámetros de control predeterminados. Para la distribución taxonómica de los *hits*, MG-RAST emplea el algoritmo contig LCA (*lowest common ancestor*). Este algoritmo busca una única entidad taxonómica consenso en cada secuencia. Para la anotación comienza por el nivel taxonómico más bajo con un valor mínimo de confianza definido, por ejemplo si no puede anotar a nivel de especie pasa al nivel inmediato superior y buscará anotar a nivel de género.

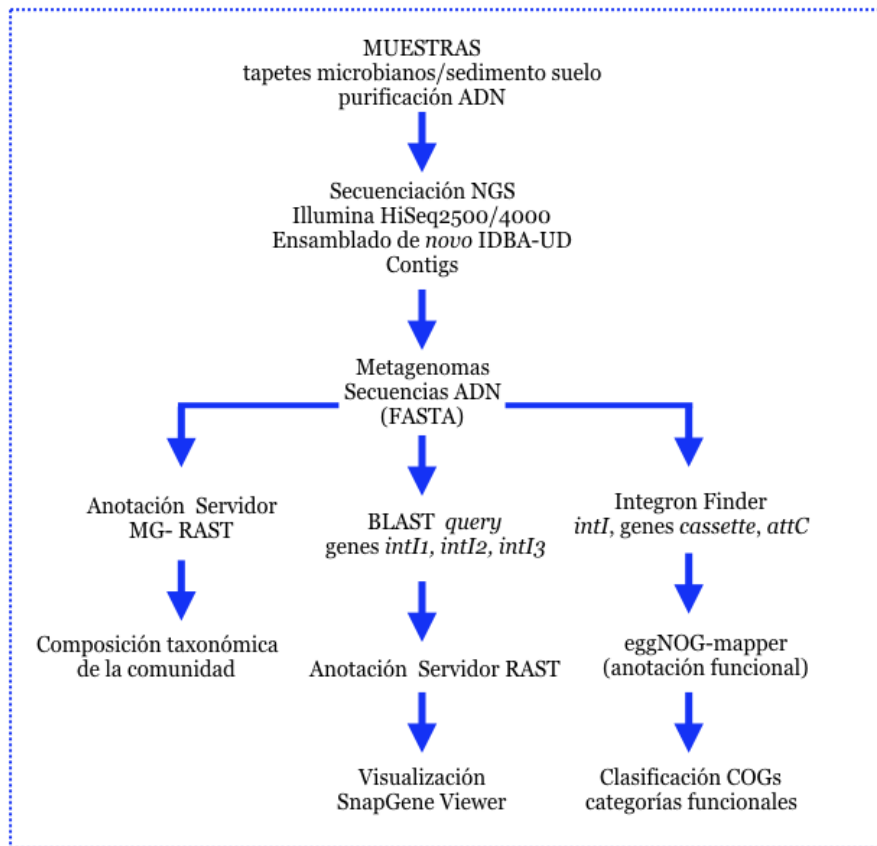


Figura 3.5 :Diagrama que describe los pasos implementados para el análisis del *dataset* de metagenomas antárticos.

3.3.2. Resultados

i) Identificación de genes que codifican para integrasas de integrones (IntI) en metagenomas antárticos

Los sitios de muestreo incluyeron zonas con marcada influencia humana como las zonas próximas a las bases de operaciones, sitios donde se establecen los asentamientos de animales y otros lugares que forman parte de las zonas especialmente protegidas (zonas ASPA) presuntamente prístinas. Asimismo, el análisis incluyó distintos tipos de muestra (tapetes microbianos y sedimento/suelo). En la Tabla 3.2 se resumen los resultados obtenidos con I-F para el análisis del *dataset* de metagenomas. Se obtuvieron un total de 3431 secuencias codificantes, que incluyeron 53 secuencias de enzimas integrón integrasas (IntI). De las 53 IntI encontradas, 34 fueron clasificados como integrones completos que constan del gen *intI* y al menos un sitio *attC*. Las 19 IntI restantes cayeron dentro del grupo Ino (integrón vacío) y el resto de las CDS incluye el grupo CALIN (GCs sin la integrasa asociada) y otros ORFs que carecen del sitio *attC* por lo que no es posible clasificarlos como GCs.

Tabla 3.2: Resultados obtenidos con el programa I-F en el *dataset* de metagenomas antárticos (la información de los sitios se describe en la tabla S2).

Sitio	Nº contigs	Nº integrones completos	Nº Ino	Nº CALIN
Drake (Metagenoma 1)	14345	0	0	42
Espejo (Metagenoma 2)	11311	0	1	149
Halfthree-point (Metagenoma 3)	10227	0	0	53
Pista (Metagenoma 4)	19554	0	2	43
Base Carlini I (Metagenoma 5)	62551	3	4	119
Base Carlini II (Metagenoma 6)	47446	1	0	143

Esperanza (Metagenoma 7)	35849	0	1	104
Primavera (Metagenoma 8)	53174	2	1	114
San Martín (Metagenoma 9)	17447	1	1	44
ANCO10 (Metagenoma 10)	18089	3	2	205
ANCO11 (Metagenoma 11)	24919	1	0	94
ANCO 12 (Metagenoma 12)	23874	6	1	284
MidGood13 (Metagenoma 13)	23997	4	1	145
ANCO14 (Metagenoma 14)	34862	1	0	70
CS1	68030	9	3	283
H3P	28119	1	2	85
IA6	54969	2	0	103

ii) Anotación de GCs de resistencia a antibióticos (Base de datos Resfam)

Se detectaron genes de resistencia a antibióticos en gran parte de los metagenomas investigados (70,58%).

El programa Integron Finder identificó un total de 26 genes *cassettes* que correspondieron a 12 metagenomas y se distribuyeron en 18 familias de genes. El resultado se resume en la Tabla suplementaria S6).

Sin embargo, todos los GCs de resistencia detectados cayeron dentro del grupo CALIN (Gene CAssette Lacking Integrase), con excepción de un gen de resistencia β -lactamasa de clase D (Resfam:RF0056).

Este GC forma parte del integrón CS1_503 (cámara séptica de la BCAA) y se describe en la sección 3.3.2.iv de este capítulo.

iii) Análisis filogenético de IntI antárticas con IntI de referencia

El alineamiento de las secuencias aminoacídicas antárticas con secuencias integrón integrasas representativas permitió identificar los motivos conservados característicos de la familia de las tirosín recombinasas. Estos motivos incluyen cuatro residuos altamente conservados (RHRY) y las regiones Box I, Box II, Parches I, II y III (Nunes-Düby et al., 1998). Además se identificó el dominio adicional de alrededor de 35 aa, específico de las enzimas integrón integrasas (IntI) (Nunes-Düby et al., 1998) (Messier & Roy, 2001). También los residuos aminoacídicos esenciales para la unión del ADN y para los procesos de recombinación están presentes en la mayoría de las secuencias, sugiriendo que representan enzimas funcionales (Messier & Roy, 2001).

Por otra parte, encontramos que tres secuencias IntI de metagenomas de sedimento de la cámara séptica de la BCAA (CS1_12849, CS1_2689, CS1_503) y cuatro de tapetes microbianos (MTG5_28909, MTG8_2561, MTG10_95, yMTG14_69269) poseen un dominio adicional extendido (variable en largo) que se localiza a continuación del Parche III. Este tipo de inserción fue identificada previamente por Rodríguez-Minguela y col., (2009) en integrones ambientales recuperados de regiones geográficas diversas. Sin embargo, se desconoce la implicancia que pueda tener este tipo de inserción en el funcionamiento y actividad de la enzima integrasa (Rodríguez-Minguela et al., 2009).

Con el fin de determinar la diversidad de proteínas integrón integrasas IntI recuperadas a partir de muestras de tapetes microbianos y sedimento/suelo empleando técnicas independientes de cultivo, se realizó un análisis filogenético con el método de máxima verosimilitud (Figura 3.6).

El alineamiento incluyó 129 secuencias IntI de la Antártida marítima y Antártida continental (74 secuencias de sedimento/suelo y 55 de tapetes microbianos) y 49 secuencias IntI de organismos conocidos, incluidas como referencia. Las secuencias obtenidas en nuestro estudio se identifican con letras mayúsculas que indican el origen de la muestra, un número ordinal propio de la secuencia, seguido por el año de colecta.

En la Tabla 3.2 se resume, para las muestras usadas en el estudio filogenético, el tipo de muestra de origen y la estrategia aplicada para su obtención. Las técnicas usadas fueron: construcción de bibliotecas de amplicones (PCR *intI*) y posterior análisis de las secuencias generadas por el método Sanger y análisis de los datos de secuenciación *shotgun* (NGS, Next Generation Sequencing) con tecnología Illumina.

Tabla 3.2: Se describe para cada caso, el tipo de muestra y la estrategia aplicada. MTG1-14: refiere a los metagenomas de la Antártida marítima del N°1 al N°9 y del continente antártico (ANCO) numerados del N°10 al N°14, descritos en la Tabla S2 (Material suplementario).

Sitio	Colecta	Origen	Estrategia aplicada	Código de colores ***
HTP (Halfthree-Point)	2008	Tapetes	PCR(<i>intI</i>)	
CN (Caleta Norma)	2013	Sedimento/suelo	PCR(<i>intI</i>)	
PD2 (Pasaje Drake)	2015	Sedimento/suelo	PCR(<i>intI</i>)	
IRJ (Isla Rey Jorge) *	2015	Sedimento/suelo	PCR(<i>intI</i>)	
CP2 (Caleta Potter)	2016	Sedimento/suelo	PCR(<i>intI</i>)	
CS1 (Cámapra Séptica-BCAA)	2017	Sedimento/suelo	NGS	
HP3 (Halfthree-Point)	2017	Sedimento/suelo	NGS	
IA6 (Isla Ardley)	2017	Sedimento/suelo	NGS	
MTG1-14 **	2015	Tapetes	NGS	

* sin datos de coordenadas geográficas

** incluye metagenomas de la Antártida marítima y continental

*** código de colores de secuencias *IntI* que integran el árbol filogenético.

IA12: corresponde a un aislamiento de *Enterobacter* sp. obtenido a partir de muestras de sedimento con fecas de pingüino en Isla Ardley.

Las secuencias integradas públicas empleadas como referencia están escritas en color negro.

Los grupos que contienen las secuencias de referencia de integrasas móviles clase 1 (*IntI1*), clase 2 (*IntI2*), clase 3 (*IntI3*), las *IntI* del género *Vibrio* sp. (clases 4 y

5, agrupadas), y nuestras secuencias más relacionadas están resaltados en color amarillo. Para facilitar la visualización del árbol, los grupos que incluyen sólo las IntI antárticas se resaltaron en dos tonos de gris.

Los diferentes grupos están señalados con números romanos y las ramas que soportan una única secuencia IntI se señalan con letras (A-G).

Las secuencias IntI identificadas en este estudio exhibieron una gran diversidad, distribuyéndose a lo largo del árbol en 28 grupos diferentes. Asimismo, varias de nuestras IntI no están relacionadas filogenéticamente con secuencias integrón integrasas conocidas y se agrupan formando clados separados.

En concordancia con lo descrito en la literatura, el análisis mostró que todas las secuencias integrón integrasas IntI forman un grupo claramente separado de las recombinasas XerC y XerD, las más cercanas dentro de la familia de la tirosín-recombinasas (Nunes-Düby et al., 1998). Además, las integrasas de clase 1 (IntI1) (Grupo XIII) y clase 3 (IntI3) (Grupo XVIII) están cercanamente emparentadas entre sí y se ubican distantes de integrasas de clase 2 (IntI2) (Grupo II), las cuales forman un grupo aparte. Asimismo, las integrasas del género *Vibrio* constituyen un grupo separado del resto de integrasas IntI (Rowe-Magnus et al., 2001).

De un total de 129 secuencias integrón integrasas identificadas, únicamente 19 están relacionadas con IntI móviles (IntI1, IntI2 e IntI3).

IntI clase 1: El grupo XIII incluyó, junto con una integrasa de referencia clase 1, ocho IntI de sedimento de Caleta Potter (CP2), una IntI de tapetes microbianos de Halfthree-Point (HP157_2008) y una integrasa clase 1 del aislamiento IA2 de *Enterobacter* sp. Todas estas secuencias exhibieron 100% de identidad con la secuencia IntI de la cepa ABNIH6 de *Acinetobacter baumannii* (EMT91539) resistente a colistina (Snitkin et al., 2013).

Asimismo se debe mencionar que la Biblioteca HP_2008 contiene 42 secuencias integrasas clase 1 idénticas a las secuencias IntI1 encontradas en la colección de aislamientos antárticos de *Enterobacter* sp. resistentes a trimetoprim (Antelo et al. 2018). A efectos de facilitar la visualización del árbol solo se muestra una secuencia representativa de las 42 IntI.

Las integrasas de clase 1 parecen ser casi idénticas entre sí, lo cual es consistente con lo reportado por la literatura, que señala la gran similitud que exhiben las integrasas clase 1 de tipo clínico.

En Caleta Potter se ubica una de las Bases científicas de Argentina por lo que se presume que en esa zona existe influencia antropogénica, sin embargo Halfthree-Point forma parte de las zonas especialmente protegidas y por lo tanto es un área de acceso restringido. Sin embargo se puede especular que la fauna presente en esa zona podría ser un factor de dispersión de esta clase de integrones

IntI clase 2: La integrasa clase 2 de *S. sonnei* de origen clínico forma el grupo II junto con dos secuencias IntI de metagenomas de la cámara séptica de la BCAA, IntI_CS1_17527 e IntI_CS1_57033. Las IntI_CS1 se ubican juntas en una rama ligeramente separada de la IntI2 de referencia sugiriendo que existe cierta diferencia con la IntI de *S. sonnei*.

Las dos integrasas de CS1 son muy similares a la secuencia IntI clase 2 de *E. coli* (WP_000071895.1).

IntI clase 3: Se puede considerar que el grupo de integrasas móviles clase 3 (grupo XVIII) se ordena en dos subgrupos, uno de los cuales incluyó las IntI3 de referencia y una integrasa de tapetes microbianos de la Antártida marítima (MTG9_2773). El otro subgrupo incluyó tres secuencias IntI de tapetes del continente y cuatro IntI de sedimento/suelo de sitios próximos a las bases de operaciones científicas (Antártida marítima). Se hace notar que el valor soporte que divide estos dos subgrupos es cero, sugiriendo que existe homogeneidad entre las secuencias.

Dos IntI (CS1_4254 y CS1_2164) son similares a una IntI de *Rhodoferrax antarcticus* (WP_076072338.1), mientras que IntI_CS1_24080 e IntI_CS1_4987, exhibieron similitud con la secuencia IntI de *Acidovorax caeni* WP_054258111.1 y de Comamonadaceae bacterium CG2_30_57_122 (OIP12783.1), respectivamente. Por otra parte, cuatro integrasas de tapetes

exhibieron similitud con diferentes integrasas. IntI_MTG5_31000 e IntI 52806 están relacionadas con la secuencia de *Acidovorax* sp. NO-1 (WP_008905967.1) y de una bacteria no clasificada del Orden Burkholderiales (OYU45514.1) respectivamente, mientras que IntI_MTG9_2773 exhibió identidad con una IntI clase 3 de *Serratia marcescens* (AAL10406.19) y la secuencia IntI_MTG9_4549 con la de *Hydrogenophaga flava* (WP_066260803.1).

IntI clase 4 y 5: Ninguna de las IntI antárticas se ubicaron dentro del grupo de integrasas del género *Vibrio*.

Asimismo, las secuencias IntI que no mostraron relación filogenética con integrasas de referencia formaron 19 grupos únicos, siete de los cuales corresponden a ramas que soportan una única secuencia IntI.

Considerando el origen geográfico, las secuencias IntI se distribuyeron en el árbol de la siguiente forma:

De siete ramas que contienen solo una secuencia IntI (señaladas con letras A-G), tres corresponden a integrasas de tapetes continentales (A, B y G) (metagenomas, secuenciación *shot-gun*) y cuatro a IntI de la región marítima (C, D, E y F) (bibliotecas de amplicones *intI*). De estas últimas, tres son IntI de sedimento (C, D y F) y una IntI de tapetes microbianos del Halfthree-Point (E).

Los doce clados restantes se dividen en tres grupos: IntI de la Antártida marítima, IntI de la Antártida continental y un grupo que incluye IntI de ambas regiones biogeográficas.

Cabe mencionar que la relación que se describe a continuación de IntI antárticas con otras integrasas corresponden al mejor *hit* determinado con el programa Blast empleando el algoritmo Blastp (Zhang et al., 2000).

a) Grupos con IntI de la Antártida marítima:

Grupo IV: agrupa 12 secuencias IntI de tapetes microbianos de Halfthree-Point que fueron recuperadas a partir de la misma biblioteca. Todas estas secuencias son iguales entre sí. Estas secuencias IntI fueron similares a una integrasa de

origen ambiental asignada a *Verrucomicrobia bacterium* (PYI76875.1). El genoma de esta bacteria fue reconstruido a partir de metagenomas de suelo de pradera (EUA) (Crits-Christoph et al., 2018).

Grupo V: incluye dos secuencias, una IntI de sedimento del Halfthree-Point (HP3_2775) (metagenoma, secuenciación *shotgun*), similar a una IntI de *Methylocaldum marinum* sp. (WP_119629835) y una IntI de sedimento de Pasaje Drake (PD2_14) (Biblioteca de amplicones, similar a la secuencia *IntI* de una bacteria no cultivada, ABD62683), identificada en metagenoma de suelo (Eslovenia) (Rodríguez-Minguela, et al., 2009).

Si bien ambas integrasas fueron recuperadas empleando estrategias diferentes, presentan en común el tipo de muestra (sedimento).

Grupo VI: formado por seis secuencias, una IntI de sedimento de Caleta Norma (Biblioteca de amplicones *intI*) relacionada con la secuencia IntI de una especie no identificada de *Verrucomicrobia* (clon IV9), encontrada en metagenoma de suelo (Crits-Christoph et al., 2018) y cinco IntI de tapetes de Halfthree-Point (Biblioteca de amplicones), todas similares a la secuencia IntI de una bacteria no cultivada de metagenoma de suelo (Eslovenia) (ABD62691) (Rodríguez-Minguela et al., 2009).

Grupo VII: incluye 20 secuencias IntI de sedimento de Caleta Norma (biblioteca de amplicones) que forman un grupo claramente separado del resto de integrasas. Estas IntI están relacionadas con integrasas ambientales identificadas en un estudio metagenómico de ambientes con distintos niveles de influencia antropogénica (Rodríguez-Minguela et al., 2009).

Los grupos VI y VII incluyen secuencias IntI de sedimento y tapetes de dos bibliotecas de amplicones (Caleta Norma y Halfthree-Point). Estos grupos junto con la secuencia CN11, comparten el mismo nodo. Analizados en conjunto se observa que las IntI de CN despliegan una mayor diversidad que la exhibida por las IntI de Halfthree-Point que son todas prácticamente iguales.

Grupo IX: incluye siete IntI de sedimento, tres metagenomas de Halfthree-Point (secuenciación *shotgun*) y cuatro de Caleta Potter (Bibliotecas de amplicones).

En este grupo se observa homogeneidad entre las secuencias, todas están relacionadas con la misma secuencia IntI de una bacteria no clasificada del phylum *Chloroflexi* (OGO47220.1), cuyo genoma fue reconstruido a partir de metagenomas de sedimento de un acuífero en EE.UU. (Anantharaman et al., 2016).

Grupo X: este grupo está representado únicamente por dos IntI de sedimentos de Caleta Potter (Biblioteca de amplicones) que exhibieron similitud con la secuencia IntI de una bacteria del género *Caldithrix* sp. (TDI82829) identificada en sedimentos marinos (EUA). Este género incluye bacterias termófilas anaerobias estrictas que han sido aisladas de chimeneas hidrotermales de fondos marinos (Miroshnichenko et al., 2003).

Grupo XI: incluye 20 IntI de sedimento recuperadas de dos bibliotecas de amplicones, 14 secuencias de la biblioteca IRJ y seis de PD2. En este caso, al igual que en el grupo VII, se observa gran homología entre las secuencias. Las 20 IntI son similares a una integrasa de origen ambiental de una bacteria no cultivada, identificada en el metagenoma de suelo agrícola (ACM88792.1) (Rodríguez-Minguela et al., 2009).

Grupo XVII: en este grupo se ubican tres IntI de tapetes microbianos, MTG4_2937, MTG4_58 y MTG5_28909 (metagenomas, secuenciación *shotgun*). Las tres IntI están relacionadas con la misma secuencia integrasa de una bacteria no clasificada del orden Verrucomicrobiales (OYVo3803.1) identificada en metagenomas de algas (Reino Unido).

b) Grupos con IntI de la Antártida continental:

Grupo XIX: este grupo es el único formado exclusivamente por secuencias de la Antártida continental. Las tres IntI de tapetes microbianos (MTG10_224,

MTG10_6152 y MTG12_6634) están relacionadas filogenéticamente con una integrasa de *Methylibium petroleiphilum* (WPO11828296).

c) Grupo con IntI marítimas y continentales:

Grupo VIII: este grupo contiene tres IntI de tapetes continentales: IntI_MTG10_1320 es similar a la secuencia IntI de *Acidobacteria bacterium* (PYS91836.1) (Crits-Christoph et al. ,2018). IntI_MTG12_6209 e IntI_MTG12_8285 presentan similitud con secuencias integrasas de *Chloroflexi bacterium* (RPI92408) y de *Anaerolineae bacterium* CG2_30_57_67 (OIO85441.1), respectivamente.

El Grupo VIII abarca también dos IntI de tapetes marítimos: IntI_MTG5_9187 e IntI_MTG6_12221, relacionadas con una secuencia integrasa de una cianobacteria, *Pseudanabaena* sp. SR411 (WP_094528421.1) y una IntI de sedimento HP3_2933 similar a una secuencia anotada como potencial transposasa de *Chloroflexi bacterium* OLB14 (KXK12064.1).

Grupo XIV: este grupo está consituído por cinco IntI de tapetes microbianos. Tres secuencias provienen de la Antártida continental: MTG11_2380, MTG13_1229 y MTG13_7622 y dos de la región marítima: MTG5_9656 y MTG8_3112.

IntI_MTG11_2380, IntI_MTG13_1229, Int_MTG13_7622 e IntI_MTG8_3112, exhibieron similitud con una secuencia integrasa de *Paucibacter* sp. KCTC 42545 (WP_082680180). IntI_MTG5_9656 está relacionada con una integrasa de *Rhizobacter* sp. Root29 (KQU70252).

Grupo XV: lo constituyen dos IntI de tapetes continentales, IntI_MTG12_6634 e IntI_MTG12_3026 y una IntI de sedimentos IntI_MTG5_26375 (Antártida marítima).

IntI_MTG12_3026 e IntI_MTG5_26375 exhibieron similitud con la secuencia integrasa de una Betaproteobacteria no clasificada (WP_082354447.1). Por otro lado, IntI_MTG12_6634 está relacionada con una integrasa de *Alicyclophilus denitrificans* (WP_094435288).

