



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS**

INSTITUTO DE ECOLOGÍA
ECOLOGÍA

**ANÁLISIS METAGENÓMICO DE TAPICES MICROBIANOS DE LA REGIÓN
ANTÁRTICA 'MCMURDO DRY VALLEYS'**

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS BIOLÓGICAS

PRESENTA:

RICARDO AUGUSTO MERCADO JUÁREZ

TUTORA PRINCIPAL DE TESIS: **DRA. LUISA ISaura FALCÓN ÁLVAREZ**

INSTITUTO DE ECOLOGÍA, UNAM

COMITÉ TUTOR: **DRA. ROCÍO JETZABEL ALCÁNTARA HERNÁNDEZ**

INSTITUTO DE GEOLOGÍA, UNAM

DR. LUIS EDUARDO SERVÍN GARCIDUEÑAS

ENES MORELIA, UNAM

MÉXICO, CD. MX. ENERO, 2020



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE ECOLOGÍA
ECOLOGÍA

**ANÁLISIS METAGENÓMICO DE TAPICES MICROBIANOS DE LA REGIÓN
ANTÁRTICA 'MCMURDO DRY VALLEYS'**

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS BIOLÓGICAS

PRESENTA:
RICARDO AUGUSTO MERCADO JUÁREZ

TUTORA PRINCIPAL DE TESIS: DRA. LUISA ISAURA FALCÓN ÁLVAREZ

INSTITUTO DE ECOLOGÍA, UNAM

COMITÉ TUTOR: DRA. ROCÍO JETZABEL ALCÁNTARA HERNÁNDEZ

INSTITUTO DE GEOLOGÍA, UNAM

DR. LUIS EDUARDO SERVÍN GARCIDUEÑAS

ENES MORELIA, UNAM

MÉXICO, CD. MX. ENERO, 2020

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS

ENTIDAD INSTITUTO DE ECOLOGÍA

OFICIO CPCB/028/2020

ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
Presente

Me permito informar a usted, que el Subcomité de Ecología y Biología Evolutiva, del Posgrado en Ciencias Biológicas en su sesión ordinaria del día 30 de septiembre 2019, aprobó el siguiente jurado para el examen de grado de **MAESTRO EN CIENCIAS BIOLÓGICAS** en el campo de conocimiento de **Ecología**, del estudiante **MERCADO JUÁREZ RICARDO AUGUSTO** con número de cuenta: **300296569** con la tesis titulada "**ANÁLISIS METAGENÓMICO DE TAPICES MICROBIANOS DE LA REGIÓN ANTÁRTICA MCMURDO DRY VALLEYS**", bajo la dirección de la **DRA. LUISA ISAURA FALCÓN ÁLVAREZ**, quedando integrado de la siguiente manera:

Presidente:	DR. ARTURO CARLOS II BECERRA BRACHO
Vocal:	DRA. ANA LAURA WEGIER BRIUOLO
Secretario:	DRA. ROCÍO JETZABEL ALCANTARA HERNÁNDEZ
Suplente:	DR. ISAAC GONZÁLEZ SANTOYO
Suplente:	DR. RENÉ CERRITOS FLORES

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPÍRITU"
Ciudad Universitaria, Cd. Mx., a 8 de enero de 2020

COORDINADOR DEL PROGRAMA


DR. ADOLFO GERARDO NAVARRO SIGÜENZA



c. c. p. Expediente del alumno

Agradecimientos Institucionales

Al Posgrado en Ciencias Biológicas perteneciente a la Universidad Nacional Autónoma de México

Al Consejo Nacional de Ciencia y Tecnología por el apoyo otorgado con una beca de posgrado para la realización de este trabajo de Maestría en Ciencias en el Posgrado en Ciencias Biológicas de la UNAM

Este estudio ha sido financiado por el Fondo Conjunto México-Chile de Cooperación Internacional "Impactos en la diversidad microbiana de ecosistemas polares frente al cambio ambiental", con apoyo de la AMEXCID y la AGCI.