Grupo XVI: este grupo, al igual que el anterior, abarca dos IntI de tapetes continentales (IntI_MTG10_1492, IntI_MTG10_95) y una IntI de tapetes marítimos (IntI_MTG7_2866). Las tres secuencias exhibieron similitud con la misma integrasa de una Betaproteobacteria no clasificada (WP_082354447.1).

d) Grupos con IntI similares a integrasas de referencia:

Por otra parte, las IntI que se relacionaron con integrasas ya descritas y conocidas se distribuyeron en nueve grupos, tres de los cuales incluyeron IntI móviles de clase 1, 2 y 3. Los grupos II, XIII y XVIII que incluyen las integrasas móviles y las IntI antárticas ya fueron descritas al inicio de esta Sección.

Grupo III: este clado está formado por la secuencia de una bacteria marina, *Sacarophagus degradans* (ABD79689), que degrada polisacáridos complejos (Weiner et al., 2008) y por una integrasa recuperada a partir de metagenomas de sedimentos (Antártida marítima) de Isla Ardley, IntI_IA6_8458. Esta secuencia exhibió >90% de identidad con la secuencia integrasa de la cepa AMac2203 de *Oceanisphaera avium* (WP_086963428.1).

Grupo XII: en este grupo se ubican las integrasas de *Xanthomonas oryzae* (WP011260557) y *Xanthomonas campestris* ATCC33913 (NP635739) junto a integrasas de sedimentos (Antártida marítima), IntI_CS1_49176 e IntI_IA6_36662. Las dos IntI exhibieron >74% de identidad con una integrasa identificada en un metagenoma de filosfera (RZA18127) (Crombie et al., 2018).

Otros dos grupos comparten un origen en común pero forman dos sub-grupos que constituyen clados distintos (XIX y XX). **El grupo XIX** está formado por tres IntI de tapetes microbianos continentales (MTG10_224, MTG10_6152 y MTG12_6634) que están cercanamente emparentados con una integrasa de *Methylibium petroleiphilum* (WP011828296).

Por su parte, **el grupo XX** está formado por secuencias de sedimentos (Antártida marítima) IntI_CS1_2689, IntI_CS1_12859, IntI_CS1_503,

IntI_MTG8_2561 y una secuencia de tapetes continentales IntI_MTG14_6926, que forman un grupo monofilético con la IntI de *T. denitrificans* (WP_011311727). El análisis de secuencia mostró que las cuatro IntI de Antártida marítima (IntI_CS1_12849, IntI_CS1_503, IntI_CS1_2689 y IntI_MTG8_2561) son similares a la secuencia integrón integrasa de *Simplicispira psychrophila* (WP_084153950.1), mientras que IntI_M14_6926 mostró similitud con la secuencia de *Polaromonas* sp. CG9_12 (WP_036804898.1).

Grupo XXI: lo conforman dos secuencias IntI de tapetes de la Antártida continental (MTG12_3385 y MTG13_7212) que se ubican en una rama próximo a una integrasa de *Nitrococcus mobilis* NB-231 (EAR20195). Las dos IntI exhibieron similitud con la integrón integrasa de *Synechococcus* sp. PCC 7335 (EDX85162.1).

Por último, **el grupo XXII** se divide en dos subgrupos, un sub-grupo contiene una rama que soporta un super-integrón de *Pseudomonas alcaligenes* (AAK73287) y el otro sub-grupo formado por dos ramas con la IntI de *Pseudomonas stutzeri* (AAN16071) y la IntI_CS1_5 de sedimentos (Antártida marítima). El análisis de secuencia indica que IntI_CS1_5 está relacionada con una integrasa de una bacteria no clasificada del Orden Pseudomonadales (OHC29044.1) (Anantharaman et al., 2016).

iv) Contexto genómico de integrones de metagenomas de tapetes microbianos y sedimento antárticos.

En esta sección se analizará el contexto genómico en el cual se ubican los integrones identificados por el programa Integron Finder y los GCS asociados. En los casos que fue posible se agregó la anotación generada por el servidor RAST. La descripción sigue el orden creciente en base al número romano del grupo correspondiente a la Figura 3.6

Cabe aclarar que solo se mencionan los grupos que incluyen integrones de tapetes y sedimentos obtenidos mediante secuenciación *shotgun*. No se incluyen las integrasas de bibliotecas de amplicones porque en este caso no fue posible

obtener información del entorno genómico del integrón. Por esta razón, no todos los grupos están mencionados en la descripción.

Grupo II

El grupo de integrasas emparentadas con integrasas de clase 2 (IntI2) incluyeron los integrones CS1_57033 y CS1_17527. Las dos secuencias IntI provienen de muestras de sedimento de la cámara séptica de la BCAA, uno de los sitios de la Isla con mayor impacto antropogénico.

Al observar el entorno genómico de estos integrones, particularmente sus GCs asociados, encontramos que de acuerdo a la anotación asignada por Integron Finder, el integrón CS1_57033 corresponde a un integrón vacío sin sitios *attC* ni GCs asociados. Por su parte, el integrón CS1_17527 posee un sitio *attC* y un ORF que codificaría para una bacterioferritina (COG2193).

Las bacterioferritinas son proteínas involucradas en la asimilación y almacenamiento de hierro, un micronutriente esencial para el crecimiento y supervivencia de la célula bacteriana. Se ha observado que la utilización de bacterioferritinas y ferritinas está relacionada con la patogénesis y con la emergencia de resistencia a antibióticos. Particularmente en *Mycobacterium tuberculosis*, la sobre-expresión de estas proteínas está directamente vinculada a la resistencia a aminoglucósidos (Sharma et al., 2017).

Grupo V

El integrón IntI_HP3_2775 posee dos sitios *attC* y dos ORFs. Un ORF (168aa) codifica para una proteína similar a una proteína con dominio de función desconocida (Pfam DUF1772) de *Luteibacter* sp. UNCMF331Sha3.1 (WP_089979935). El otro ORF (56 aa) codifica para una proteína que exhibe similitud con una metiltransferasa de *Pyrinomonas methylaliphatogenes* (WP_041978712.1).

Grupo VIII

El integrón IntI_MTG5_9187 (Grupo VIII) posee un sitio *attC* y cinco ORFs (Figura 3.7). No es posible afirmar que todos los ORFs anotados correspondan a GCs de la región variable del integrón, dado que Integron Finder identificó solo un sitio *attC* asociado al gen *intI*.

Dos ORFs codifican para proteínas con dominios de función desconocida, DUF433 (IPR007367) y DUF820 (IPR008538). Las proteínas con dominio DUF433 no han sido caracterizadas, aunque posiblemente forman parte del sistema toxina-antitoxina. Las que poseen el dominio DUF820 son proteínas hipotéticas de cianobacterias cuya función se desconoce por completo. Algunas proteínas también contienen un dominio N-terminal Clp (IPR004176) que participa en la interacción con otras proteínas (Pfam, <https://pfam.xfam.org/>).

Otro de los ORFs codifica para una proteína RadA que participaría en la reparación del ADN dañado por alquilación endógena. El 3º ORF codificaría para una endorribonucleasa, YafQ, que forma parte de la familia RelE/ParE. Esta familia de proteínas incluye varias toxinas del módulo toxina-antitoxina. YafQ, (cuyo antídoto es el componente DinJ), se asocia al ribosoma y bloquea la traducción mediante el clivaje del ARNm.

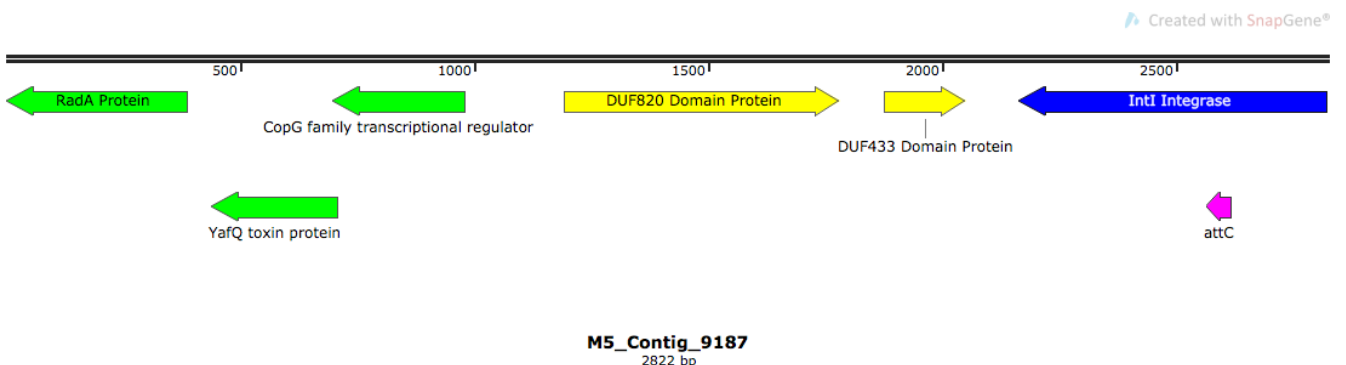


Figura 3.7: Contexto genómico del integrón identificado en el contig 9187 de MTG5 (Base Carlini I). Se muestra la enzima IntI (azul), el sitio *attC* (rosado), proteínas con dominios de función desconocida (amarillo) y proteínas que codifican para funciones varias (verde).

El integrón IntI_MTG10_1320 (Grupo VIII) (Figura 3.8), que proviene de tapetes microbianos de la región continental, posee cinco sitios *attC* y seis ORFs, dos de los cuales codifican para proteínas hipotéticas.

Los restantes ORF codificarían para:

- a) Una proteína con dominio AP2 (PF00847) similar a una proteína de *Chloroherpeton thalassium* (WP_012499248.1). AP2 es un dominio de unión al ADN que se encuentra en factores de transcripción de proteínas.
- b) Una pequeña proteína Barstar (ENOG410YOPO)(IPR000468). La proteína Barstar (Bacterial RiboNucleASE) Ribonuclease (E.C.3.1.27) forma parte del complejo Barnase/Barstar. Barnase es una ribonucleasa que es letal para la célula cuando se expresa sin su inhibidor Barstar. El inhibidor se une y bloquea el sitio activo de la ribonucleasa, previniendo de esta forma que la ribonucleasa dañe el ARN celular luego de ser sintetizado.
- c) Una proteína con dominio C-terminal CinA tipo S con función RecA, que participaría en la reparación de ADN (anotación RAST Server).
- d) Una trasposasa de la familia IS630.

No es posible afirmar que los ORFs mencionados en los ítems c y d correspondan a GCs ya que no tienen sitios *attC* asociados.

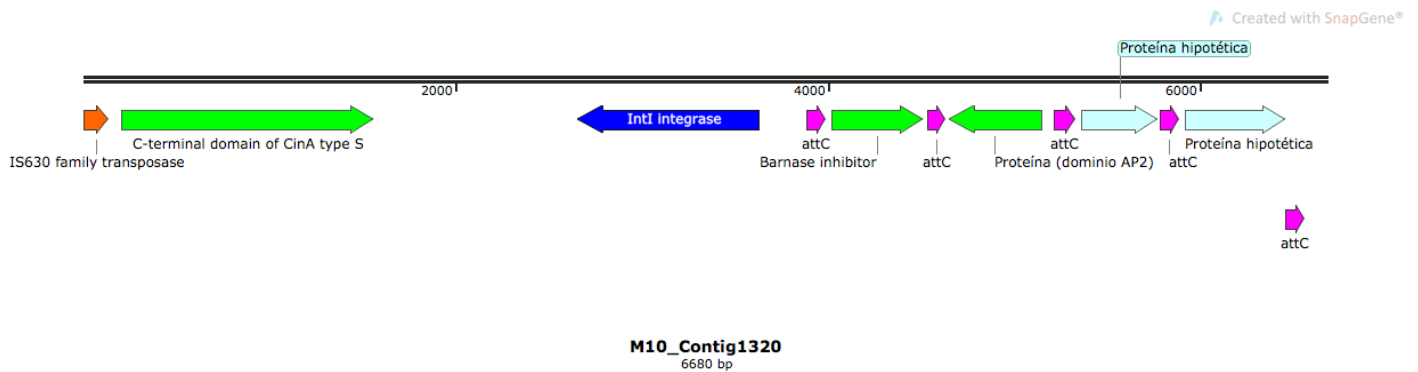


Figura 3.8: Contexto genómico del integrón identificado en el contig 1320 de MTG10. Se muestra la IntI (azul), sitios *attC* (rosado), proteínas hipotéticas (celeste), proteínas con funciones diversas (verde) y una trasposasa (naranja).

Grupo XII

Las secuencias relacionadas con integrasas de *Xanthomonas oryzae* (WP011260557) y *Xanthomonas campestris* ATCC33913 (NP635739) (Grupo XII) incluyeron un integrón vacío IntI_CS1_49176 y un integrón completo IntI_IA6_36662 con un sitio *attC* y un ORF que codificaría para una N-acetiltransferasa (familia GNAT) de *Lachnoclostridium* sp. SNUG30370 (WP_106988712).

Grupos XV y XVI

Los grupos XV y XVI están formados por tres integrones cada uno. El grupo XV posee un Ino (MTG5_26375) y dos integrones completos provenientes de la misma muestra de tapetes continentales: MTG12_6634 (sin ORFs asociados) y MTG12_3026 con un sitio *attC* y una trasposasa de 177 aa de largo, cuyo ortólogo más cercano pertenece a una especie del género *Shewanella* (ENOG410XYP3). El grupo XVI posee tres integrones vacíos, Ino (MTG7_2866, MTG10_1492 y MTG10_95).

Grupo XVII

El grupo XVII está formado por tres IntI vacíos Ino de tapetes microbianos de la Antártida marítima MTG4_58, MTG4_2937 y MTG5_28909).

Grupo XVIII

Las secuencias relacionadas con IntI clase 3 forman dos subgrupos, uno incluye el integrón IntI_MTG9_2773 y las IntI3 clínicas de *K. pneumoniae* sp. y *S. marcescens* sp.

Al analizar el entorno genómico de MTG9_2773 encontramos un sitio *attC* asociado a un ORF que codifica para una fosfolipasa A2 (PLA2). Las PLA2 son una familia de enzimas lipolíticas que hidrolizan los enlaces acil-éster de lípidos de membrana. Los productos resultantes (lisofosfolípidos y ácidos grasos libres) participan en rutas de señalización celular y son además precursores de factores que median la respuesta inflamatoria en el huésped (Sitkiewicz et al., 2007). Se ha observado que algunas PLA2 que funcionan como factores de virulencia en patógenos humanos (*Pseudomonas aeruginosa* y *Streptococcus* Grupo A), se presume que han sido adquiridas mediante eventos de THG (Sitkiewicz et al., 2007).

Contiguo al locus de PLA2 se encuentra una trasposasa IS110. La familia IS110 con dominio DEDD (Asp, Glu, Asp, Asp) es la única clase de trasposasas que se conoce codifica para RuvC, una resolvasa que participa en el clivaje de los intermediarios de Holliday durante el proceso de recombinación homóloga (Siguier et al., 2014).

De esta forma, la presencia de un elemento móvil en el entorno del integrón sugiere que podría estar implicado en la movilización de los GCs asociados. Sin embargo, en este caso los ORFs codifican para proteínas hipotéticas por lo que no es posible especular sobre las implicancias que tendría la movilización de estos genes.

El otro subgrupo incluyó siete secuencias IntI de tapetes y sedimento colectadas en sitios próximos a las bases de operaciones científicas. En este subgrupo se

incluyen dos integrones vacíos (MTG5_52806 y MTG9_4540) y una IntI con un sitio *attC* sin ORF asociado (MTG5_31000).

Según la anotación asignada por el servidor RAST, en el contig 4540 de MTG9 se identificaron dos ORFs que codifican para proteínas osmoprotectantes de la familia de transportadores ABC (YehW e YehX). Se ha observado en células de *E. coli*, que el operón *yehZYXW* codifica para un sistema de transportadores ABC que es inducido en situaciones de estrés ambiental y cuya función es actuar como osmoprotectante (Checroun & Gutierrez, 2004).

También se ubican cuatro IntI que provienen de la cámara séptica y se distribuyen de la siguiente forma: un integron vacío (CS1_24080) y tres integrones completos con un sitio *attC* asociado.

CS1_4987 y CS1_4254 poseen un ORF cada uno, que codifica para proteínas hipotéticas y CS1_2164 contiene un ORF que codifica para una proteína involucrada en la biosíntesis de purina (PurH) (COG0138). Además, según la anotación del servidor RAST contiguo a PurH se encuentra una dihydrouridina sintasa, una probable proteína de membrana (YPO3302) y una proteína (YchF) de la familia de las GTPasas que desempeñan una función clave en la síntesis de proteínas (Figura 3.9). En este caso el único ORF que posee un sitio *attC* asociado es el que codifica para la proteína PurH y puede considerarse que forma parte del arreglo de GCs del integrón.

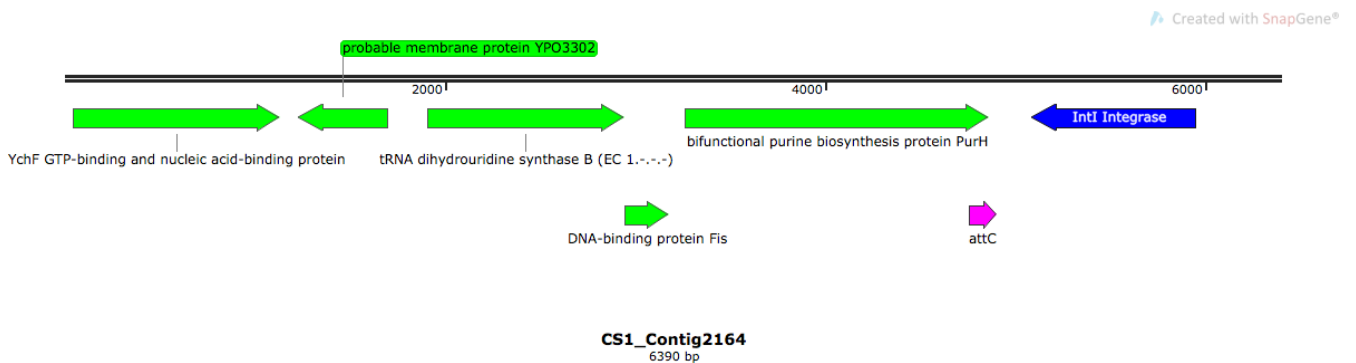


Figura 3.9: Se describe el contexto genómico del integrón identificado en el contig 2164 de CS1. Se muestra la integrasa IntI (azul), el sitio *attC* (rosado) y proteínas con funciones varias (verde).

Los ORF presentes en el entorno del integrón CS1_2164 fueron muy similares a nivel nucleotídico a los genes correspondientes, presentes en la cepa DSM24876 de *Rhodospirillum rubrum*.

R. antarcticus es un miembro del filum *Betaproteobacteria*, y representa una de las cuatro especies del género *Rhodospirillum*, una bacteria psicrófila púrpura no sulfurosa (fototrófica anoxigénica). Este grupo de bacterias presentan una gran plasticidad metabólica, lo que les permite crecer tanto como autótrofos o como quimiótrofos (Baker et al., 2017).

La cepa ANT.BR^T de *R. antarcticus*, aislada de matas microbianas en la Isla Ross en la región de Mc Murdo (continente Antártico), fue la primer bacteria fotótrofa anoxigénica psicrófila en ser secuenciada (Baker et al., 2017). El análisis de su genoma reveló que posee un plásmido (198,6 kb) con 21 genes de ARNt, (1/3 de los genes de ARNt de su genoma). Todos los genes de ARNt del plásmido son el duplicado de anticodones presentes en el cromosoma y se desconoce que función cumplen en la traducción de *R. antarcticus*. Se especula que el ensamblaje de genes de ARNt potencialmente degenerados sea una estrategia de adaptación al

frío, dado que la misma organización ha sido observada en otras bacterias psicrófilas (Math et al., 2012). La presencia de un gran número de genes de ARNt está relacionada con tasas elevadas de traducción que aceleran las reacciones biosintéticas y favorecen el crecimiento, normalmente enlentecido a bajas temperaturas. De esta forma, el total de genes ARNt en el genoma de un organismo psicrófilo funcionaría como un indicador de adaptación al ambiente (Math et al., 2012).

Las dos cepas de *R. antarcticus* (DSM24876 y ANT.BR^T) poseen los genes que codifican para las enzimas DuSA, DusSB y DusC implicadas en la modificación de la dihidrouridina en moléculas de ARNt, aumentando la flexibilidad conformacional de los ARNt.

El entorno genómico del integrón CS1_2164 es muy similar a los genes correspondientes de la cepa DSM24876 de *Rhodospirillum rubrum*. Los ORFs localizados en el entorno del integrón que codifican para el regulador transcripcional Fis, la enzima DusSB, la proteína PurH y el regulador redox YchF están también presentes en ambas cepas de *R. antarcticus*, sugiriendo que existiría cierta homología entre el entorno genómico del integrón y la región correspondiente en dichas cepas.

Grupo XIX

Tres integrones de tapetes microbianos continentales que comparten origen común con la secuencia integrasa de *Methylobium petroleiphilum* (WP011828296) (Grupo XIX) corresponden a: un integrón vacío, MTG12_6634 y dos integrones completos de una misma muestra, MTG10. El integrón del contig MTG10_224 posee dos sitios *attC* y un ORF que codifica para una proteína de membrana (COG0600). MTG10_6152 posee un sitio *attC* y dos ORFs que codifican para una proteína con dominio de función desconocida (Pfam DUF1173) y para una N-acetil-transferasa.

Grupo XX

Por otra parte, cinco integrones completos forman un grupo monofilético con la secuencia IntI de *T. denitrificans* (WP_011311727).

El integrón de CS1_2689 tiene tres sitios *attC* y cuatro ORFs, dos de los cuales codifican para proteínas hipotéticas, los otros tres codificarían para el factor tipo A von Willebrand (vWA) (COG4245), para una chaperonina (COG4248) y para una fosfoproteína fosfatasa (ENOG411YVJ) (inferido por la base de datos egg-NOG)(Figura 3.10).

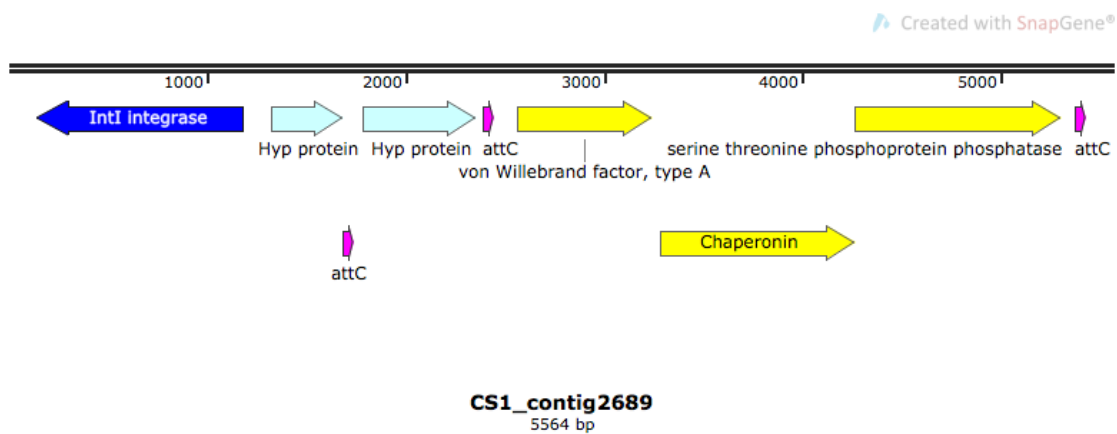


Figura 3.10: Se describe el contexto genómico del integrón identificado en el contig 2689 de CS1. Se muestra la IntI (azul), los sitios *attC* (rosado), dos proteínas hipotéticas (celeste), una chaperona, una fosfatasa y una proteína con dominio vWA(amarillo).

El integrón IntI_CS1_12849 posee dos sitios *attC* y ningún ORF asociado mientras que el integrón de CS1_503 contiene tres sitios *attC* y dos ORFs que codifican para una β -lactamasa de clase D (Resfam:RFO056) y para una proteína hipotética.

El integrón de MTG8_2561 contiene un sitio *attC* y un ORF que codifica para una antitoxina hipotética (COG3636). Asimismo, según la anotación del servidor RAST, contiguo a la antitoxina se encuentran ORFs que codificarían para una

proteína hipotética, una prolina deshidrogenasa (E.C.1.5.99.8) y un regulador de respuesta a leucina (Figura 3.11).

A su vez, el integrón de MTG14_6926 posee un sitio *attC* y tres ORFs que codifican para dos proteínas hipotéticas y una proteína con dominio de función desconocida (Pfam DUF2511),(ENOG410Y8JA). La familia DUF2511 está conservada en bacterias pero se desconoce su función.

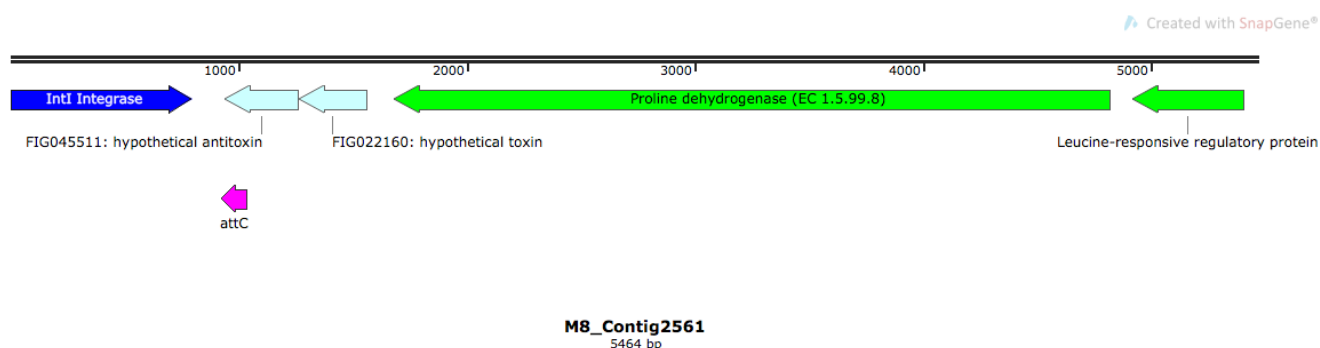


Figura 3.11: Se describe el contexto genómico del integrón identificado en el contig 2561 de MTG8. Se muestra la IntI (azul), los sitios *attC* (rosado), dos ORFs que codifican para dos proteínas hipotéticas (celeste), dos ORFs que codifican para una prolina deshidrogenasa y un regulador de leucina (verde). La IntI se orienta en sentido inverso a los GCs.

Grupo XXI

Dos integrones completos (MTG12_3385 y MTG13_7212) comparten un origen común con la secuencia integrasa de *Nitrococcus mobilis* NB-231 (Grupo XXI). El integrón de MTG12_3385 posee tres sitios *attC* y cuatro ORFs que codifican para una proteína hipotética, una proteína con dominio DUF820 (Pfam DUF820), (ENOG4111HPD) y dos proteínas de membrana de la familia de transportadores ABC (COG1463 y COG1127) (Figura 3.12).

Las proteínas con dominio DUF820 están ampliamente extendidas en cianobacterias y se encuentran con menor frecuencia en otras bacterias. Se cree que pertenecen a una superfamilia de nucleasas.

El integrón de MTG13_7212 tiene dos sitios *attC* y dos ORFs que codificarían para una peroxirredoxina (COG1125) y una endonucleasa de restricción de la familia Uma2.

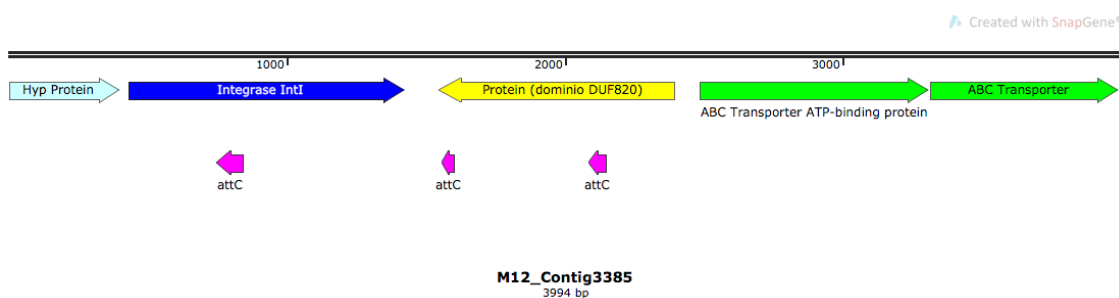


Figura 3.12: Se describe el contexto genómico del integrón identificado en el contig 3385 de MTG12. Se muestra la IntI (azul), los sitios *attC* (rosado), una proteína con dominio de función desconocida (amarillo), una proteína hipotética (celeste) y proteínas que codifican para transportadores (verde).

Grupo XXII

El integrón más largo identificado en este estudio pertenece a CS1_5 y posee 15 sitios *attC* y seis ORFs. (Figura 3.13). IntI_CS1_5 forma un grupo (XXII) con un superintegrón de *Pseudomonas alcaligenes* (AAK73287) y una IntI de *Pseudomonas stutzeri* (AAN16071).

Los integrones del género *Pseudomonas* están presentes únicamente en un pequeño número de especies y posiblemente hayan sido adquiridos mediante THG en un período tardío de su historia evolutiva. El número y composición de los GCs varía entre las especies, pueden tener desde diez GCs (*P. stutzeri*) hasta más de 32 (*P. alcaligenes*) (Boucher et al., 2007).

Todos los ORFs identificados en el integrón de CS1_5 codifican para proteínas hipotéticas, excepto un ORF (385 pb) que codificaría para una penicilin acilasa II

(EC. 3.5.1.11), una β -lactam-acilasa que hidroliza las penicilinas naturales (anotación RAST Server).

El número de sitios *attC* (n=15) presentes en el integrón CS1_5 y su relación filogenética con especies del género *Pseudomonas* sp. que portan integrones en su cromosoma permitiría inferir que se trata de un integrón cromosómico o incluso de un super-integrón. Poder determinar las funciones codificadas por los ORFs ayudaría a la especulación sobre la naturaleza de este integrón.

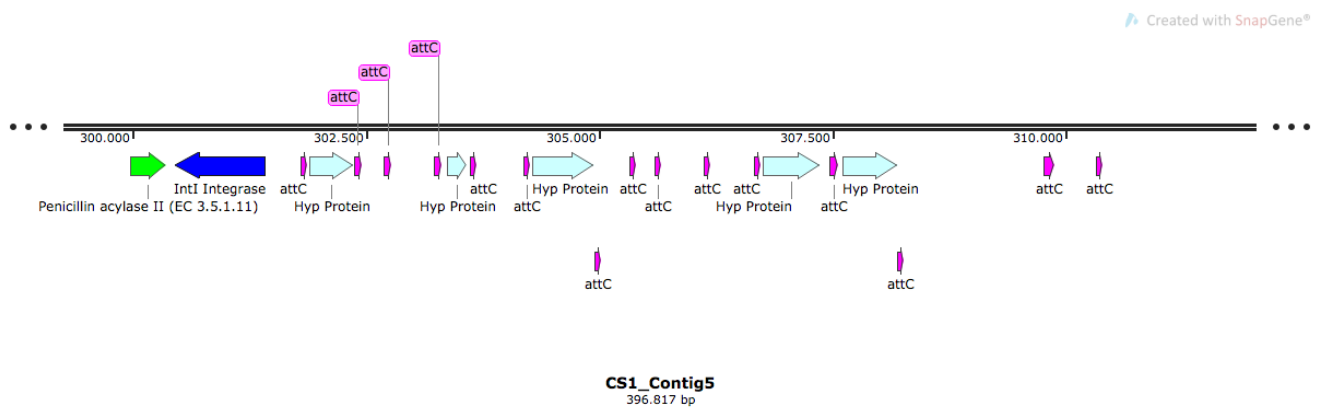


Figura 3.13: Contexto genómico del integrón completo identificado en el contig 5 de CS1. Se muestran la integrasa IntI (azul), proteínas hipotéticas (celeste), una penicilina acilasa (verde) y los sitios *attC* (rosado).

Por último, las ramas A, B y G incluyen secuencias IntI de tapetes de la Antártida continental, MTG12_12744, MTG13_3715 y MTG13_10203, respectivamente. El integrón de MTG12_12744 posee un sitio *attC* y un ORF que codifica para una proteína hipotética. La IntI exhibe >67% de identidad con una integrón integrasa de un organismo del phylum *Planctomycetes* (*Zavarzinella formosa*, WP_020475816.1). Por otra parte, MTG13_3715 contiene un integrón vacío con una secuencia IntI relacionada con una integrasa de una bacteria no clasificada del phylum Verrucomicrobia (PYL98288). La rama G contiene un integrón completo, MTG13_10203 con un sitio *attC* y un ORF que codifica para una proteína hipotética. La integrasa IntI_MTG13_10203 está relacionada con la secuencia IntI de *Candidatus Schekmanbacteria bacterium* GWA2_38_9 (OGL39197).

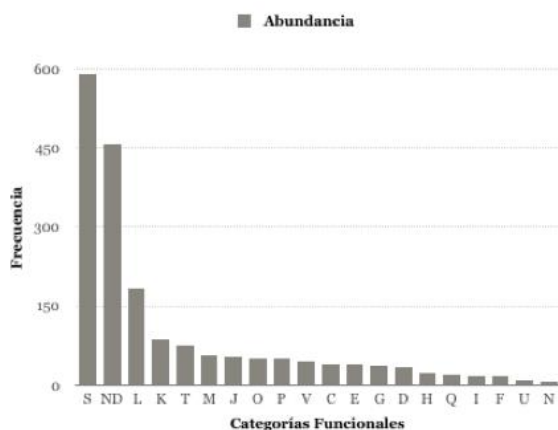
v) Anotación funcional de GCs y ORFs con eggNOG-mapper

En el histograma de la Figura 3.14 se muestra la distribución de la abundancia de GCs y ORFs para las 20 categorías de COGs asignadas por el programa eggNOG-mapper. Se observa que las representaciones de las categorías jerárquicas COGs en el *dataset* de metagenomas antárticos correspondieron a procesos relacionados con:

- 1) Almacenamiento de información y procesamiento: 17,04% (incluye las categorías de transcripción, replicación, recombinación y reparación, procesamiento y modificación de ARN).
- 2) Procesos de señalización celular: 14,67% (mecanismos de transducción de señales, mecanismos de defensa, modificaciones postranscripcionales, división, movilidad y envoltura celular).
- 3) Metabolismo: 12,98% (incluye las categorías de transporte y metabolismo de nucleótidos, aminoácidos, carbohidratos, lípidos, coenzimas, metabolitos secundarios y producción de energía).
- 4) Categorías de función desconocida o pobremente categorizadas representaron el 55,03%.

Tal como era esperable, el número de secuencias asignadas a categorías de genes cuya función se desconoce resultó ser muy elevado, representando más del 50% del total.

Por otra parte, analizando la totalidad de GCs y ORFs que fueron clasificados dentro de las categorías funcionales COGs (n=1895), solamente una pequeña fracción (n=29) están asociados al gen *intI* y por lo tanto son clasificados como integrones completos. El resto de los ORFs identificados constituyen CALIN (CAssette Lacking INtegrase) u ORFs sin sitio *attC* asociado.



Categoría funcional	Funciones	Abundancia
S	Función desconocida	591 (13,19%)
ND	N/A	457 (24,11%)
L	Replicación, reparación y recombinación	184 (9,7%)
K	Transcripción	86 (4,53%)
T	Mecanismos de transducción de señales	75 (3,95%)
M	Pared celular, membrana y biogénesis de la envoltura	56 (2,95%)
J	Traducción, estructura del ribosoma y biogénesis	53 (2,79%)
O	Modificación postraduccional, recambio de proteínas y chaperonas	51 (2,69%)
P	Transporte y metabolismo de iones inorgánicos	51 (2,69%)
V	Mecanismos de defensa	46 (2,42%)
C	Producción y conversión de energía	39 (2,05%)
E	Transporte y metabolismo de aminoácidos	39 (2,05%)
G	Transporte y metabolismo de carbohidratos	37 (1,95%)
D	Control del ciclo celular, división celular y partición del cromosoma	33 (1,74%)
H	Transporte y metabolismo de coenzimas	24 (1,26%)
Q	Biosíntesis, transporte y catabolismo de metabolitos secundarios	20 (1,05%)
I	Transporte y metabolismo de lípidos	19 (1,00%)
F	Transporte y metabolismo de nucleótidos	17 (0,89%)
U	Tráfico intracelular, secreción y transporte vesicular	10 (0,52%)
N	Movilidad celular	7 (0,36%)

Figura 3.14: Se describe la abundancia de genes para cada una de las 20 categorías funcionales de genes ortólogos (COG) obtenidas a partir del análisis de 17 metagenomas antárticos.

Los integrones completos cuyos GCs codifican para funciones asignadas a las categorías funcionales COGs, están distribuidos en cinco metagenomas de la Antártida continental (ANCO10, ANCO11, ANCO12, ANCO13 y ANCO14) y tres metagenomas de la Antártida marítima: Base Carlini I (MTG5), Isla Ardley (IA6) y la cámara séptica de la BCAA (CS1).

Asimismo, se realizó una representación gráfica (*heatmap*) de la distribución de abundancia de GCs y ORFs para comparar los perfiles funcionales de cada metagenoma. La abundancia se muestra mediante la gama de colores de azul (máximo) a blanco (mínimo).

El análisis incluyó 847 ORFs (se omitieron las categorías con función desconocida y pobremente categorizadas).

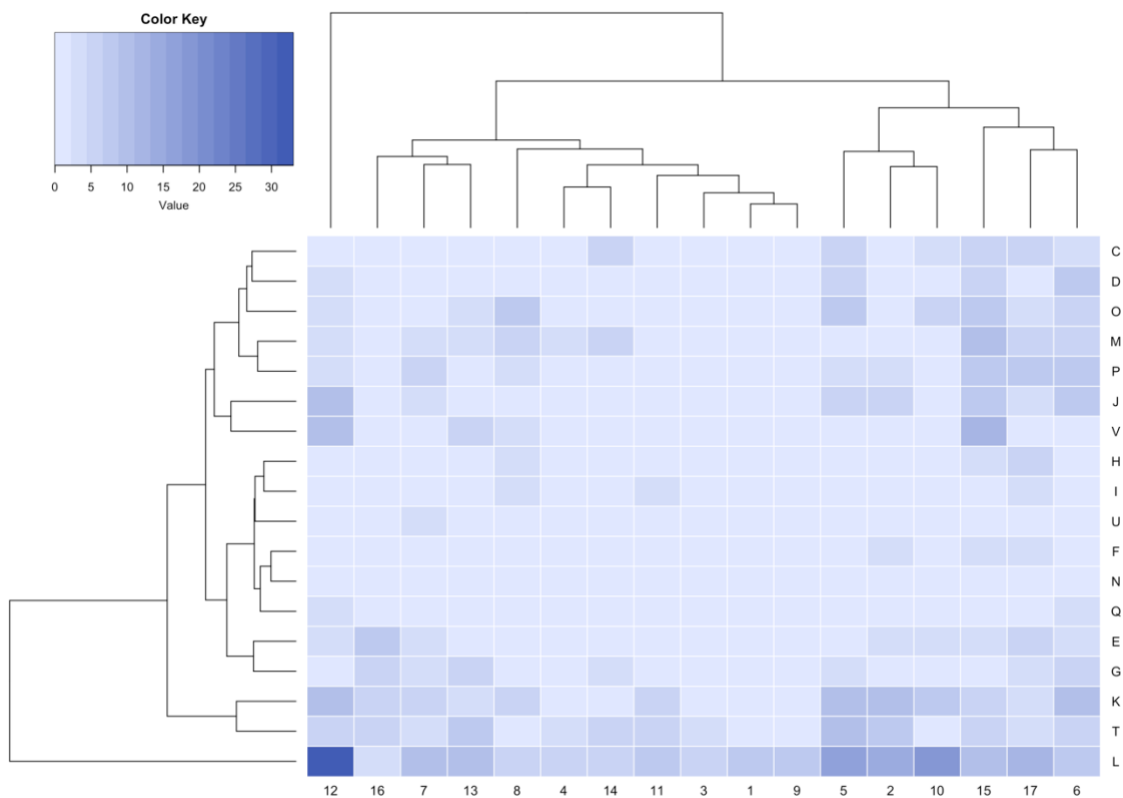


Figura 3.15: *Heatmap* y dendrograma que representa la distribución de la abundancia relativa de los ORFs para los 17 metagenomas (eje x) dentro de las categorías funcionales COGs (eje y).

El dendrograma muestra que los metagenomas se asocian en tres grupos diferentes:

- a) Metagenomas: N°2 (Drake_2), N°5 (Carlini I), N°6 (Carlini II), N°10 (ANCO10), N°15 (CS1) y N°17 (IA6).

- b) Metagenomas: N°1 (Espejo_1), N°3 (HP3), N°4 (Pista_4), N°7 (Esperanza_7), N°8 (Primavera_8), N°9 (San Martín_9), N°11 (ANCO11), N°13 (ANCO13) y N°16 (HP3).
- c) Metagenoma N°12 (ANCO12) que se ubica solo claramente separado del resto.

No se observa un patrón de distribución de los perfiles funcionales que asocie los metagenomas en base a la región geográfica (Antártida marítima y Antártida continental).

En lo que refiere a las categorías funcionales COGs, se observa que la categoría L (involucrada en los procesos de replicación, reparación y recombinación) está enriquecida en todo el *dataset* de metagenomas. Particularmente, la mayoría de las secuencias corresponden a genes que codifican para elementos móviles como transposasas, secuencias de inserción e integrasas. La gran proporción de elementos móviles podría ser un indicio de que existen eventos de THG en las comunidades microbianas presentes en los diferentes metagenomas.

Dentro de la categoría L también se identificaron secuencias que codifican para sistemas CRISPR (*clustered regularly interspaced short palindromic repeat*). Los CRISPR cumplen funciones de protección contra elementos genéticos móviles (como plásmidos conjugativos y elementos transponibles) y se considera que forman parte del sistema inmune de la célula.

Se identificaron también numerosos transportadores del tipo ABC, que incluyen principalmente transportadores de aminoácidos, iones inorgánicos y metabolitos secundarios. Asimismo, transportadores relacionados con mecanismos de defensa (categoría funcional V) están presentes en los metagenomas N°8 (Primavera 8), metagenoma N°12 (ANCO12), metagenoma N° 13 (ANCO13) y metagenoma N°15 (CS1).

Asimismo en varios metagenomas se encontraron genes que codifican para acetiltransferasas (categoría funcional J), particularmente de la superfamilia GNAT. La relevancia de las N-acetiltransferasas radica en su contribución a la emergencia de la resistencia antibiótica entre bacterias patógenas. La

modificación de los aminoglucósidos mediante la N-acetilación de grupos amino disminuye la afinidad del antibiótico por el sitio blanco (Dyda et al., 2000).

En la mayoría de las muestras (con excepción de los metagenomas N°1, N°3, N°6, N°9 y N°16) se observa un enriquecimiento de las categorías que incluyen procesos involucrados en el control del ciclo celular, división celular, partición del cromosoma, biogénesis de la pared celular, traducción, modificación post-traduccional, recambio de proteínas y producción de energía, sugiriendo que las comunidades microbianas allí presentes se encontrarían activas.

vi) Composición taxonómica de la comunidad de metagenomas antárticos.

Los datos metagenómicos generados mediante secuenciación *shotgun* fueron depositados en el servidor MG-RAST para la anotación de las secuencias. De esta forma el *dataset* se compone de 14 metagenomas de tapetes microbianos y dos metagenomas de sedimento/suelo (IA6 y HP3). El número de secuencias que se obtuvo para cada sitio, el número de pares de bases (Nº pb), el largo promedio y porcentaje de contenido GC (%GC) se describen en la Tabla suplementaria S3 y S4.

Los sitios de muestreo se pueden clasificar en tres zonas biogeográficas que denominamos arbitrariamente Zona 1 (Isla Rey Jorge), Zona 2 (Península Antártica) y Zona 3 (Continente Antártico). En la figura 3.16, la Zona 1 está representada en 3.16 A, la Zona 2 en 3.16 B y la Zona 3 en 3.16 C.

Considerando que las muestras son diferentes en cuanto a su composición, los metagenomas de sedimento/suelo serán analizados de forma separada a los tapetes microbianos.

Analizando en conjunto la distribución taxonómica de la comunidad de tapetes microbianos, se observa que a nivel de phyla, los grupos más abundantes estuvieron representados por miembros de *Cyanobacteria*, *Bacteroidetes*, *Proteobacteria*, *Verrucomicrobia* y *Firmicutes*. Mientras que *Cyanobacteria* fue el filo dominante en siete metagenomas (Drake, HTP3, Carlini I, Carlini II, Esperanza, Primavera y MidGood 13), *Bacteroidetes* lo fue en cinco (Espejo2, Pista 4, San Martín 9, ANCO 11 y ANCO 14) y *Proteobacteria* solo en dos (ANCO 10 y ANCO 12) (Figura 3.16, A, B y C).

A nivel de clase, el grupo mayoritario en once metagenomas lo constituyeron las secuencias no clasificadas derivadas de Cianobacterias (28%-59%), mientras que *Flavobacteriia*, *Cytophagia* y *Beta-Proteobacteriia* fueron la clase predominante en Pista4, San Martín9 y ANCO10 (B012-2015-14), respectivamente (Figura suplementaria S1).

Descripción de la Zona 1: Isla Rey Jorge

a) Metagenomas de tapetes microbianos (Figura 3.16).

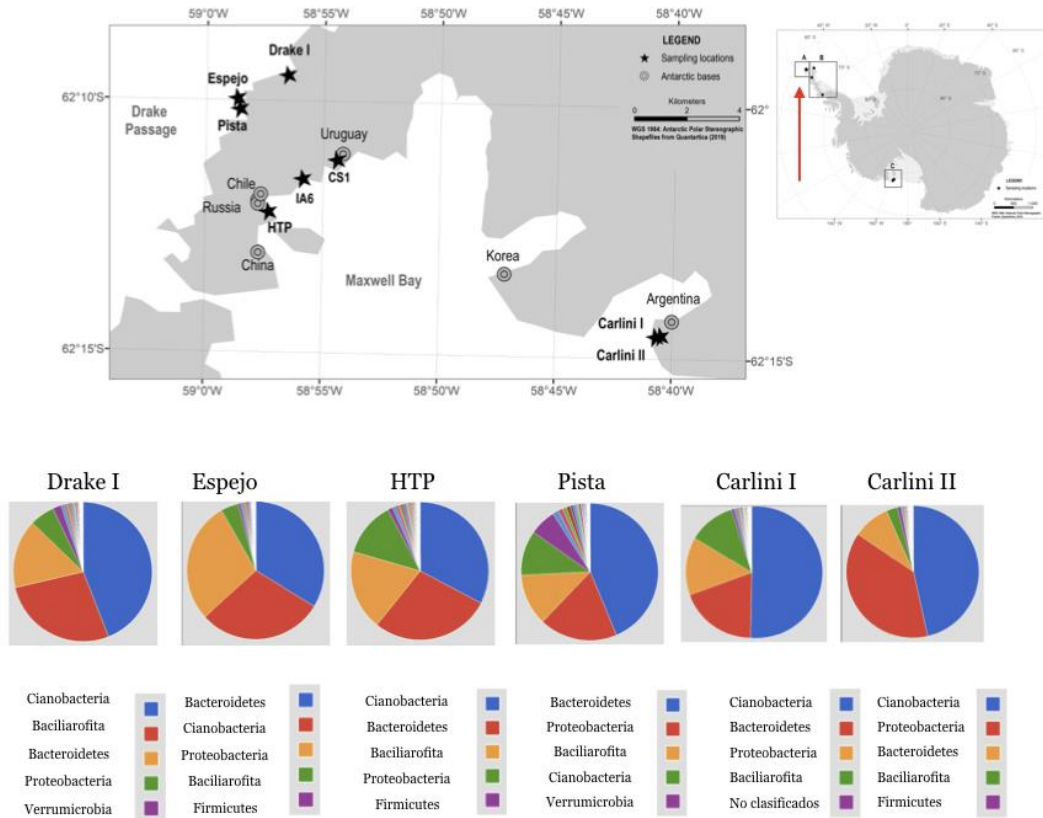


Figura 3.16A: Localización geográfica y distribución taxonómica de las comunidades microbianas de los metagenomas de Isla Rey Jorge. (Anotación Servidor MG-RAST)(Meyer et al., 2008).

La composición de tapetes microbianos a nivel de phylum indica que los grupos más abundantes correspondieron a *Cyanobacteria*, *Bacteroidetes*, *Proteobacteria*, *Verrumicrobia* y *Firmicutes*. También incluye el grupo *Baciliariofita* que pertenece al dominio Eucaria. El filo *Baciliariofita* comprende a las Diatomeas que contienen cloroplastos. La abundancia de

cloroplastos/Cianobacterias está relacionada con la proximidad de los sitios a la zona costera (Lopatina et al., 2016).

Si analizamos la composición de la comunidad a nivel de clase, se observa que el grupo más abundante en cinco de los seis sitios correspondió a las secuencias no clasificadas derivadas de Cianobacteria (37-56%) y a la clase *Flavobacteria* (30,88%) en un solo sitio (Pista 4). El resto de los grupos más representados está integrado por *Betaproteobacteria*, *Cytophagia*, *Verrumicrobiae*, *Bacteroidia* y *Spinghobacteria*.

b) Metagenomas de sedimento/suelo (Figura 3.16B)

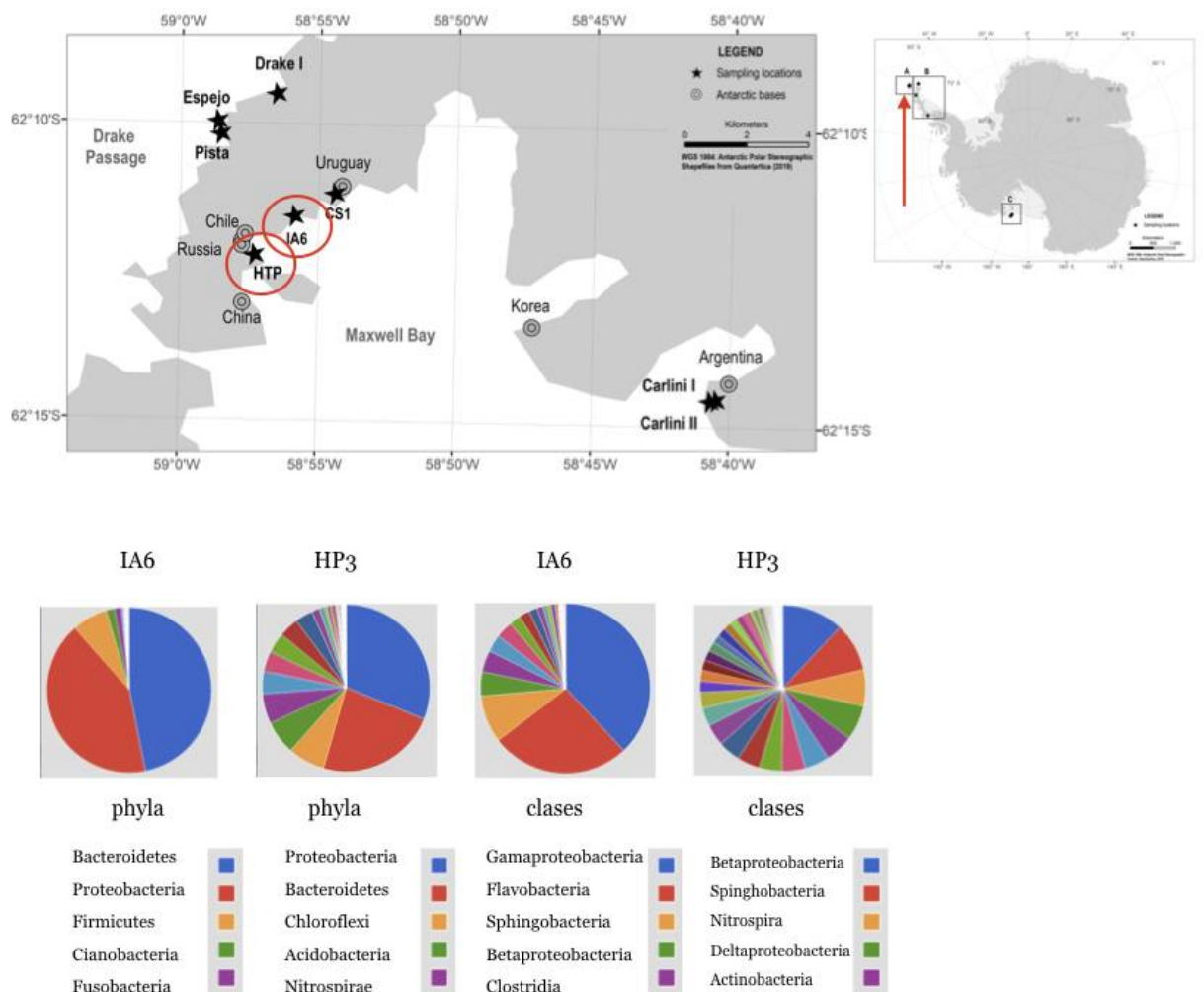


Figura 3.16B: Localización geográfica y distribución taxonómica de las comunidades de metagenomas de sedimento/suelo de IA6 y HP3 (Isla Rey Jorge). (Anotación Servidor MG-RAST)(Meyer et al., 2008)

El análisis de la composición taxonómica de IA6 y HP3 revela que existen diferencias en cuanto al grado de diversidad presente en cada metagenoma, tanto a nivel de phyla como de clases. Sin embargo, en ambos metagenomas los phyla dominantes estuvieron representados por miembros del grupo *Bacteroidetes* y *Proteobacteria*. En IA6 los *Bacteroidetes* constituyen el 46,85% y en HP3 un 23,21%. Mientras que *Proteobacteria* representa el 41,71% en IA6 y el 31,10% en HP3.

Los phyla presentes en HP3 incluyen los grupos *Chloroflexi*, *Acidobacteria*, *Nitrospirae*, *Actinobacteria*, *Cyanobacteria*, *Firmicutes*, *Verrucomicrobia* y *Gemmatimonadetes*. Además incluye miembros de catorce clases diferentes que representan entre el 2,18% y el 12,05%, siendo las clases más numerosas Betaproteobacteria, Sphingobacteria, Nitrospira, Deltaproteobacteria, Actinobacteria, Cytophagia y Alphaproteobacteria.

Mientras tanto IA6 muestra una composición relativamente diferente a HP3, donde Gamaproteobacteria representa la clase dominante (38,09%), seguida de Fusobacteria (26,43%) y Sphingobacteria (9,09%) y en menor proporción se encuentran las clases Betaproteobacteria, Clostridia, Bacterioidia y Cytophagia.

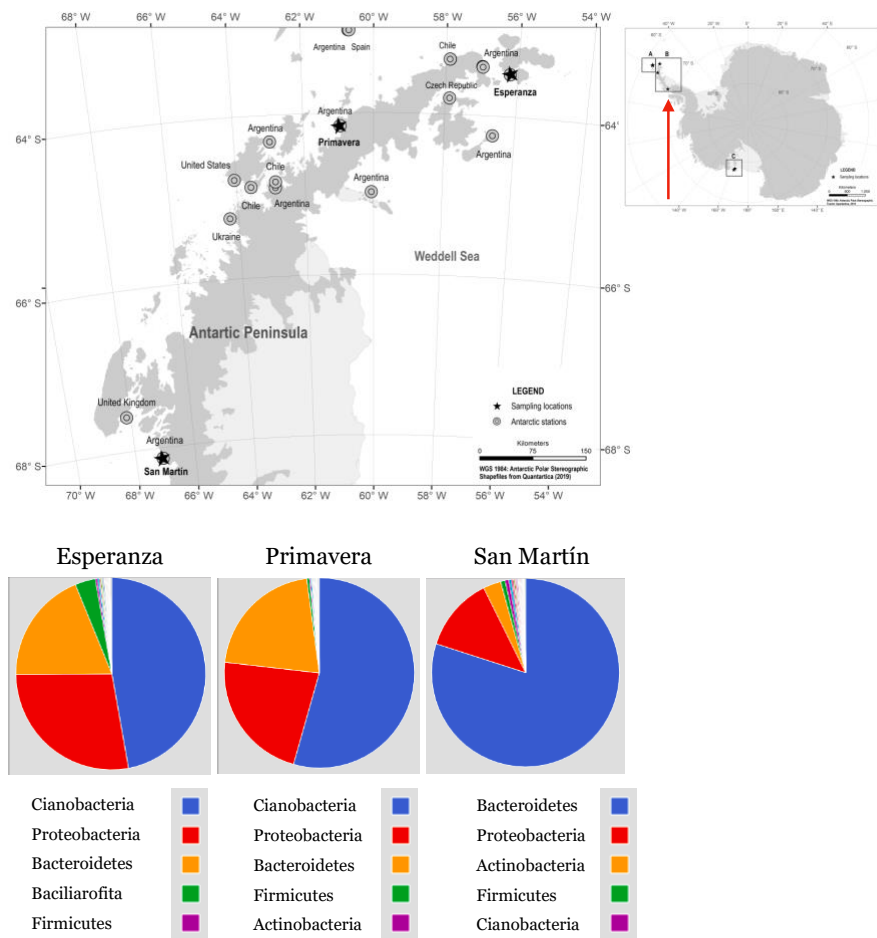


Figura 3.16 C: Localización geográfica y distribución taxonómica de la composición de la comunidad de metagenomas de la Península Antártica. (Anotación Servidor MG-RAST)(Meyer et al., 2008)

En la Zona 2: Península Antártica (Figura 3.16C), el phylum dominante fue *Cyanobacteria* en dos de los tres sitios, Esperanza (47,12%) y Primavera (54,38%), seguido por el grupo de *Proteobacteria* y *Bacteroidetes*, mientras que en San Martín, *Bacteroidetes* representó casi el 80% de las secuencias. Por otra parte, las secuencias no clasificadas derivadas de *Cyanobacteria* fueron mayor al 50% en Esperanza y Primavera, mientras que en San Martín la clase dominante fue Cytophagia que incluye varias especies psicrofílas. La clase Cytophagia (único Orden Cytophagales) son bacterias Gram-negativas, que producen pigmentos carotenoides y flexirubinas.

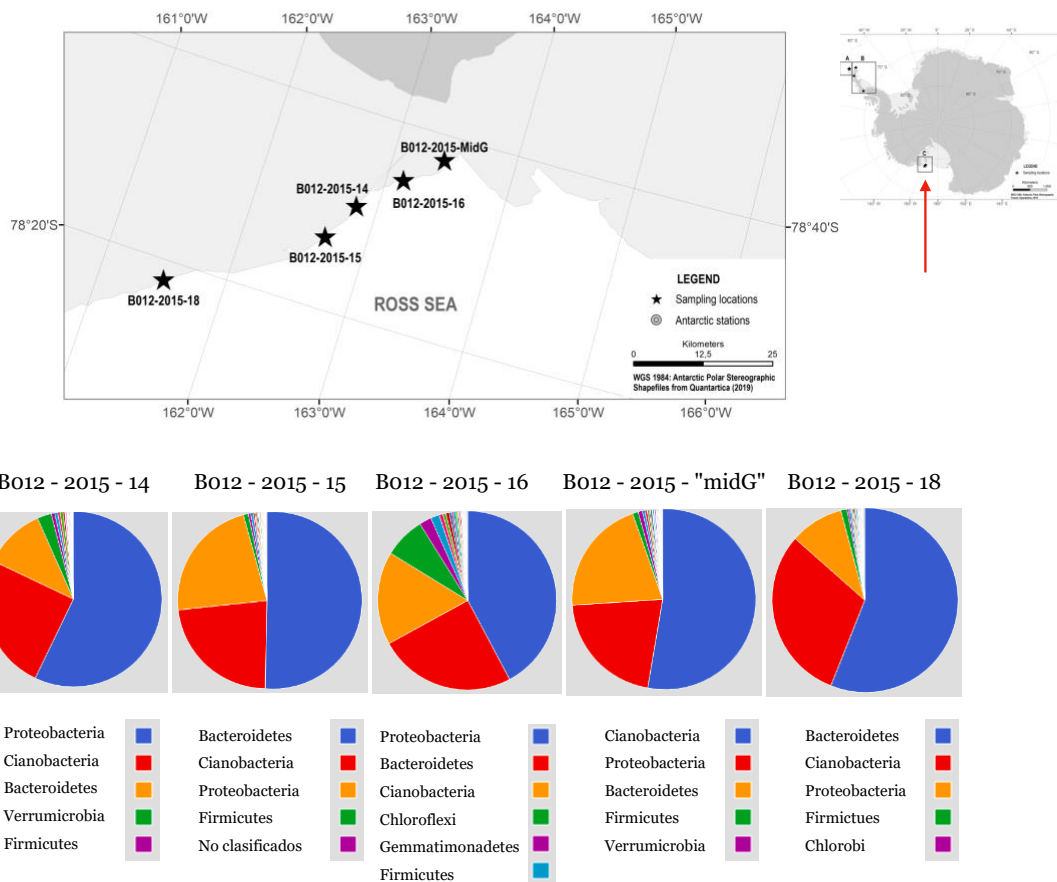


Figura 3.16 D: Localización geográfica y distribución taxonómica de la composición de la comunidad de metagenomas del continente antártico. (Anotación Servidor MG-RAST) (Meyer et al., 2008)

Zona 3 (Continente Antártico): los tres phyla predominantes en las cinco muestras de esta zona incluyen al grupo de *Proteobacteria*, *Cyanobacteria* y *Bacteroidetes*. En tres muestras *Firmicutes* ocupó el cuarto lugar (ANCO 11, ANCO 13 y ANCO 14), *Verrumicrobia* en ANCO 10 y *Chloroflexi* en ANCO 12. A nivel de clase, las secuencias no clasificadas derivadas de *Cyanobacteria* fue el grupo mayoritario en cuatro de las muestras, llegando a constituir casi el 60% del total en B012-2015-16 (ANCO 12) y el segundo grupo en ANCO 10 (B012-2015-14), en este último la clase Betaproteobacteria ocupó el primer lugar. Alfa, Beta, Delta y Gama Proteobacteria se detectaron en todas las muestras, así como la clase Cytophagia y Spingobacteria.

En concordancia con estudios previos desarrollados en diferentes regiones biogeográficas de la Antártida, los grupos predominantes incluyen representantes de las phyla *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, *Acidobacteria* y *Cyanobacteria* (Tytgat et al., 2014)(Lopatina et al. 2016).

Asimismo, *Proteobacteria* y *Actinobacteria* son los phyla dominantes, reflejando posiblemente la versatilidad metabólica de los integrantes de estos dos grupos. También en once metagenomas las secuencias no clasificadas derivadas de Cianobacterias constituyeron el grupo más representado sugiriendo la presencia de especies novedosas que aún no han sido caracterizadas. Cianobacteria fue el filo dominante en el 50% de las muestras de tapetes microbianos.

3.3.3- Discusión

Para analizar los resultados obtenidos en este trabajo existen algunos factores a considerar. Por un lado se debe tener en cuenta que las muestras analizadas son de origen diverso (tapetes microbianos y sedimento/suelo). Por otra parte, la metodología empleada para recuperar los genes de interés no fue la misma, en un caso se aplicó secuenciación *shotgun* y en el otro se construyeron bibliotecas de amplicones de un fragmento del gen *intI*.

En el presente análisis se determinó la relación filogenética de 129 secuencias integrasas antárticas con integrasas representativas de integrones móviles e integrones cromosómicos.

Se construyó un árbol de Máxima Verosimilitud con buen soporte que mostró una buena estructuración y mucha divergencia en las secuencias integrón integrasas. El 74,42% de las secuencias antárticas se distribuyeron en 19 grupos bien definidos y separados de secuencias integrasas conocidas, sugiriendo que podría tratarse de bacterias cuyos integrones aún no han sido caracterizados y por lo tanto no están disponibles en las bases de datos de integrasas.

El resto de las IntI se agruparon por un lado con integrasas móviles de tipo clínico (clases 1, 2 y 3) (n=19), mientras que 14 secuencias IntI se ubicaron formando clústers con distintas integrasas de especies conocidas que portan integrones cromosómicos o super-integrones.

Ninguna de las secuencias identificadas en este estudio agruparon con integrasas del género *Vibrio* (clase 4 y clase 5).

Analizando las distintas estrategias empleadas, se destaca que no se recuperó ningún integrón clase 1 de metagenomas de tapetes y sedimento secuenciados mediante *shotgun*.

Los integrones clase 1 de tipo clínico están frecuentemente asociados a genes de resistencia a antibióticos, metales pesados y genes que confieren resistencia a desinfectantes (genes *qac* que codifican para compuestos del amonio cuaternario). Asimismo, están ampliamente distribuidos en organismos patógenos y comensales. Por su localización plasmídica pueden dispersarse

rápidamente entre bacterias mediante THG. Por esta razón, la abundancia relativa de la variante clínica del gen *intI1* es empleada como marcador de contaminación antropogénica (M. R. Gillings et al., 2015). Tomando en cuenta estas consideraciones resulta llamativo que ninguna de las integrasas recuperadas de la muestra de la cámara séptica de la BCAA (CS1) caiga dentro del grupo de IntI1.

En el estudio previo que desarrollé empleando metodologías dependientes del cultivo (medio LB) en muestras colectadas en la Isla Rey Jorge (Antártida marítima), logramos identificar integrones clase 1 de tipo “clínico”. Estos integrones están asociados a determinantes de resistencia a antibióticos localizados en plásmidos conjugativos de enterobacterias psicrotolerantes (Antelo et al., 2018). Los genes *intI1* ampliamente conservados, fueron identificados en aislamientos recuperados a partir de muestras de sedimento y tapetes colectadas en distintos sitios de la Isla durante diferentes campañas organizadas por el IAU (Instituto Antártico Uruguayo). Los cinco aislamientos exhibieron un perfil genómico similar y los plásmidos presentaron las mismas características genotípicas y fenotípicas.

Cuando construimos bibliotecas de amplicones del gen *intI1* empleando los cebadores originalmente diseñados por Mazel et al. (2000), se recuperaron integrasas clase 1 idénticas a las identificadas en las enterobacterias psicrotolerantes (Antelo et al., 2015). Las secuencias IntI1 fueron idénticas entre sí y todas exhibieron identidad con integrasas de origen clínico. Además de estos integrones de tipo clínico, se obtuvo una gran diversidad de genes integrón integrasas (*intI*) que presentaron similitud con integrasas ambientales de diversos ecosistemas de regiones geográficas remotas (Antelo et al., 2015). Cabe mencionar que el par de cebadores IntI1F-IntIR1 no está dirigido específicamente a integrones clase 1 “clínicos”. En ese sentido, una gran proporción de las secuencias recuperadas fueron similares a integrasas descritas en ambientes muy diversos o formaron grupos únicos separados de las secuencias conocidas. Varios grupos de IntI mostraron relación filogenética con secuencias recuperadas de diferentes ambientes, incluyendo regiones de clima templado, el trópico y el Ártico (Rodríguez-Minguela et al., 2009).

Posiblemente, las diferentes estrategias empleadas en estos análisis hayan influido en los resultados alcanzados. La estrategia dependiente de cultivo incluyó el cultivo de microorganismos en un medio de cultivo rico (Muller-Hinton) suplementado con antibiótico (Trimetoprim), introduciendo de esta forma un fuerte sesgo en la selección de bacterias resistentes. Más aún si consideramos que los integrones de clase 1 se encuentran comúnmente asociados a genes que confiere resistencia a este antibiótico.

Por otra parte, el abordaje metagenómico mediante la construcción de bibliotecas de amplicones (*intI1*) permitió obtener una mayor diversidad de genes *intI*, recuperando integrones ambientales además de los integrones de tipo “clínico”. Asimismo, fue posible identificar secuencias IntI que formaron clados separados de integrasas conocidas, sugiriendo que se trataba de secuencias novedosas.

Sin embargo, cuando se analizaron metagenomas empleando secuenciación *shotgun* y eliminando el sesgo de los cebadores, no se obtuvo ningún integrón clase 1 (clínico o ambiental). Esto indica, por un lado, la importancia del sesgo introducido con los cebadores así como también, probablemente, la poca abundancia de genes *intI1* presentes en las muestras (tapetes).

En lo que refiere a integrasas de clase 2, únicamente dos secuencias de sedimento de la cámara séptica CS1_17527 y CS1_57033 agruparon con una secuencia IntI2 de *Shigella sonnei* (AAT78891) de origen clínico.

Los integrones clase 2 han sido documentados en bacterias patógenas de animales y de humanos y también en bacterias comensales. En la mayoría de los casos conocidos, la IntI2 no es funcional debido a la presencia de un codón stop prematuro que interrumpe el gen *intI2*. Sin embargo, en dos oportunidades se identificaron integrasas funcionales idénticas en aislamientos de *Providencia stuartii* cepa C202 (Barlow et al., 2006) y *E. coli* cepa 8157 (Marquez et al., 2008), aunque solo en esta última la funcionalidad de la integrasa fue demostrada experimentalmente. El reemplazo del codón stop ocre 179 (TAA) por un codón de glutamina (CAA) permitió la expresión de un integrón funcional. Asimismo Rodríguez-Minguela y col., (2009) analizaron la presencia de integrones clase 2 en ecosistemas naturales expuestos a diferentes grados de influencia antropogénica. En este trabajo, empleando técnicas independientes de cultivo,

encontraron once ORFs que codificarían para enzimas IntI2 potencialmente funcionales, ya que presentaban la sustitución del codón ocre 179 por el codón del residuo de glutamina .

En nuestro caso, el alineamiento parcial de IntI de CS1_17527 y CS1_57033 con secuencias IntI2 funcionales *Providencia stuartii* C202 (ABG21674.1) y de *E. coli* 8157 (EU780012.1) y no funcionales *S. sonnei* (AAT72891.1), permitió determinar que las IntI de CS1 presentaban también el residuo de glutamina en la posición 179 en lugar del codón stop comúnmente encontrado en IntI2 convencionales, sugiriendo que podría tratarse de integrones clase 2 funcionales.

Por último en lo que refiere a integrones móviles, se recuperaron ocho IntI de tapetes y sedimentos relacionadas con integrasas de clase 3 (IntI3). Estas secuencias provienen de cuatro sitios diferentes (Antártida marítima) cercanos a las Bases de operaciones, incluyendo la cámara séptica de la BCAA. Las IntI que forman este grupo exhiben cierta diversidad de secuencia entre ellas, mostrando un patrón de agrupamiento entre los diferentes sitios. Considerando que son escasos los reportes referidos a integrones de esta clase que han sido caracterizados en detalle, resulta llamativo la cantidad de integrasas clase 3 detectadas en este estudio.

La identificación de integrasas antárticas dispares a las integrasas previamente descritas es consistente con la idea de que los integrones son elementos frecuentes en el genoma bacteriano y no están presentes únicamente en ambientes clínicos o sometidos a contaminación crónica. La variedad de funciones que codifican los GCs refuerza la idea de que estos elementos genéticos están involucrados en el reclutamiento de genes que le permiten al hospedero adaptarse rápidamente a nichos ambientales diferentes e impredecibles.

3.3.4-Conclusiones

En este estudio se investigó la presencia de genes integrón integrasa (*intI*) en muestras de tapetes microbianos y sedimento/suelo de la Antártida marítima y continental. Los primers utilizados permitieron amplificar 182 *intI* que incluyeron integrones de clase 1 (*intI1*), clase 2 (*intI2*) y clase 3 (*intI3*) e integrones presumiblemente cromosómicos.

Todas las secuencias integrasas IntI1, IntI2 y una secuencia IntI3 exhibieron similitud con integrasas de origen clínico. El resto de las secuencias, incluyendo siete IntI3, se relacionaron con integrasas recuperadas de ambientes naturales de orígenes diversos.

El análisis *in silico* de metagenomas antárticos de tapetes microbianos y sedimento/suelo, empleando el programa Integron Finder permitió identificar 49 genes *intI* (31 integrones completos y 18 integrones vacíos Ino). De los 49 genes *intI*, dos correspondieron a integrones clase 2 (*intI2*) y ocho a integrones de clase 3 (*intI3*). No se identificó ningún integrón de clase 1 (*intI1*).

Por otra parte, mediante la construcción de bibliotecas de amplicones empleando el par de cebadores *intI1F-intI1R*, se recuperaron 134 secuencias integrasas IntI, de las cuales 50 correspondieron a integrasas clase 1 (IntI1). Las 84 secuencias restantes se relacionaron con integrasas de origen ambiental que exhibieron exhibiendo una gran diversidad.

El alineamiento de secuencias de aminoácidos deducidas permitió identificar todos los motivos conservados característicos de la familia de las integrón integrasas, incluyendo el dominio adicional esencial para la actividad de la enzima. Además, dos secuencias IntI2 de sedimento de la cámara séptica de la BCAA carecen del codón stop (ocre 179) frecuentemente encontrado en integrones clase 2 convencionales. La presencia de los dominios conservados y la ausencia del codón ocre sugiere que podría tratarse de integrasas funcionales.

La gran diversidad de secuencias integrasas obtenidas en este trabajo sugiere que la estrategia utilizada empleando herramientas de secuenciación masiva es el abordaje más adecuado para estudios de diversidad en comunidades microbianas de ecosistemas naturales.

Las secuencias IntI relacionadas con integrones móviles se recuperaron a partir de muestras de sitios próximos a las Bases de operaciones y por lo tanto en zonas con alto grado de influencia antropogénica.

Las *intI* que formaron grupos con integrasas de especies que portan integrones cromosómicos incluyeron secuencias de tapetes microbianos y sedimento/suelo tanto de la Antártida marítima como continental. Asimismo, las secuencias que formaron grupos únicos no relacionados filogenéticamente con integrasas conocidas, incluyeron en su mayoría integrasas identificadas a partir de bibliotecas de amplicones *intI* y en menor grado integrasas de metagenomas.

Material suplementario

1.-Tablas

Tabla S1: Sitios y coordenadas geográficas de las muestras analizadas en este estudio.

Muestra	Campaña	Posición	Sitio	Descripción
HP19	Mar/2008	S 62 13 29 W 58 57 9	Halfthree-Point	Mata marrón agua de deshielo estancada
BCH	Ene/2012	S 62 12 942 W 58 57 830	Base china	Mata roja y verde
PN	Ene/2012	S 62 10 96,5 W 58 51 86,5	s/d	Mata roja
ALC	Mar/2012	S 62 11 21,9 W 58 55 45,8	Lago congelado (tanques rusos)	Agua lago congelado

CN11	Ene/2013	S 62 11 13,6 W 58 54 48,8	Caleta Norma	Sedimento asociado a briofitas
GC5	Ene/2013	S 62 10 58 W 58 53 24	Glaciar Collins	Cañada de deshielo
LCTR1	Ene/2013	S 62 11 15 W 58 55 06	Lago congelado (tanques rusos)	Hielo, sedimento
IA1	Ene/2013	S 62 12 45,9 W 58 54 57,3	Isla Ardley	Nieve con coloración verde-rojiza (próximo al refugio)
IA2	Ene/2013	S 62 12 43 W 58 55 14	Isla Ardley	Nieve con coloración verde-rojiza
IA3	Ene/2013	S 62 12 45,5 W 58 55 1,2	Isla Ardley	Cañada de deshielo (próximo al refugio)
IA4	Ene/2013	S 62 12 43 W 58 55 1,8	Isla Ardley	Sedimento con fecas de pingüino y musgo
22.5	Ene/2014	S 62 10.701 W58 54.662	Lagos aledaños Caleta Norma	Hielo con algas rojizas
23.1	Ene/2014	S 62 10 42.4 W 58 55 12.5	Pasaje Drake	Mata roja con sedimentos
23.3	Ene/2014	S 62 10 42.4 W 58 55 12.5	Drake	Arena, agua de mar
23.5	Ene/2014	S 62 09 38 W 58 56 35,7	Drake	Agua de mar, algas rojas y sedimento
23.8	Ene/2014	S 62 10 03.5 W 58 55 36.0	Pasaje Drake	Mata marrón-rojiza, cañada sobre pedregullo
23.9	Ene/2014	s/d	Drake	Mata roja y sedimento
24.1	Ene/2014	s/d	Camino a tanques rusos	Agua lago congelado
24.2	Ene/2014	S 62 11 25,4 W 58 55 04	Caleta Norma	Agua de mar y arena
24.3	Ene/2014	S 62 11 31 W 58 55 33,1	Caleta Norma	Agua de mar y arena
24.4	Ene/2014	W 62 11 39.2 S 58 56 16.2	Tanques rusos	Cañada deshielo
26.2	Ene/2014	S 62 12 47.9 W58.54 58.6	Isla Ardley	Hielo con algas verdes
26.4	Ene/2014	S 62 12 45.4 W 58 55 16.3	Isla Ardley	Mata negra con musgo y sedimento
26.5	Ene/2014	S 62 12 45.4 W58 55 16.3	Isla Ardley	Mata negra con musgo
26.7	Ene/2014	S 62 12 44,6 W 58 55 19,8	Isla Ardley	Agua de cañada y musgo
26.8	Ene/2014	S 62 12 36,5 W 58 55 38,7	Isla Ardley	Matas y sedimento
HP 2.8	Mar/2014	S 62 13 48 W 58 58 54	Halfthree-Point	Sedimento
HP 2.9	Mar/2014	S 62 13 48	Halfthree-Point	Sedimento

		W 58 58 54		
HP 2.10	Marzo/2014	S 62 13 48 W 58 58 54	Halfthree-Point	Sedimento
HP 3.2	Mar/2014	S 62 13 48 W 58 58 54	Halfthree-Point	Sedimento
HP 3.3	Mar/2014	S 62 13 48 W 58 58 54	Halfthree-Point	Sedimento
HP 4	Mar/2014	S 62 13 48 W 58 58 54	Halfthree-Point	Sedimento
LR1	Abr/2015	S 62 11 28,39 W 58 58 37,70	Lago cercano a Base Rusa.	Agua de deshielo
PI2	Abr/2015	S 62 11 35,01 W 58 58 29,27	Aeropuerto Base Chilena	Agua de deshielo
BD3	Abr/ 2015	S 62 11 7,91 W 58 59 35,16	Costa Drake	Agua de mar
IN4	Abr/2015	S 62 11 3,72 W 58 59 29,77	Aeródromo Base Chile	Agua de deshielo
PB5	Abr/ 2015	S 62 11 10,3 W 58 54 11,3	BCAA	Agua de mar
LU6	Abr/ 2015	S 62 11 9,9 W 58 54 50,6	Lago Uruguay	Agua dulce
CD7	Abr/2015	S 62 09 26,4 W 58 56 13	Costa Drake	Agua de mar
GC5	Abr/2015	S 62 10 58 W 58 53 24	Glaciar Collins	Cañada de deshielo
PD2a	Abr/2015	S 62 09 03 W 58 56 27	Pasaje Drake	Agua de mar
FE	Abr/2015	S 62 11 35,01 W 58 58 29,27	Aeródromo Base Chile	Sedimento
CH39	Abril/2015	s/d	Isla Rey Jorge	Sedimento
PD2b	Abril/2015	S 62 09 03 W 58 56 27	Pasaje Drake	Sedimento
IA1M	Abril/2015	S 62 12 35,78 W 58 55 41,83	Isla Ardley	Sedimento, agua de deshielo briofitas y fecas
IA2M	Abril/2015	S 62 12 35,83 W 58 56 27	Isla Ardley	Sedimento y agua de deshielo
IA3M	Abril/2015	S 62 12 36,88 W 58 55 48,54	Isla Ardley	Sedimento, agua de deshielo briofitas
TR1	Abril/2015	S 62 11' 38,32'' W 58 56' 15,6''	Drake	Sedimento (tanques combustible)
TC2	Abril/2015	S 62 11 05,8 W 58 58 20,0	Bahía Collins	Sedimento (tanques combustible)
IA1S	Enero/2016	s/d	Isla Ardley	Costa próximo a la pingüinera

IA2S	Enero/2016	s/d	Isla Ardley	Agua de deshielo, tapetes y fecas.
IA1K	Febrero/2016	S 62 12 36,8 W 58 56 27,5	Isla Ardley	Tapete microbiano (agua deshielo)
D1K	Febrero/2016	S 62 09 53,8 W 58 56 12,8	Drake	Matas microbianas
PN1K	Febrero/2016	S 62 10`57,3 W 58 52 00,4	Punta Nebles	Lago de deshielo y sedimento
PN2K	Febrero/2016	S 62 11 3 W 58 51 30	Punta Nebles	Lago sobre piedras
CP1P	Febrero/2016	S 62 14 35 W 58 39 99	Caleta Potter (Base Carlini)	Matas microbianas
CP2P	Febrero/2016	S 62 14 35 W 58 39 99	Caleta Potter (Base Carlini)	Matas microbianas
CP3P	Febrero/2016	S 62 14 35 W 58 39 99	Caleta Potter	Sedimento
PE4P	Febrero/2016	S 62 14 34 W 58 39 86	Punta Elefante	Sedimento
CP5P	Febrero/2016	S 62 14 35 W 58 39 99	Caleta Potter (Base Carlini)	Matas microbianas
CP6P	Febrero/2016	S 62 14 18.3 W 58 40 27	Caleta Potter (Laguna Glaciar)	Sedimento
CP7P	Febrero/2016	S 62 14 34 W 58 39 86	Caleta Potter (Base Carlini)	Sedimento
BCAA3	Febrero/2016	S 62 11 1,6 W 58 54 21,9	Base Artigas	Sedimento

Tabla S2: Sitios y coordenadas geográficas de los metagenomas de tapetes microbianos y sedimento/suelo analizados en este estudio.

ANCO = Antártida continental

Metagenomas	Sitio de muestreo	Tipo de muestra	Coordenadas geográficas	Coordenadas geográficas	Mapa
Drake 1	Drake 1 (abril/2015)	Tapete microbiano (cañada de deshielo)	62° 9'31.7" S	58° 56'31.5" W	2A
Espejo 2	Espejo (abril/2015)	Tapete microbiano (cañada de deshielo)	62° 9'50.002" S	58° 58'46.552" W	2A
HTP 3	Haltthree-Point 3 (abril/2015)	Tapete microbiano	62° 11'59.918" S	58° 57'2.659" W	2A

		(cañada de deshielo)			
Pista 4	Pista (abril/2015)	Tapete microbiano (cañada de deshielo)	62° 10' 8.663" S	58° 58' 58.903" W	2A
Carlini I 5	Carlini I (abril/2015)	Tapete microbiano (cañada de deshielo)	62° 14' 35" S	58° 39' 99" W	2A
Carlini II 6	Carlini II (abril/2015)	Tapete microbiano (cañada de deshielo)	62° 14' 34" S	58° 39' 86" W	2A
Esperanza 7	Esperanza (abril/2015)	Tapete microbiano (cañada de deshielo)	63° 24' 280' S	56° 59' 825' W	2B
Primvaera 8	Primavera (abril/2015)	Tapete microbiano (cañada de deshielo)	64° 09' 22.5" S	60° 57' 30.7" W	2B
San Martín 9	San Martín (bril/2015)	Tapete microbiano (cañada de deshielo)	68° 07' 45.66" S	67° 06' 02.74" W	2B
ANCO 10 B012 - 2015 - 14	B012-2015-14 (abril/2015)	Tapete microbiano (cañada de deshielo)	78° 01' 399" S	63° 55' 049" W	2C
ANCO 11 B012 - 2015 - 15	B012-2015-15 (abril/2015)	Tapete microbiano (cañada de deshielo)	78° 01' 389" S	63° 54' 938" W	2C
ANCO 12 B012 - 2015 - 16	B012-2015-16 (abril/2015)	Tapete microbiano (cañada de deshielo)	78° 01' 380" S	63° 54' 117" W	2C
ANCO 13 B012 - 2015 - 16 Mid Good	B012-2015 Mid Garwood (abril/2015)	Tapete microbiano (cañada de deshielo)	78° 01' 504" S	64° 06' 034" W	2C
ANCO 14 B012 - 2015 - 18	B012-2015-18 (abril/2015)	Tapete microbiano (cañada de deshielo)	77° 39.66" S	63° 05.511" W	2C
CSC1	Cámara séptica BCAA (abril/2017)	Sedimento/suelo	62° 11' 3,5' S	58° 54' 19" W	2A
HP3	Halfthree-point (abril/2017)	Sedimento/suelo	62° 13,2' 9" S	58 57,9" W	2A

IA6	Isla Ardely (abril/2017)	Sedimento/suelo	62° 12,4' 3"S	58° 55,1' 4 "W	2A
-----	-----------------------------	-----------------	---------------	----------------	----

Tabla S3: Datos de secuenciación obtenidos para los metagenomas de tapetes microbianos. que tuvieron un promedio de $7,8 \pm 2$ Gb que representan 109 Gpb de ADN secuenciado. Los datos crudos pasaron por un filtro de calidad y fueron ensamblados de *ново* empleando IDBA-UD, con secuencias k-mer de entre 120-150 de largo (Peng et al. 2012).

Muestra	Tamaño	Nº de secuencias	Nº de pb	Largo promedio	% contenido GC
Drake Metagenoma 1	25,7 MB	14345	26.277.856	1,832 ± 4,459 bp	46 ± 9 %
Espejo Metagenoma 2	23,6 MB	11311	24.258.161	2,145 ± 3,264 bp	49 ± 11 %
HTP3 Metagenoma 3	28,1 MB	17047	27,221,748	1,597 ± 2,892 bp	42 ± 12 %
Pista Metagenoma 4	42,4 MB	19554	40,889,758	2,091 ± 4,755 bp	43 ± 14 %
Carlini I Metagenoma 5	107,1 MB	62551	98,991,768	1,583 ± 2,627 bp	40 ± 15 %
Carlini II Metagenoma 6	81,4 MB	47446	83,192,253	1,753 ± 2,171 bp	42 ± 8 %
Esperanza Metagenoma 7	68,1 MB	35849	69,755,911	1,946 ± 4,666 bp	46 ± 10 %
Primvaera Metagenoma 8	82,2 MB	53174	83.777.672	1,576 ± 2,762 bp	46 ± 11 %
San Martín Metagenoma 9	64,9 MB	17447	67,226,753	3,853 ± 10,671 bp	46 ± 12 %
Bo12 - 2015 - 14 ANCO 10	49,5 MB	10089	47,693,188	2,637 ± 5,037 bp	52 ± 18 %
Bo12 - 2015 15 ANCO 11	73,1 MB	24919	75.528.956	3,031 ± 5,222 bp	44 ± 12 %
Bo12 - 2015 - 16 ANCO 12	54,9 MB	23874	56,525,696	2,368 ± 2,992 bp	50 ± 12 %
Bo12 - 2015 - 16MidGood ANCO 13	56,6 MB	23997	58,290,664	2,429 ± 4,994 bp	47 ± 11 %
Bo12 - 2015 - 18 ANCO 14	98,1 MB	34862	101,298,892	2,906 ± 6,350 bp	38 ± 8 %
HP3	32,5 MB	33,072,408 bp	1,154 ± 1,158 bp	51 ± 11 %	28644

IA6	70,7 MB	72,183,621 bp	1,258 ± 1,705 bp	42 ± 10 %	57342
-----	---------	---------------	------------------	-----------	-------

Tabla S4: Datos de secuenciación obtenidos para los metagenomas de sedimento/suelo.

Muestra	Nº total de bases (pb)	Nº total de reads	% GC	% AT	Q20 (%)	Q30(%)
IA6	2,310,652,548	22,877,748	42,81	57,18	94,71	91,60
HTP3	2,261,211,028	22,388,228	56,46	43,54	94,30	90,92
CS1	2,711,260,968	26,844,168	47,73	52,26	94,11	90,60

Tabla S5: Se describen los aislamientos obtenidos en medio R2A y medio marino empleados para la búsqueda de integrones clase 1,2 y 3, utilizando los cebadores descriptos por Mazel y col. (2000).

Muestra ID	Obs.	Aislamientos	PCR <i>intI1</i>	PCR <i>intI2</i>	PCR <i>intI3</i>
22.5	Hielo con algas rojas	1,2,3,4,5,6,7,8	-	-	-
23.1	Mata roja	1A,2A,3A,4A,5A,6A	+	-	-
23.3	Mata roja y sedimento (Pasaje Drake)	2A,3A,4A,5A,1B,2B, 1C,2C,4C,2D,3D, 5D,1E,2E,3E,5E,1F,2F,3F,4F,5F,1G	-	n/d	n/d
23.5	Algas rojas y sedimento	1A,2A,3B,4C,5C,6C,2D	+	-	-

23.8	Mata marrón-rojiza, cañada sobre pedregullo	1,2,3,4,5,6,7,8	-	-	-
23.9	Mata roja y sedimento	1B,2B,3B,2C,3C	-	-	-
24.1	Lago congelado (Tanques rusos)	1A,2A,3A, 4B,5B,6B, 7C,8C,9C,10D,11D,12D 13E,14E,16E, 17F,18F,19F,17F,18F,19F,20F,25G,26G,27H,28H,29H	-	n/d	n/d
24.3	Agua de mar y arena (Caleta Norma)	1A,2A,3A,1B2,4B	-	-	-
26.2	Nieve y algas verdes (Isla Ardley)	1A,2A,3A,4A,2B,5B,6B,7B,8B	-	-	-
26.5	Mata negra con musgo (Isla Ardley)	3A,4A,5A,6A,7A,8A	-	-	-
26.7	Agua de cañada	1,2,3,4,5,6,7,8,9,10	-	n/d	n/d
26.8	Matas y sedimento	1C,2C,3C,4C,1D,2D,4D,2E,3E	-	n/d	n/d
IA ₃	Cañada de deshielo próximo al refugio (Isla Ardley)	1A,2A,3A,4A,5A ,6B,7,9B,10B,11B 12C,13C,117D,18D,19D,20D,21D,4C,15C,16C 22E,23E,24E,25E,26E 27F,28F,29F	-	-	-
HP2.10	Sedimento	4,5,6,7	-	-	-
HP2.9	Sedimento	2,4,7	-	-	-
HP2.10	Sedimento	1,2,3,4,5	-	-	-
HP3.2	Sedimento	1,2,3,4,5	-	-	-
HP4	Sedimento (Halfthree-Point)	1A,2A,3A,4A,5A, 1I,2I,4I,14J,15J,25J	-	-	-
ALC	Agua lago congelado	2,3,4,6,7	-	-	-
BCH	Agua de mar	1,2,3,4,5,6,7,8,9	-	-	-
PN	Agua dulce	1,2,3,4,5,6,7,8,9	-	-	-
SLC	Sedimento lago congelado	1,2,3	-	-	-
LR1	Agua deshielo	1D,1E	-	+	-
PI2	Agua deshielo	2A, 2B,2C	-	+	-
BD3	Agua de mar	1H,2H	-	-	-
IN4	Agua de deshielo	1G,2G,3G	-	-	-
PB5	Agua de mar (BCAA)	1F,2F	-	+	-
LU6	Agua dulce	1I,2I	-	+	-
CD7	Agua de mar	7J	-	-	-

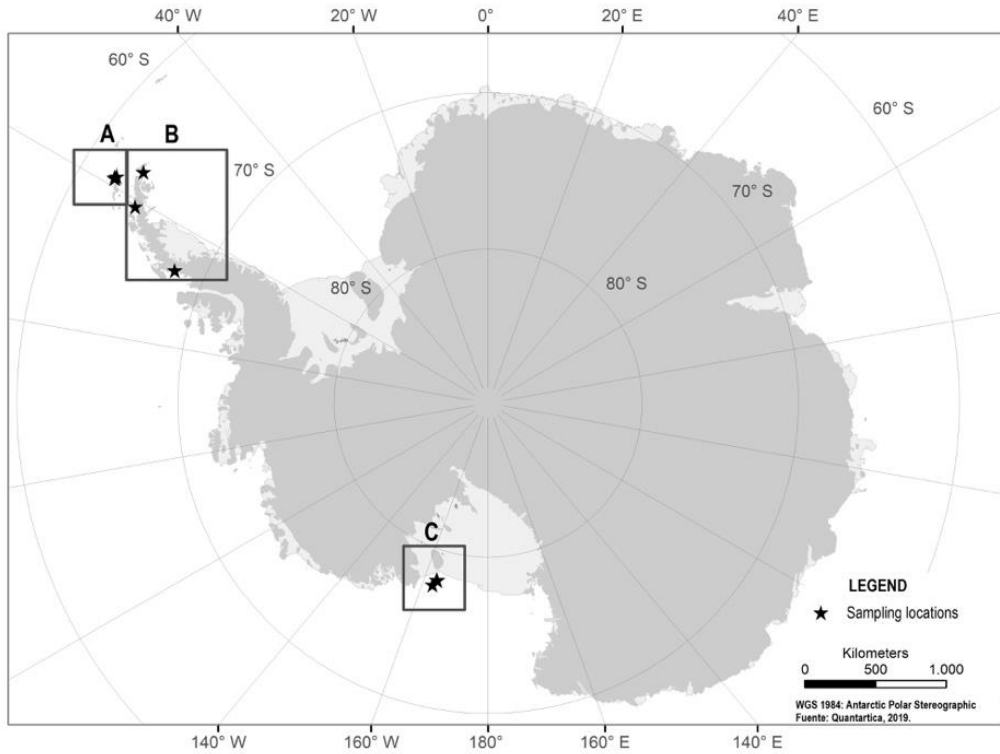
Tabla S6: Se resume la familia de genes de resistencia anotados por Resfam y el tipo de mecanismo de acción para cada uno de los genes identificados por el programa Integron Finder.

Muestra	ID Resfams	Familia/Descripción	Clasificación del antibiótico/Mecanismos de acción
Espejo2_7271	RF0067	Emr 23S ribosomal ARN metiltransferasa/ rRNA metiltransferasa confiere resistencia antibiótica	RNA Metiltransferasa
Espejo2_866	RF0104	MFS/facilitador mayor superfamilia (MFS), bomba de eflujo de antibiótico	transportador MFS
HTP3_2	RF0023	adeS / Gen modulador eflujo del antibiótico regulador AdeABC	Otros eflujos
Carlini I_1260	RF0023	adeS / Gen modulador eflujo del antibiótico regulador AdeABC	Otros eflujos
Carlini I_6546	RF0155	vanS / regulador transcripcional de genes de resistencia a van glicopéptidos	Resistencia a glicopéptidos/Gen modulador de resistencia
Carlini II_13063	RF0074	DNA topoisomerasa resistencia a fluoroquinolonas	Resistencia a quinolonas
Carlini II_2017	RF0023	adeS/ Gen modulador eflujo del antibiótico regulador AdeABC	Otros eflujos
Carlini II_2316	RF0074	DNA topoisomerasa resistencia a fluoroquinolonas	Resistencia a quinolonas
Carlini II_7179	RF0114	ramA/gen modulador del eflujo antibiótico	Gen modulador de resistencia
Esperanza 7_1205	RF0113	Qnr: proteína de resistencia a quinolonas	Resistencia a quinolonas
Primavera 8_27	RF0101	mexX Subunidad de bomba de eflujo que confiere resistencia a antibiótico	RND eflujo del antibiótico
ANCO10_124	RF0154	VanR: activador transcripcional	Resistencia a glicopéptidos/Gen modulador de resistencia

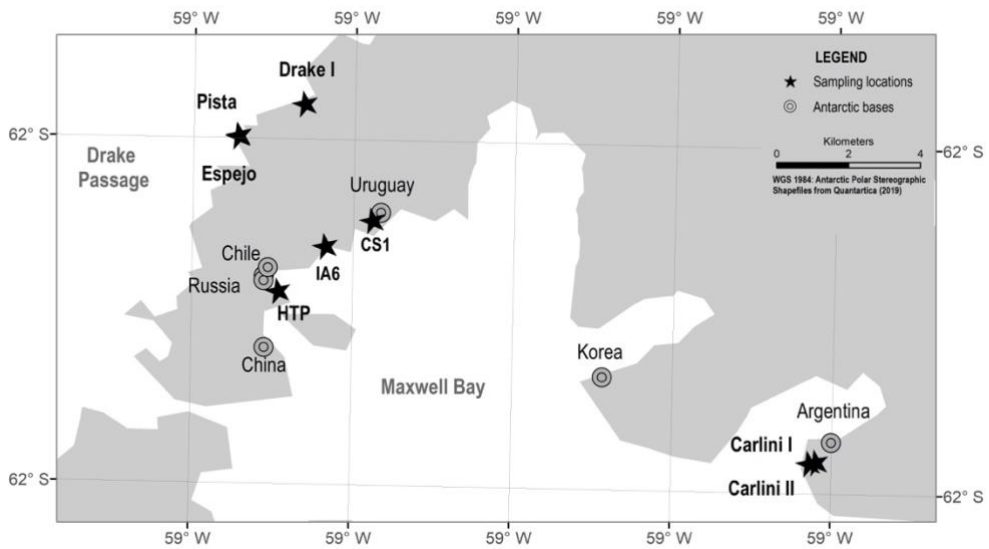
		regulador de VanA, VanH y VanX	
ANCO10_7316	RF0003	Aminoglicósido acetiltransferasa AAC3-I	Enzima modificadora de aminoglicósidos/Acetiltransferasa
ANCO11_3257	RF0033	Aminoglicósido fosfotransferasa APH3	Enzima modificadora de aminoglicósidos/Fosfotransferasa
ANCO12_133	RF0155	vanS regulador transcripcional de genes de resistencia a van glicopéptidos	Regulador transcripcional de van glicopéptidos genes de resistencia
ANCO12_3311	RF0155	vanS regulador transcripcional de genes de resistencia a van glicopéptidos	Resistencia a glicopéptidos/Gen modulador de resistencia
ANCO12_3319	RF0107	msbA cassette de unión -ATP (ABC) bomba de eflujo del antibiótico	Transportador ABC
ANCO12_8028	RF0155	vanS regulador transcripcional de genes de resistencia a van glicopéptidos	Resistencia a glicopéptidos/Gen modulador de resistencia
ANCO12_8785	RF0066	emrE Pequeña bomba de resistencia multidroga, bomba de eflujo de antibiótico (SMR)	Otros eflujos
ANCO14_2801	RF0153	vanH Alfa ceto-ácido D- específica deshidrogenasa que sintetiza D-lactasa	Resistencia a glicopéptidos
HP3_1193	RF0052	Cloramfenicol fosfotransferasa_CPT	Resistencia a cloramfenicol/Fosfotransferasa
HP3_492	RF0091	marA/ Factor de transcripción que induce la bomba de eflujo MDR AcrAB	Gen modulador de resistencia
CS1_503	RF0056	Clase D/Clase D- betalactamasa	Beta-lactamasa
CS1_2587	RF0066	emrE Pequeña bomba de resistencia multidroga, bomba de eflujo de antibiótico (SMR)	Otros eflujos
CS1_61281	RF0066	emrE Pequeña bomba de resistencia multidroga,	Otros eflujos

		bomba de eflujo de antibiótico (SMR)	
CS1_7893	RF0027	ANT3 Aminoglucósido nucleotidiltransferasa	Enzima modificadora de aminoglucósidos/Nucleotidiltransferasa

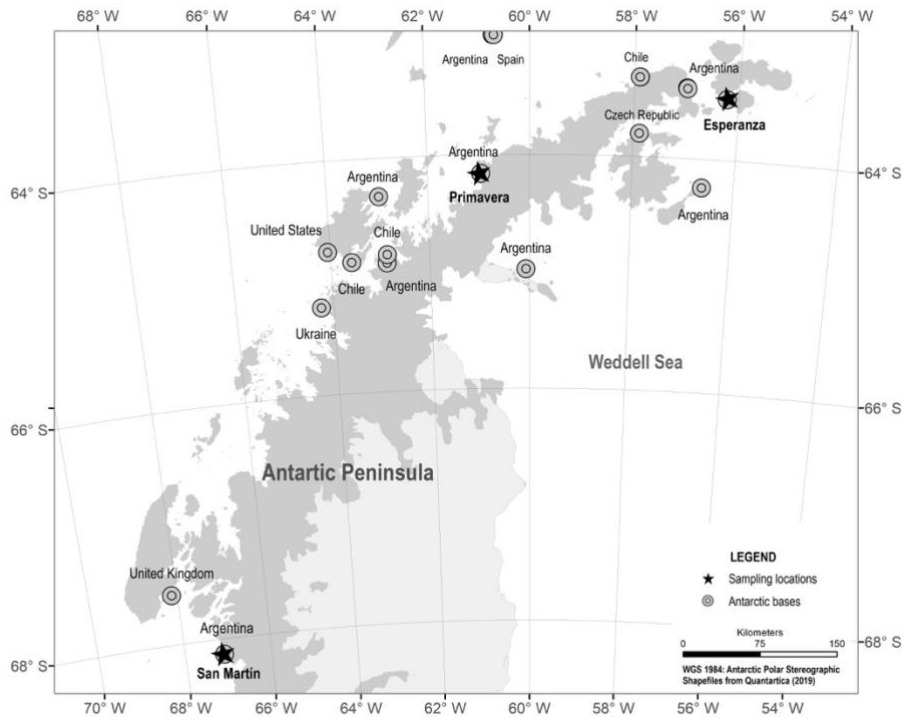
2.-Sitios de muestreo



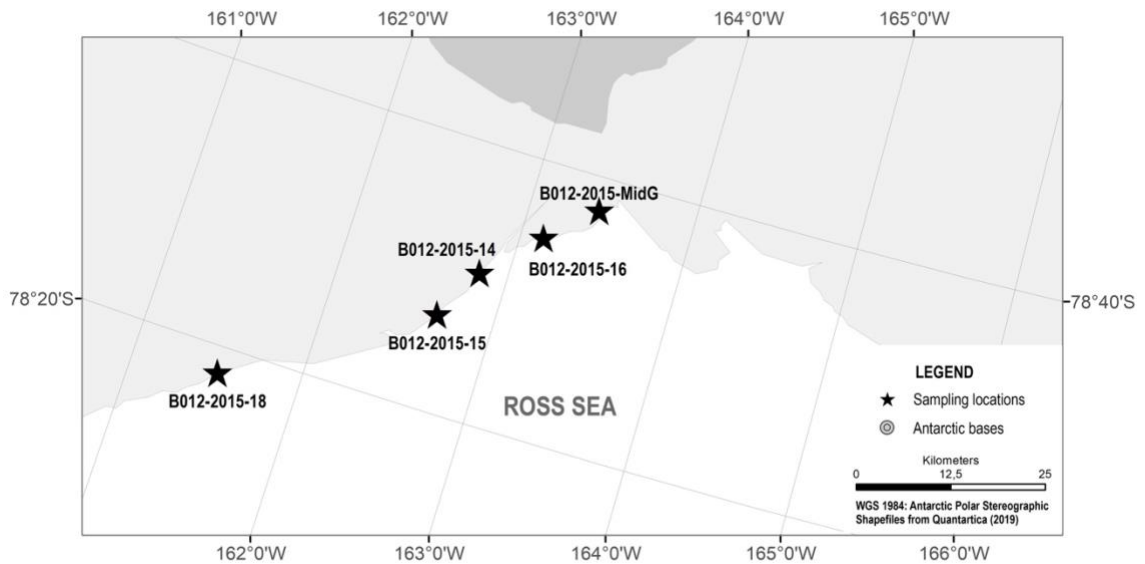
Mapa 1: Sitios de muestreo en el continente y Península Antártica.



Mapa 2A: Isla Rey Jorge



Mapa 2B: Península Antártica



Mapa 2C: Continente Antártico

Matas microbianas y briofitas



3.- Composición taxonómica de la comunidad

Figura S1: Composición de la comunidad de tapetes microbianos a nivel de clase. A) Isla Rey Jorge. B) Península Antártica. C) Continente Antártico.

Figura S1: A) Isla Rey Jorge

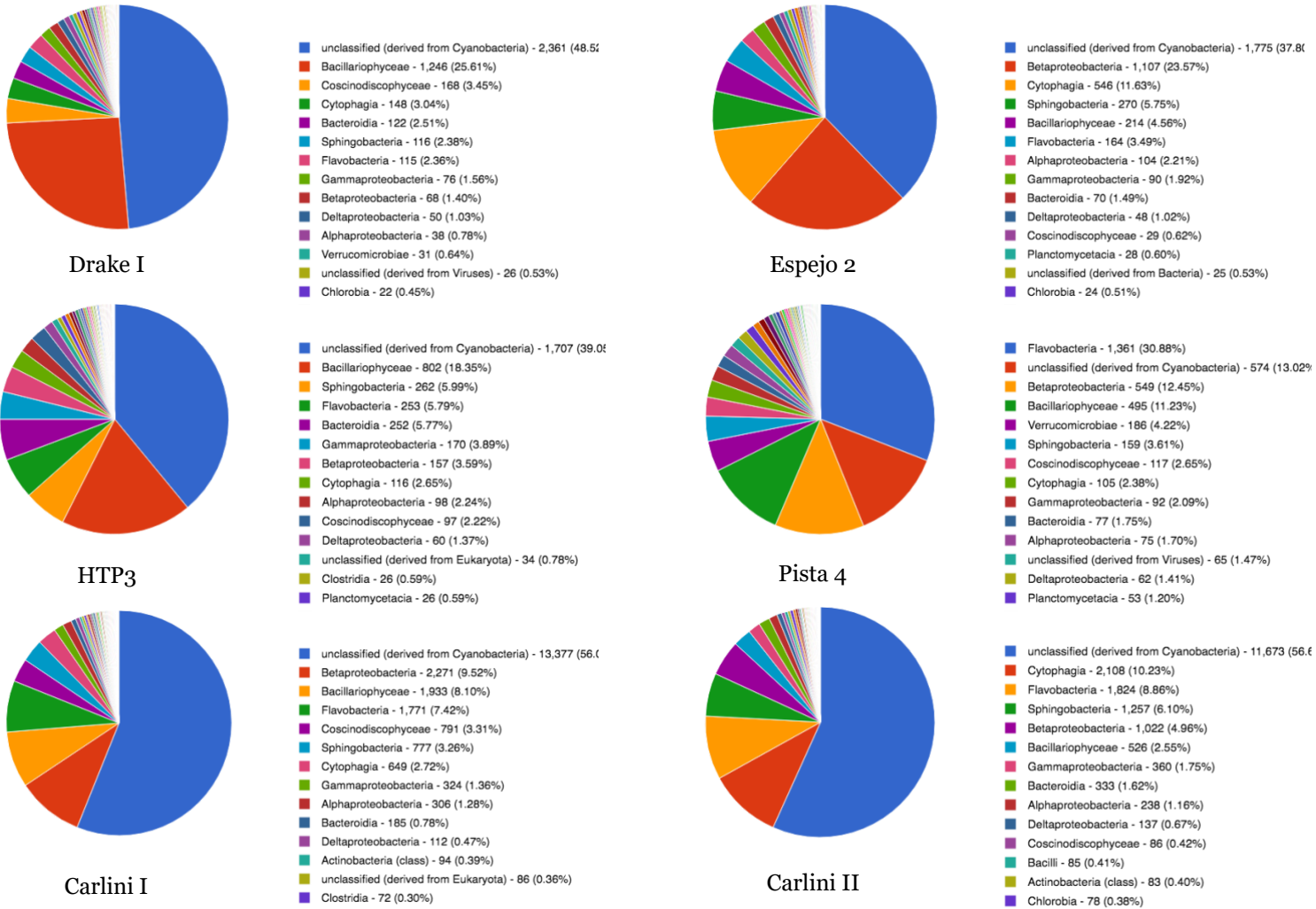
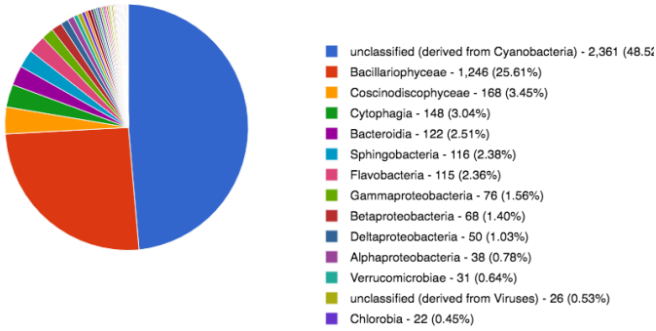
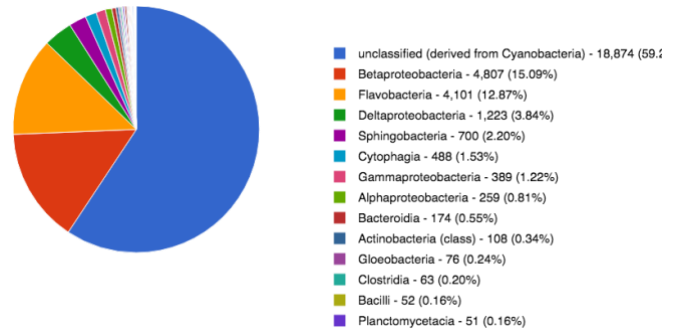


Figura S1 B) Península Antártica

Esperanza 7



Primavera 8



San Martín 9

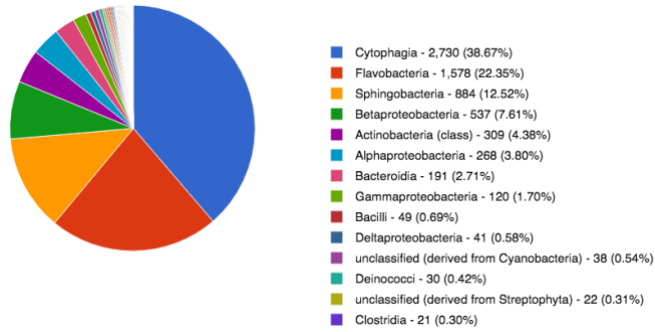
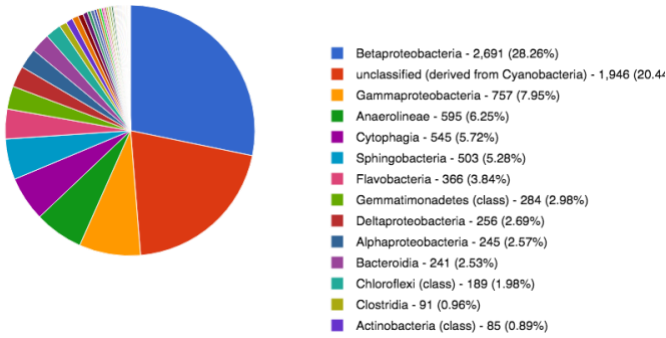
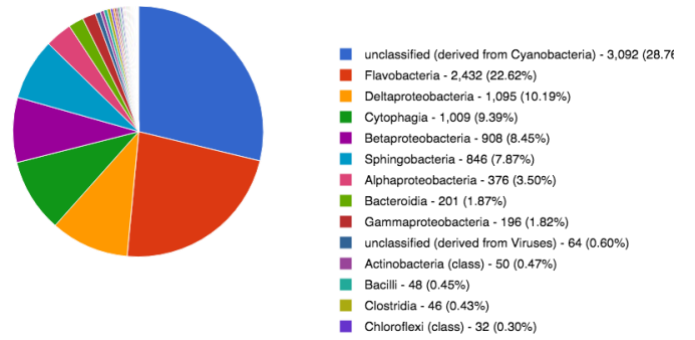


Figura S1 C) Continente Antártico

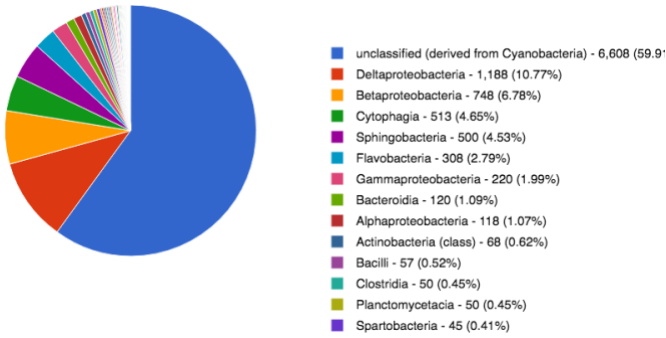
Bo12 - 2015 - 14



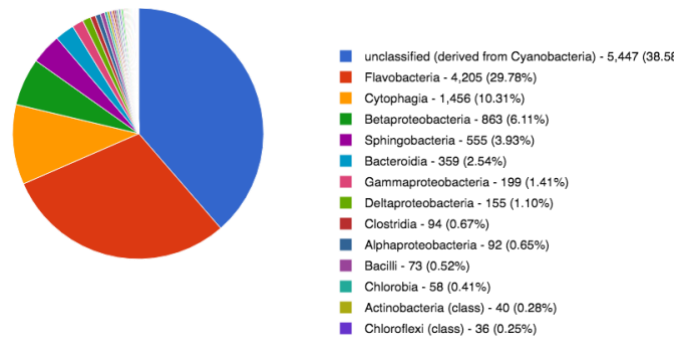
Bo12 - 2015 - 15



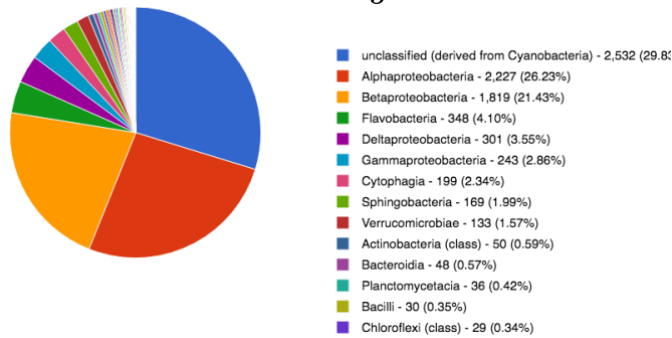
Bo12 - 2015 - 16



Bo12 - 2015 - 17 MidGood



Bo12 - 2015 - 18



4.-Composición de medios de cultivo

Medio marino	Composición g/L
Bactopeptona	5
Extracto de levadura	1
Sales marinas	40
Na ₂ HPO ₄ (g)	0,008
NH ₄ NO ₃ (g)	0,0016
Citrato férrico de amonio	0,10
Agar	15
pH 7	7,6

Medio R2A	Composición g/L
Extracto de levadura	0,5
Proteosa peptona N ^o 3	0,5
Casamino ácidos	0,5
Glucosa	0,5
Almidón soluble	0,5
K ₂ HPO ₄	0,3
Piruvato de sodio	0,3
MgSO ₄ .7H ₂ O	0,05
Agar	15,0

Referencias Bibliográficas

- Alcántara-Hernández, R. J., Centeno, C. M., Ponce-Mendoza, A., Batista, S., Merino-Ibarra, M., Campo, J., & Falcón, L. I. (2014). Characterization and comparison of potential denitrifiers in microbial mats from King George Island, Maritime Antarctica. *Polar Biology*, 37(3), 403–416. <http://doi.org/10.1007/s00300-013-1440-3>
- Anantharaman, K., Brown, C. T., Hug, L. A., Sharon, I., Castelle, C. J., Probst, A. J., ... Banfield, J. F. (2016). Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nature Communications*, 7, 13219. <http://doi.org/10.1038/ncomms13219>
- Antelo, V., Guerout, A. M., Mazel, D., Romero, V., Sotelo-Silveira, J., & Batista, S. (2018). Bacteria from Fildes Peninsula carry class 1 integrons and antibiotic resistance genes in conjugative plasmids. *Antarctic Science*, 30(1), 22–28. <http://doi.org/10.1017/S0954102017000414>
- Antelo, V., Romero, H., & Batista, S. (2015). Detection of integron integrase genes on King George Island, Antarctica. *Advances in Polar Science*, 26(1), 30–37. <http://doi.org/10.13679/J.ADVPS.2015.1.00030>
- Arakawa, Y., Murakami, M., Suzuki, K., Ito, H., Wacharotayankun, R., Ohsuka, S., ... Ohta, M. (1995). A novel integron-like element carrying the metallo- β -lactamase gene bla(IMP). *Antimicrobial Agents and Chemotherapy*, 39(7), 1612–1615. <http://doi.org/10.1128/AAC.39.7.1612>
- Aziz, R. K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R. A., ... Zagnitko, O. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*, 9, 1–15. <http://doi.org/10.1186/1471-2164-9-75>
- Baharoglu, Z., Bikard, D., & Mazel, D. (2010). Conjugative DNA transfer induces the bacterial SOS response and promotes antibiotic resistance development through integron activation. *PLoS Genetics*, 6(10), 1–10. <http://doi.org/10.1371/journal.pgen.1001165>
- Baharoglu, Z., Garriss, G., & Mazel, D. (2013). Multiple Pathways of Genome

- Plasticity Leading to Development of Antibiotic Resistance. *Antibiotics*, 2(2), 288–315. <http://doi.org/10.3390/antibiotics2020288>
- Baker Jennifer, Riester, C., Zhao, T., Swingley, W., Sherman, C., Asao, M., ... Blankenship, R. (2017). Genome Sequence of *Rhodospirillum rubrum* ANT.BRT; A Psychrophilic Purple Nonsulfur Bacterium from an Antarctic Microbial Mat. *Microorganisms*, 5(1), 8. <http://doi.org/10.3390/microorganisms5010008>
- Barlow, R. S., & Gobius, K. S. (2006). Diverse class 2 integrons in bacteria from beef cattle sources. *Journal of Antimicrobial Chemotherapy*, 58(6), 1133–1138. <http://doi.org/10.1093/jac/dkl423>
- Biskri, L., Bouvier, M., Guérout, A.-M., Boissard, S., & Mazel, D. (2005). Comparative study of class 1 integron and *Vibrio cholerae* superintegron integrase activities. *Journal of Bacteriology*, 187(5), 1740–50. <http://doi.org/10.1128/JB.187.5.1740-1750.2005>
- Biskri, L., & Mazel, D. (2003). Erythromycin esterase gene *ere(A)* is located in a functional gene cassette in an unusual class 2 integron. *Antimicrobial Agents and Chemotherapy*, 47(10), 3326–3331. <http://doi.org/10.1128/AAC.47.10.3326-3331.2003>
- Blakely, G. W. (2014). *Mechanisms of Horizontal Gene Transfer and DNA Recombination. Molecular Medical Microbiology: Second Edition* (Vol. 1–3). Elsevier Ltd. <http://doi.org/10.1016/B978-0-12-397169-2.00015-9>
- Bölter, M. (2011). Soil development and soil biology on King George Island. *Journal of Soil Science*, 32(2), 105–116. <http://doi.org/10.2478/v10183>
- Boto, L. (2010). Horizontal gene transfer in evolution: facts and challenges. *Proceedings of the Royal Society B: Biological Sciences*, 277(1683), 819–827. <http://doi.org/10.1098/rspb.2009.1679>
- Boucher, Y., Labbate, M., Koenig, J. E., & Stokes, H. W. (2007). Integrons: mobilizable platforms that promote genetic diversity in bacteria. *Trends in Microbiology*, 15(7), 301–9. <http://doi.org/10.1016/j.tim.2007.05.004>
- Bouvier, M., Ducos-Galand, M., Loot, C., Bikard, D., & Mazel, D. (2009). Structural features of single-stranded integron cassette *attC* sites and their

- role in strand selection. *PLoS Genetics*, 5(9).
<http://doi.org/10.1371/journal.pgen.1000632>
- Boyd, E. F., Almagro-Moreno, S., & Parent, M. A. (2009). Genomic islands are dynamic, ancient integrative elements in bacterial evolution. *Trends in Microbiology*, 17(2), 47–53. <http://doi.org/10.1016/j.tim.2008.11.003>
- Brown, H. J., Stokes, H. W., & Hall, R. M. (1996). The integrons In0, In2, and In5 are defective transposon derivatives. *Journal of Bacteriology*, 178(15), 4429–4437. <http://doi.org/10.1128/jb.178.15.4429-4437.1996>
- Buchfink, B., Xie, C., & Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12(1), 59–60.
<http://doi.org/10.1038/nmeth.3176>
- Callejas, C., Azziz, G., Souza, E. M., Gill, P. R., & Batista, S. (2018). Prokaryotic diversity in four microbial mats on the Fildes Peninsula, King George Island, maritime Antarctica. *Polar Biology*, 1.
<http://doi.org/10.1007/s00300-018-2256-y>
- Callejas, C., Gill, P. R., Catalán, A. I., Azziz, G., Castro-Sowinski, S., & Batista, S. (2011). Phylotype diversity in a benthic cyanobacterial mat community on King George Island, maritime Antarctica. *World Journal of Microbiology and Biotechnology*, 27(6). <http://doi.org/10.1007/s11274-010-0578-1>
- Cambray, G., Guerout, A.-M., & Mazel, D. (2010). Integrons. *Annual Review of Genetics*, 44(1), 141–166. <http://doi.org/10.1146/annurev-genet-102209-163504>
- Cambray, G., Sanchez-alberola, N., Campoy, S., Guerin, É., Re, S. Da, González-zorn, B., ... Erill, I. (2011). Prevalence of SOS-mediated control of integron integrase expression as an adaptive trait of chromosomal and mobile integrons. *Mobile DNA*, 2(1), 6. <http://doi.org/10.1186/1759-8753-2-6>
- Carattoli, A. (2009). Resistance Plasmid Families in. *Society*, 53(6), 2227–2238.
<http://doi.org/10.1128/AAC.01707-08>
- Carattoli, A., Aschbacher, R., March, A., Larcher, C., Livermore, D. M., & Woodford, N. (2010). Complete nucleotide sequence of the IncN plasmid pKOX105 encoding VIM-1, QnrS1 and SHV-12 proteins in

- Enterobacteriaceae from Bolzano, Italy compared with IncN plasmids encoding KPC enzymes in the USA. *Journal of Antimicrobial Chemotherapy*, 65(10), 2070–2075. <http://doi.org/10.1093/jac/dkq269>
- César, C. E., Machón, C., De La Cruz, F., & Llosa, M. (2006). A new domain of conjugative relaxase TrwC responsible for efficient oriT-specific recombination on minimal target sequences. *Molecular Microbiology*, 62(4), 984–996. <http://doi.org/10.1111/j.1365-2958.2006.05437.x>
- Chavda, K. D., Chen, L., Fouts, D. E., Sutton, G., Brinkac, L., Jenkins, S. G., ... Kreiswirth, B. N. (2016). Comprehensive Genome Analysis of Carbapenemase-Producing *Enterobacter* spp.: New Insights into Phylogeny, Population Structure, and Resistance Mechanisms. *MBio*, 7(6). <http://doi.org/10.1128/mBio.02093-16>
- Checroun, C., & Gutierrez, C. (2004). σ -Dependent regulation of yehZYXW, which encodes a putative osmoprotectant ABC transporter of *Escherichia coli*. *FEMS Microbiology Letters*, 236(2), 221–226. <http://doi.org/10.1016/j.femsle.2004.05.046>
- Claridge, I. B. C. (*ü) and G. G. C. (2009). Permafrost Soils, *Soil Biology*. *Permafrost Soils, Soil Biology*, 78, 175–183. <http://doi.org/10.1007/978-3-540-69371-0>
- Collis, C. M., Grammaticopoulos, G., Briton, J., Stokes, H. W., & Hall, R. M. (1993). Site-specific insertion of gene cassettes into integrons. *Molecular Microbiology*, 9(1), 41–52. <http://doi.org/10.1111/j.1365-2958.1993.tb01667.x>
- Collis, C. M., & Hall, R. M. (1992). Gene cassettes from the insert region of integrons are excised as covalently closed circles. *Molecular Microbiology*, 6(19), 2875–85. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1331702>
- Collis, C. M., & Hall, R. M. (1995). Expression of antibiotic resistance genes in the integrated cassettes of integrons. *Antimicrobial Agents and Chemotherapy*, 39(1), 155–162. <http://doi.org/10.1128/AAC.39.1.155>
- Collis, C. M., Kim, M., Partridge, S. R., Hall, R. M., & Stokes, H. W. (2002).

- Characterization of the Class 3 Integron and the Site-Specific Recombination System It Determines Characterization of the Class 3 Integron and the Site-Specific Recombination System It Determines. *Society*, 184(11), 3017–3026. <http://doi.org/10.1128/JB.184.11.3017>
- Conlan, S., Thomas, P. J., Deming, C., Park, M., Lau, A. F., Dekker, J. P., ... Segre, J. A. (2014). Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing Enterobacteriaceae. *Science Translational Medicine*, 6(254). <http://doi.org/10.1126/scitranslmed.3009845>
- Constins, O. S. X. T., Hochhut, B., Lotfi, Y., Mazel, D., Shah, M., Woodgate, R., ... Faruque, S. M. (2001). Molecular Analysis of Antibiotic Resistance Gene Clusters in *Vibrio cholerae* O139 and Molecular Analysis of Antibiotic Resistance Gene Clusters in *Vibrio cholerae* O139 and O1 SXT Constins. *Antimicrobial Agents and Chemotherapy*, 45(11), 2991–3000. <http://doi.org/10.1128/AAC.45.11.2991>
- Convey, P. (2010). Terrestrial biodiversity in Antarctica – Recent advances and future challenges. *Polar Science*, 4(2), 135–147. <http://doi.org/10.1016/J.POLAR.2010.03.003>
- Convey, P., & Stevens, M. I. (2007). Antarctic biodiversity. *Science*, 317(5846), 1877–8. <http://doi.org/10.1126/science.1147261>
- Cowan, D. A. (2014). *Antarctic Terrestrial Microbiology*. *Antarctic Terrestrial Microbiology*. <http://doi.org/10.1007/978-3-642-45213-0>
- Crits-Christoph, A., Diamond, S., Butterfield, C. N., Thomas, B. C., & Banfield, J. F. (2018). Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. *Nature*, 558(7710), 440–444. <http://doi.org/10.1038/s41586-018-0207-y>
- Crombie, A. T., Larke-Mejia, N. L., Emery, H., Dawson, R., Pratscher, J., Murphy, G. P., ... Murrell, J. C. (2018). Poplar phyllosphere harbors disparate isoprene-degrading bacteria. *Proceedings of the National Academy of Sciences of the United States of America*. <http://doi.org/10.1073/pnas.1812668115>

- Cury, J., Jové, T., Touchon, M., Néron, B., & Rocha, E. P. (2016a). Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Research*, *44*(10), 4539–4550. <http://doi.org/10.1093/nar/gkw319>
- Darling, C. A., & Siple, P. A. (1941). Bacteria of Antarctica. *Journal of Bacteriology*, *42*(1), 83–98.
- de la Cruz, F., & Davies, J. (2000). Horizontal gene transfer and the origin of species: lessons from bacteria. - PubMed - NCBI. *Trends in Microbiology*, (1985), 128–133. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/?term=10707066>
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., ... Gascuel, O. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research*, *36*(Web Server issue), 465–469. <http://doi.org/10.1093/nar/gkn180>
- Dyda, F., Klein, D. C., & Hickman, A. B. (2000). GCN5-RELATED N-ACETYLTRANSFERASES: A Structural Overview. *Annual Review of Biophysics and Biochemical Structure* Vol. 29:81-103 <https://doi/10.1146/annurev.biophys.29.1.81>
- Eikmeyer, F., Hadiati, A., Szczepanowski, R., Wibberg, D., Schneiker-bekel, S., Rogers, L. M., ... Schlüter, A. (2012). Plasmid The complete genome sequences of four new IncN plasmids from wastewater treatment plant effluent provide new insights into IncN plasmid diversity and evolution. *Plasmid*, *68*(1), 13–24. <http://doi.org/10.1016/j.plasmid.2012.01.011>
- Elsaied, H., Stokes, H. W., Kitamura, K., Kurusu, Y., Kamagata, Y., & Maruyama, A. (2011). Marine integrons containing novel integrase genes, attachment sites, attI, and associated gene cassettes in polluted sediments from Suez and Tokyo Bays. *ISME Journal*, *5*(7), 1162–1177. <http://doi.org/10.1038/ismej.2010.208>
- Elsaied, H., Stokes, H. W., Yoshioka, H., Mitani, Y., & Maruyama, A. (2014). Novel integrons and gene cassettes from a Cascadian submarine gas-hydrate-bearing core. *FEMS Microbiology Ecology*, *87*(2), 343–56. <http://doi.org/10.1111/1574-6941.12227>

- Erdmann, S., Tschitschko, B., Zhong, L., Raftery, M. J., & Cavicchioli, R. (2017). A plasmid from an Antarctic haloarchaeon uses specialized membrane vesicles to disseminate and infect plasmid-free cells. *Nature Microbiology*, *2*(10), 1446–1455. <http://doi.org/10.1038/s41564-017-0009-2>
- Erill, I., Campoy, S., & Barbé, J. (2007). Aeons of distress: An evolutionary perspective on the bacterial SOS response. *FEMS Microbiology Reviews*, *31*(6), 637–656. <http://doi.org/10.1111/j.1574-6976.2007.00082.x>
- Escudero, J. A. (2015). The Integron: Adaptation On Demand. *Microbiol Spectrum*, *3*(2). <http://doi.org/10.1128/microbiolspec>
- Escudero, J. A., Loot, C., & Mazel, D. (2018). *Integrans as Adaptive Devices*. Molecular Mechanisms of Microbial Evolution, Grand Challenges in Biology and Biotechnology, https://doi.org/10.1007/978-3-319-69078-0_9
- Escudero, J. A., Loot, C., Parissi, V., Nivina, A., Bouchier, C., & Mazel, D. (2016). Unmasking the ancestral activity of integron integrases reveals a smooth evolutionary transition during functional innovation. *Nature Communications*, *7*, 1–12. <http://doi.org/10.1038/ncomms10937>
- Fluit, A. C., & Schmitz, F. J. (2004a). Resistance integrons and super-integrans. *Clinical Microbiology and Infection*, *10*(4), 272–288. <http://doi.org/10.1111/j.1198-743X.2004.00858.x>
- Frost, L. S., Leplae, R., Summers, A. O., & Toussaint, A. (2005). Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology*, *3*(9), 722–732. <http://doi.org/10.1038/nrmicro1235>
- Galperin, M. Y., Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2015). Expanded Microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Research*, *43*(D1), D261–D269. <http://doi.org/10.1093/nar/gku1223>
- Ganzert, L., Lipski, A., Hubberten, H. W., & Wagner, D. (2011). The impact of different soil parameters on the community structure of dominant bacteria from nine different soils located on Livingston Island, South Shetland Archipelago, Antarctica. *FEMS Microbiology Ecology*, *76*(3), 476–491. <http://doi.org/10.1111/j.1574-6941.2011.01068.x>

- Garrity, G. M., Chai, B., Wang, Q., McGarrell, D. M., Cole, J. R., Farris, R. J., ... Cardenas, E. (2006). The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Research*, *35*(Database), D169–D172. <http://doi.org/10.1093/nar/gkl889>
- Gibson, M. K., Forsberg, K. J., & Dantas, G. (2015). Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *The ISME Journal*, *9*(1), 207–216. <http://doi.org/10.1038/ismej.2014.106>
- Gillings, M., Boucher, Y., Labbate, M., Holmes, A., Krishnan, S., Holley, M., & Stokes, H. W. (2008). The evolution of class 1 integrons and the rise of antibiotic resistance. *Journal of Bacteriology*, *190*(14), 5095–5100. <http://doi.org/10.1128/JB.00152-08>
- Gillings, M. R. (2014). Integrons: past, present, and future. *Microbiology and Molecular Biology Reviews : MMBR*, *78*(2), 257–77. <http://doi.org/10.1128/MMBR.00056-13>
- Gillings, M. R. (2017). Lateral gene transfer, bacterial genome evolution, and the Anthropocene. *Annals of the New York Academy of Sciences*, *1389*(1), 20–36. <http://doi.org/10.1111/nyas.13213>
- Gillings, M. R., Gaze, W. H., Pruden, A., Smalla, K., Tiedje, J. M., & Zhu, Y.-G. (2015). Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *The ISME Journal*, *9*(6), 1269–79. <http://doi.org/10.1038/ismej.2014.226>
- Gillings, M. R., Holley, M. P., & Stokes, H. W. (2009). Evidence for dynamic exchange of qac gene cassettes between class 1 integrons and other integrons in freshwater biofilms. *FEMS Microbiology Letters*, *296*(2), 282–288. <http://doi.org/10.1111/j.1574-6968.2009.01646.x>
- Gillings, M. R., Krishnan, S., Worden, P. J., & Hardwick, S. A. (2008). Recovery of diverse genes for class 1 integron-integrases from environmental DNA samples. *FEMS Microbiology Letters*, *287*(1), 56–62. <http://doi.org/10.1111/j.1574-6968.2008.01291.x>
- Gillings, M. R., Xuejun, D., Hardwick, S. A., Holley, M. P., & Stokes, H. W.

- (2009). Gene cassettes encoding resistance to quaternary ammonium compounds: a role in the origin of clinical class 1 integrons? *The ISME Journal*, 3(2), 209–215. <http://doi.org/10.1038/ismej.2008.98>
- Glass, E. M., & Meyer, F. (2011). The Metagenomics RAST Server: A Public Resource for the Automatic Phylogenetic and Functional Analysis of Metagenomes. *Handbook of Molecular Microbial Ecology I: Metagenomics and Complementary Approaches*, 8, 325–331. <http://doi.org/10.1002/9781118010518.ch37>
- Gootz, T. D., Lescoe, M. K., Dib-hajj, F., Dougherty, B. A., He, W., Della-latta, P., & Huard, R. C. (2009). Genetic Organization of Transposase Regions Surrounding bla KPC Carbapenemase Genes on Plasmids from Klebsiella Strains Isolated in a New York City Hospital , 53(5), 1998–2004. <http://doi.org/10.1128/AAC.01355-08>
- Guindon, S., Dufayard, J., & Lefort, V. New Algorithms and Methods to Estimate Maximum Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, 59(3):307-21, 2010. [http://doi.org/10.1016/0003-6870\(83\)90219-3](http://doi.org/10.1016/0003-6870(83)90219-3).
- GRINDLEY, N. D. F. (2002). The Movement of Tn 3 -Like Elements : Transposition and Cointegrate Resolution. *Mobile DNA II*, 272–302.
- Grindley, N. D. F., Whiteson, K. L., & Rice, P. A. (2006). Mechanisms of Site-Specific Recombination. *Annual Review of Biochemistry*, 75(1), 567–605. <http://doi.org/10.1146/annurev.biochem.73.011303.073908>
- Guérout, A.-M., Iqbal, N., Mine, N., Ducos-Galand, M., Van Melderen, L., & Mazel, D. (2013). Characterization of the phd-doc and ccd toxin-antitoxin cassettes from Vibrio superintegrons. *Journal of Bacteriology*, 195(10), 2270–83. <http://doi.org/10.1128/JB.01389-12>
- Hall, R. M., Brookes, D. E., & Stokes, H. W. (1991). Site-specific insertion of genes into integrons: role of the 59-base element and determination of the recombination cross-over point. *Molecular Microbiology*, 5(8), 1941–1959. <http://doi.org/10.1111/j.1365-2958.1991.tb00817.x>

- Hernández, J., Stedt, J., Bonnedahl, J., Molin, Y., Drobni, M., Calisto-Ulloa, N., ... Olsen, B. (2012). Human-associated extended-spectrum β -lactamase in the Antarctic. *Applied and Environmental Microbiology*, 78(6), 2056–2058. <http://doi.org/10.1128/AEM.07320-11>
- Holmes, A. J., Gillings, M. R., Nield, B. S., Mabbutt, B. C., Nevalainen, K. M. H., & Stokes, H. W. (2003). The gene cassette metagenome is a basic resource for bacterial genome evolution. *Environmental Microbiology*, 5(5), 383–94. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12713464>
- Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., & Bork, P. (2017). Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Molecular Biology and Evolution*, 34(8), 2115–2122. <http://doi.org/10.1093/molbev/msx148>
- Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., ... Bork, P. (2016). EGGNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Research*, 44(D1), D286–D293. <http://doi.org/10.1093/nar/gkv1248>
- Iqbal, N., Guérout, A. M., Krin, E., Le Roux, F., & Mazel, D. (2015). Comprehensive functional analysis of the 18 *Vibrio cholerae* N16961 toxin-antitoxin systems substantiates their role in stabilizing the superintegron. *Journal of Bacteriology*, 197(13), 2150–2159. <http://doi.org/10.1128/JB.00108-15>
- Jeremy E. Koenig, Christine Sharp¹, Marlena Dlutek¹, Bruce Curtis¹, Michael Joss², Yan Boucher³, W. F. D. (2009). Integron Gene Cassettes and Degradation of Compounds Associated with Industrial Waste: The Case of the Sydney Tar Ponds. *PLoS ONE*, 4(4), e5276. <http://doi.org/10.1371/journal.pone.0005276>
- Jové, T., Da Re, S., Denis, F., Mazel, D., & Ploy, M. C. (2010). Inverse correlation between promoter strength and excision activity in class 1 integrons. *PLoS Genetics*, 6(1). <http://doi.org/10.1371/journal.pgen.1000793>

- Jové, T., Da Re, S., Tabesse, A., Gassama-Sow, A., & Ploy, M. C. (2017). Gene expression in class 2 integrons is sos-independent and involves two Pc promoters. *Frontiers in Microbiology*, 8(AUG), 1–11.
<http://doi.org/10.3389/fmicb.2017.01499>
- Jungblut, A. D., & Neilan, B. A. (2010). NifH gene diversity and expression in a microbial mat community on the McMurdo Ice Shelf, Antarctica. *Antarctic Science*, 22(02), 117. <http://doi.org/10.1017/S0954102009990514>
- Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2014). Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Research*, 42(D1), 199–205.
<http://doi.org/10.1093/nar/gkt1076>
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780.
<http://doi.org/10.1093/molbev/mst010>
- Kent, W. J. (2002). BLAT - The BLAST-like alignment tool. *Genome Research*, 12(4), 656–664. <http://doi.org/10.1101/gr.229202>. Article published online before March 2002
- Kim, S. R., & Komano, T. (1992). Nucleotide sequence of the R721 shufflon. *Journal of Bacteriology*, 174(21), 7053–7058.
<http://doi.org/10.1128/jb.174.21.7053-7058.1992>
- Kloesges, T., Popa, O., Martin, W., & Dagan, T. (2011). Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths. *Molecular Biology and Evolution*, 28(2), 1057–1074. <http://doi.org/10.1093/molbev/msq297>
- Komárek, J., & Komárek, J. (2013). Phenotypic and ecological diversity of freshwater coccoid cyanobacteria from maritime Antarctica and islands of NW Weddell Sea. I. Synechococcales, 3(2), 130–143.
<http://doi.org/10.5817/CPR2013-2-14>
- Komarek, O., & Komarek, J. (1999). Diversity of freshwater and terrestrial habitats and their oxyphototroph microflora in the Arctowski Station region,

- South Shetland Islands. *Polish Polar Research*, 20(3), 259–282.
- Koonin, E. V., Makarova, K. S., & Aravind, L. (2002). Horizontal Gene Transfer in Prokaryotes: Quantification and Classification. *Annual Review of Microbiology*, 55(1), 709–742.
<http://doi.org/10.1146/annurev.micro.55.1.709>
- Lane, D.J. 1991. 16S/23S rRNA sequencing. In: Nucleic acid techniques in bacterial systematics. Stackebrandt, E., and Goodfellow, M., eds., John Wiley and Sons, New York, NY, pp. 115-175
- Larsen, N., Brodie, E. L., Andersen, G. L., Huber, T., DeSantis, T. Z., Keller, K., ... Hu, P. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology*, 72(7), 5069–5072. <http://doi.org/10.1128/aem.03006-05>
- Larsson, A. (2014). AliView : a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, 30(22), 3276–3278.
<http://doi.org/10.1093/bioinformatics/btu531>
- Lederberg 1946 Conjugation.pdf. (n.d.). *Nature* Vol. 158.
- Léon, G., & Roy, P. H. (2009). Potential role of group IIC-attC introns in integron cassette formation. *Journal of Bacteriology*, 191(19), 6040–6051.
<http://doi.org/10.1128/JB.00674-09>
- Liebert, C. A., Hall, R. M., & Summers, A. O. (1999). Transposon Tn21, flagship of the floating genome. *Microbiology and Molecular Biology Reviews : MMBR*, 63(3), 507–22. Retrieved from
<http://www.ncbi.nlm.nih.gov/pubmed/10477306>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC103744>
- Loot, C., Nivina, A., Cury, J., Escudero, J. A., Ducos-Galand, M., Bikard, D., ... Mazela, D. (2017). Differences in integron cassette excision dynamics shape a trade-off between evolvability and genetic capacitance. *MBio*, 8sRo(2).
<http://doi.org/10.1128/mBio.02296-16>
- Lopatina Anna, Krylenkov, V., Medvedeva, S., Severinov, K., Lopatina, A., Shmakov, S., & Logacheva, M. D. (2016). Metagenomic Analysis of

- Bacterial Communities of Antarctic Surface Snow. *Frontiers in Microbiology*, 7(March), 1–13. <http://doi.org/10.3389/fmicb.2016.00398>
- Macdonald, D., Demarre, G., Bouvier, M., & Mazel, D. (2006). Structural basis for broad DNA specificity in integron recombination, *440*(7088).
- Marquez, C., Labbate, M., Ingold, A. J., Chowdhury, P. R., Ramírez, M. S., Centrón, D., ... Stokes, H. W. (2008). Recovery of a functional class 2 integron from an Escherichia coli strain mediating a urinary tract infection. *Antimicrobial Agents and Chemotherapy*, 52(11), 4153–4154. <http://doi.org/10.1128/AAC.00710-08>
- Maruyama, A., Kitamura, K., Nakamura, T., Elsaied, H., Stokes, H. W., & Fuse, H. (2007). Novel and diverse integron integrase genes and integron-like gene cassettes are prevalent in deep-sea hydrothermal vents. *Environmental Microbiology*, 9(9), 2298–2312. <http://doi.org/10.1111/j.1462-2920.2007.01344.x>
- Math, R. K., Jin, H. M., Kim, J. M., Hahn, Y., Park, W., Madsen, E. L., & Jeon, C. O. (2012). Comparative genomics reveals adaptation by Alteromonas sp. SN2 to marine tidal-flat conditions: Cold tolerance and aromatic hydrocarbon metabolism. *PLoS ONE*, 7(4). <http://doi.org/10.1371/journal.pone.0035784>
- Mazel, D. (2006). Integrons: agents of bacterial evolution. *Nature Reviews Microbiology*, 4(8), 608–620. <http://doi.org/10.1038/nrmicro1462>
- Mazel, D., Dychinco, B., Webb, V. ., & Davies, J. (2000). Antibiotic resistance in the ECOR collection:integrons and identification of a novel aad gene. *Antimicrobial Agents and Chemotherapy*, 44(6), 1568–1574. <http://doi.org/10.1128/AAC.44.6.1568-1574.2000>.
- Mazel, D., Dychinco, B., Webb, V. a, & Davies, J. (1998). A distinctive class of integron in the Vibrio cholerae genome. *Science (New York, N.Y.)*, 280(5363), 605–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9554855>
- Messier, N., & Roy, P. H. (2001). Integron integrases possess a unique additional domain necessary for activity. *Journal of Bacteriology*, 183(22),

6699–706. <http://doi.org/10.1128/JB.183.22.6699-6706.2001>

- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E., Kubal, M., ... Edwards, R. (2008). The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, 9(1), 386. <http://doi.org/10.1186/1471-2105-9-386>
- Miroshnichenko, M. L., Kostrikina, N. A., Chernyh, N. A., Pimenov, N. V., Tourova, T. P., Antipov, A. N., ... Bonch-Osmolovskaya, E. A. (2003). *Caldithrix abyssi* gen. nov., sp. nov., a nitrate-reducing, thermophilic, anaerobic bacterium isolated from a Mid-Atlantic ridge hydrothermal vent, represents a novel bacterial lineage. *International Journal of Systematic and Evolutionary Microbiology*, 53(1), 323–329. <http://doi.org/10.1099/ijs.0.02390-0>
- Moura, A., Soares, M., Pereira, C., Leitao, N., Henriques, I., & Correia, A. (2009). INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. *Bioinformatics*, 25(8), 1096–1098. <http://doi.org/10.1093/bioinformatics/btp105>
- Nawrocki, E. P., & Eddy, S. R. (2013). Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*, 29(22), 2933–2935. <http://doi.org/10.1093/bioinformatics/btt509>
- Nemergut, D. R., Martin, a P., & Schmidt, S. K. (2004). Integron diversity in heavy-metal-contaminated mine tailings and inferences about integron evolution. *Applied and Environmental Microbiology*, 70(2), 1160–8. <http://doi.org/10.1128/AEM.70.2.1160>
- Nield, B. S., Holmes, A. J., Gillings, M. R., Recchia, G. D., Mabbutt, B. C., Nevalainen, K. M. H., & Stokes, H. W. (2001). Recovery of new integron classes from environmental DNA. *FEMS Microbiology Letters*, 195(1), 59–65. [http://doi.org/10.1016/S0378-1097\(00\)00546-2](http://doi.org/10.1016/S0378-1097(00)00546-2)
- Nield, B. S., Willows, R. D., Torda, A. E., Gillings, M. R., Holmes, A. J., Nevalainen, K. M., ... Mabbutt, B. C. (2004). New enzymes from environmental cassette arrays: functional attributes of a phosphotransferase and an RNA-methyltransferase. *Protein Sci*, 13(6),

- 1651–1659. <http://doi.org/10.1110/ps.04638704\13/6/1651> [pii]
- Nivina, A., Escudero, J. A., Vit, C., Mazel, D., & Loot, C. (2016). Efficiency of integron cassette insertion in correct orientation is ensured by the interplay of the three unpaired features of attC recombination sites. *Nucleic Acids Research*, *44*(16), 7792–7803. <http://doi.org/10.1093/nar/gkw646>
- Nunes-Düby, S. E., Kwon, H. J., Tirumalai, R. S., Ellenberger, T., & Landy, A. (1998). Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucleic Acids Research*, *26*(2), 391–406. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9421491>
- Ochman, H., Lawrence, J. G., & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature*, *405*(6784), 299–304. <http://doi.org/10.1038/35012500>
- Partridge, S. R. (2011). Analysis of antibiotic resistance regions in Gram-negative bacteria. *FEMS Microbiology Reviews*, *35*(5), 820–855. <http://doi.org/10.1111/j.1574-6976.2011.00277.x>
- Partridge, S. R., Tsafnat, G., Coiera, E., & Iredell, J. R. (2009). Gene cassettes and cassette arrays in mobile resistance integrons: Review article. *FEMS Microbiology Reviews*, *33*(4), 757–784. <http://doi.org/10.1111/j.1574-6976.2009.00175.x>
- Peeters, K., Ertz, D., & Willems, A. (2011). Culturable bacterial diversity at the Princess Elisabeth Station (Utsteinen, Sør Rondane Mountains, East Antarctica) harbours many new taxa. *Systematic and Applied Microbiology*, *34*(5), 360–7. <http://doi.org/10.1016/j.syapm.2011.02.002>
- Peng, Y., Leung, H. C. M., Yiu, S. M., & Chin, F. Y. L. (2012). IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, *28*(11), 1420–1428. <http://doi.org/10.1093/bioinformatics/bts174>
- Phan, M. D., Forde, B. M., Peters, K. M., Sarkar, S., Hancock, S., Stanton-Cook, M., ... Schembri, M. A. (2015). Molecular characterization of a multidrug resistance IncF plasmid from the globally disseminated escherichia coli ST131 clone. *PLoS ONE*, *10*(4), 1–17.

<http://doi.org/10.1371/journal.pone.0122369>

- Power, M. L., Samuel, A., Smith, J. J., Stark, J. S., Gillings, M. R., & Gordon, D. M. (2016). *Escherichia coli* out in the cold: Dissemination of human-derived bacteria into the Antarctic microbiome. *Environmental Pollution*, *215*, 58–65. <http://doi.org/10.1016/j.envpol.2016.04.013>
- Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., & Glöckner, F. O. (2007). SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, *35*(21), 7188–7196. <http://doi.org/10.1093/nar/gkm864>
- Ramírez-Fernández, L., Trefault, N., Carú, M., & Orlando, J. (2019). Seabird and pinniped shape soil bacterial communities of their settlements in Cape Shirreff, Antarctica. *PLoS ONE*, *14*(1), 1–22. <http://doi.org/10.1371/journal.pone.0209887>
- Ramírez, M. S., Piñeiro, S., Centrón, D., Kaufman, S., Kovensky, J., Vay, C., ... Predari, S. (2010). Novel insights about class 2 integrons from experimental and genomic epidemiology. *Antimicrobial Agents and Chemotherapy*, *54*(2), 699–706. <http://doi.org/10.1128/AAC.01392-08>
- Ramírez, M. S., Quiroga, C., & Centrón, D. (2005). Novel rearrangement of a class 2 integron in two non-epidemiologically related isolates of *Acinetobacter baumannii*. *Antimicrobial Agents and Chemotherapy*, *49*(12), 5179–5181. <http://doi.org/10.1128/AAC.49.12.5179-5181.2005>
- Rapa, R. A., & Labbate, M. (2013). The function of integron-associated gene cassettes in *Vibrio* species: The tip of the iceberg. *Frontiers in Microbiology*, *4*(DEC), 1–7. <http://doi.org/10.3389/fmicb.2013.00385>
- Reasoner, D. J., & Geldreich, E. E. (1985). A new medium for the enumeration and subculture of bacteria from potable water. *Applied and Environmental Microbiology*, *49*(1), 1–7.
- Recchia, G. D., & Hall, R. M. (1995). Gene cassettes: A new class of mobile element. *Microbiology*, *141*(12), 3015–3027. <http://doi.org/10.1099/13500872-141-12-3015>

- Rodríguez-Minguela, C. M., Apajalahti, J. H. A., Chai, B., Cole, J. R., & Tiedje, J. M. (2009). Worldwide prevalence of class 2 integrases outside the clinical setting is associated with human impact. *Applied and Environmental Microbiology*, *75*(15), 5100–5110. <http://doi.org/10.1128/AEM.00133-09>
- Rossolini, G. M., D'Andrea, M. M., & Mugnaioli, C. (2008). The spread of CTX-M-type extended-spectrum β -lactamases. *Clinical Microbiology and Infection*, *14*(SUPPL. 1), 33–41. <http://doi.org/10.1111/j.1469-0691.2007.01867.x>
- Rowe-Magnus, D. a., Guerout, A.-M., & Mazel, D. (2002). Bacterial resistance evolution by recruitment of super-integron gene cassettes. *Molecular Microbiology*, *43*(6), 1657–1669. <http://doi.org/10.1046/j.1365-2958.2002.02861.x>
- Rowe-Magnus, D. A., Guerout, A.-M., Ploncard, P., Dychinco, B., Davies, J., & Mazel, D. (2001). The evolutionary history of chromosomal super-integrations provides an ancestry for multiresistant integrons. *Proceedings of the National Academy of Sciences*, *98*(2), 652–657. <http://doi.org/10.1073/pnas.98.2.652>
- Rowe-magnus, D. a, Guerout, A., Biskri, L., Bouige, P., & Mazel, D. (2003). Comparative Analysis of Superintegrons : Engineering Extensive Genetic Diversity in the Vibrionaceae Comparative Analysis of Superintegrons : Engineering Extensive Genetic Diversity in the Vibrionaceae. *Genome Research*, 428–442. <http://doi.org/10.1101/gr.617103>
- Rowe-Magnus, D. a, & Mazel, D. (2001). Integrons: natural tools for bacterial genome evolution. *Current Opinion in Microbiology*, *4*(5), 565–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11587934>
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning : a laboratory manual*. Cold Spring Harbor Laboratory. Retrieved from <https://trove.nla.gov.au/work/13615226?q&versionId=45221481>
- Sharma, D., & Bisht, D. (2017). Role of Bacterioferritin & Ferritin in M. tuberculosis Pathogenesis and Drug Resistance: A Future Perspective by Interactomic Approach. *Frontiers in Cellular and Infection Microbiology*,

- 7(June), 1–5. <http://doi.org/10.3389/fcimb.2017.00240>
- Sheperd, A., Ivins, E., Rignot, E., Smith, B., van den Broeke, M., & Velicogna, I. (2018). Mass balance of the Antarctic Ice Sheet. *Nature*, *558*, 219–234.
- Siegele, D. A. (2005). GUEST COMMENTARIES Universal Stress Proteins in *Escherichia coli*. *Journal of Bacteriology*, *187*(18), 6253–6254. <http://doi.org/10.1128/JB.187.18.6253>
- Siguier, P., Gourbeyre, E., & Chandler, M. (2014). Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiology Reviews*, *38*(5), 865–891. <http://doi.org/10.1111/1574-6976.12067>
- Sitkiewicz, I., Stockbauer, K. E., & Musser, J. M. (2007). Secreted bacterial phospholipase A2 enzymes: better living through phospholipolysis. *Trends in Microbiology*, *15*(2), 63–69. <http://doi.org/10.1016/j.tim.2006.12.003>
- Smillie, C., Garcillan-Barcia, M. P., Francia, M. V., Rocha, E. P. C., & de la Cruz, F. (2010). Mobility of Plasmids. *Microbiology and Molecular Biology Reviews*, *74*(3), 434–452. <http://doi.org/10.1128/MMBR.00020-10>
- Snitkin, E. S., Palmore, T. N., Zelazny, A. M., Murray, P. R., Segre, J. A., & Gupta, J. (2013). Genomic insights into the fate of colistin resistance and *Acinetobacter baumannii* during patient treatment. *Genome Research*, *23*(7), 1155–1162. <http://doi.org/10.1101/gr.154328.112>
- Sorum, H., Roberts, M. C., & Crosa, J. H. (1992). Identification and cloning of a tetracycline resistance gene from the fish pathogen *Vibrio salmonicida*. *Antimicrobial Agents and Chemotherapy*, *36*(3), 611–615. <http://doi.org/10.1128/AAC.36.3.611>
- Stokes, H. W., & Hall, R. M. (1989). A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. *Molecular Microbiology*, *3*(12), 1669–1683. <http://doi.org/10.1111/j.1365-2958.1989.tb00153.x>
- Stokes, H. W., Holmes, A. J., Nield, B. S., Holley, M. P., Nevalainen, K. M. H., Mabbutt, B. C., & Gillings, M. R. (2001). Gene Cassette PCR: Sequence-Independent Recovery of Entire Genes from Environmental DNA. *Applied and Environmental Microbiology*, *67*(3–12), 5240–5246.

<http://doi.org/10.1128/AEM.67.11.5240-5246.2001>

Stokes, H. W., Nesbø, C. L., Holley, M., Bahl, M. I., Gillings, M. R., & Boucher, Y. (2006). Class 1 integrons potentially predating the association with *tn402*-like transposition genes are present in a sediment microbial community. *Journal of Bacteriology*, *188*(16), 5722–30.

<http://doi.org/10.1128/JB.01950-05>

Stokes, H. W., O’Gorman, D. B., Recchia, G. D., Parsekhian, M., & Hall, R. M. (1997). Structure and function of 59-base element recombination sites associated with mobile gene cassettes. *Molecular Microbiology*, *26*(4), 731–745. <http://doi.org/10.1046/j.1365-2958.1997.6091980.x>

Straka, R. P., & Stokes, J. L. (1960). Psychrophilic Bacteria from Antarctica. *Journal of Applied Microbiology*, *80*(5), 622–625.

Tatusov, R. L., Koonin, E. V., & Lipman, D. J. (1997). A Genomic Perspective on Protein Families. *Science (New York, N.Y.)*, *631*(1997), 631–637.

<http://doi.org/10.1126/science.278.5338.631>

Terauds, A., & Lee, J. R. (2016). Antarctic biogeography revisited: updating the Antarctic Conservation Biogeographic Regions. *Diversity and Distributions*, *22*(8), 836–840. <http://doi.org/10.1111/ddi.12453>

Turner, J., Colwell, S. R., Marshall, G. J., Lachlan-Cope, T. A., Carleton, A. M., Jones, P. D., ... Iagovkina, S. (2005). Antarctic climate change during the last 50 years. *International Journal of Climatology*, *25*(3), 279–294.

<http://doi.org/10.1002/joc.1130>

Turner, J., Overland, J. E., & Walsh, J. E. (2007). An Arctic and Antarctic perspective on recent climate change. *INTERNATIONAL JOURNAL OF CLIMATOLOGY Int. J. Climatol*, *27*, 277–293.

<http://doi.org/10.1002/joc.1406>

Tytgat, B., Verleyen, E., Obbels, D., Peeters, K., De Wever, A., D’hondt, S., ... Willems, A. (2014). Bacterial diversity assessment in antarctic terrestrial and aquatic microbial mats: A comparison between bidirectional pyrosequencing and cultivation. *PLoS ONE*, *9*(6).

<http://doi.org/10.1371/journal.pone.0097564>

- Valdespino-Castillo, P. M., Cerqueda-García, D., Espinosa, A. C., Batista, S., Merino-Ibarra, M., Taş, N., ... Falcón, L. I. (2018). Microbial distribution and turnover in Antarctic microbial mats highlight the relevance of heterotrophic bacteria in low-nutrient environments. *FEMS Microbiology Ecology*, 94(9). <http://doi.org/10.1093/femsec/fiy129>
- Van Melderen, L., & De Bast, M. S. (2009). Bacterial toxin-Antitoxin systems: More than selfish entities? *PLoS Genetics*, 5(3). <http://doi.org/10.1371/journal.pgen.1000437>
- Weiner, R. M., Taylor, L. E., Henrissat, B., Hauser, L., Land, M., Coutinho, P. M., ... Hutcheson, S. (2008). Complete genome sequence of the complex carbohydrate-degrading marine bacterium, *Saccharophagus degradans* strain 2-40T. *PLoS Genetics*, 4(5). <http://doi.org/10.1371/journal.pgen.1000087>
- Wilke, A., Harrison, T., Wilkening, J., Field, D., Glass, E. M., Kyrpides, N., ... Meyer, F. (2012). The M5nr: A novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. *BMC Bioinformatics*, 13(1). <http://doi.org/10.1186/1471-2105-13-141>
- Wong, J. J. W., Lu, J., & Glover, J. N. M. (2012). Relaxosome function and conjugation regulation in F-like plasmids - a structural biology perspective. *Molecular Microbiology*, 85(4), 602–617. <http://doi.org/10.1111/j.1365-2958.2012.08131.x>
- Xu, H., Davies, J., & Miao, V. (2007). Molecular characterization of class 3 integrons from *Delftia* spp. *Journal of Bacteriology*, 189(17), 6276–6283. <http://doi.org/10.1128/JB.00348-07>
- Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A Greedy Algorithm for Aligning DNA Sequences. *Journal of Computational Biology*, 7(1–2), 203–214. <http://doi.org/10.1089/10665270050081478>