A mi comité tutor: Dra. Luisa Isaura Falcón Álvarez, Dra. Rocío Jetzabel Alcántara Hernández, Dr. Luis Eduardo Servín Garcidueñas

Agradecimientos a título personal

A mi mamá Guadalupe Juárez, prima Rocío Cortez, tía Rosa María Juárez, abuelo Agustín Juárez y primo Alejandro Rodea, amigos y familiares

A la Universidad Nacional Autónoma de México

Al Instituto de Ecología, UNAM

A la Dra. Luisa Isaura Falcón Álvarez, Dra. Rocío Jetzabel Alcántara Hernández, Dr. Luis Eduardo Servín Garcidueñas

A los miembros del jurado: Dra. Rocío Jetzabel Alcántara Hernández, Dra Ana Laura Wegier Briuolo, Dr. Arturo Carlos Il Becerra Bracho, Dr. René Cerritos Flores y Dr. Isaac González Santoyo

Se agradece el apoyo técnico y logístico de M. en C. Osiris Gaona Pineda

A los colegas, Dr. Edward J. Carpenter, Dr. Stephen C. Cary, Jill A. Sohm, Troy E. Gunderson, Douglas G. Capone por la colecta de los tapetes en Dry Valley's

Al Instituto Antártico Chileno, Instituto Antártico Uruguayo, Dirección Nacional del Antártico de Argentina, y al programa Antártico de la National Science Foundation, EUA por el apoyo en la colecta de muestras.

A los miembros del Laboratorio de Ecología Bacteriana

Índice de figuras y tablas

Figura 1. Posición geográfica de McMurdo Dry Valley, imagen satelital en mosaico LIMA (Landsat Image Mosaic of Antarctica) de la zona de McMurdo Dry Valley (Antarctic Digital Database Map Viewer SCAR) y paisaje de las corrientes efímeras, formadas en el verano austral por el deshielo de los glaciares, donde están asociados tapetes microbianos en Dry Valley (Don A. Cowan Creative Commons Attribution License (CC BY)).....	8
Figura 2. Ubicación de "MDV" en el continente Antártico (derecha, círculo azul) e imagen satelital Landsat/Copernicus Google Earth ® de los lugares donde fueron colectados los tapetes microbianos.	15
Figura 3. Procedimiento general para el análisis de los metagenomas.	22
Figura 4. Gráfica de composición taxonómica entre los tapetes microbianos a nivel phylum.....	24
Figura 5. Filogenia basada en secuencias predichas 18S ARNr en los metagenomas MDV (S010,S011,S012,S013,S014).	26
Figura 6. Diagrama de Venn comparativo de los KO no repetidos anotados entre los metagenomas.....	28
Figura 7. Gráfico de barras de la asignación funcional de los KO compartidos entre los metagenomas.....	29
Figura 8. Gráfico de barras comparativo de la asignación funcional de los KO únicos de cada metagenoma.	30
Figura 9. Mapa de calor de la relación entre la taxonomía (ARQ=arqueas, BAC=bacterias, EUC=eucariontes) a nivel phylum y la asignación de funciones en los metabolismos generales indicados, entre los metagenomas. La intensidad de color indica un mayor número de KO asignados al metabolismo indicado.....	33
Tabla 1. Información de los metagenomas recuperados después de la secuenciación.....	14
Tabla 2. Valores de diversidad de los tapetes microbianos a partir del nivel género.	25
Tabla 3. Matriz de disimilitud de abundancia relativa de géneros potenciales entre los tapetes microbianos.....	25
Tabla 4. Matriz de disimilitud de presencia/ausencia de géneros potenciales entre los tapetes microbianos.....	25
Tabla 5. Resumen de la anotación contra los ortólogos KEGG (KO) de secuencias codificantes predichas entre los metagenomas.	27
Tabla 6. Presencia/Ausencia de las capacidades fotosintéticas y de quimiosíntesis entre los metagenomas a partir del mapeo de las secuencias codificantes predichas en los módulos KEGG. verde = presencia.....	31
Tabla 7. Presencia/Ausencia de la combinación de asignación de módulos KEGG entre los metagenomas. verde = presencia.....	32

Tabla 8. 'Bins' recuperados y asignados taxonómicamente, en verde los potenciales genomas a recuperar.....	34
--	----

Índice	
Índice de figuras y tablas	I
Índice	III
Resumen	1
Abstract	2
Introducción	3
Generalidades del sitio de estudio	6
Generalidades de los tapetes microbianos	9
Justificación.	10
Hipótesis	11
Objetivo	11
Objetivos particulares	11
Antecedentes	12
Métodos	13
Obtención de secuencias metagenómicas	13
Análisis de diversidad basado en lecturas	15
Análisis basado en anotación de secuencias (ensamble)	18
Recuperación de 'MAG's' ('Binning')	21
Resultados	22
Análisis de diversidad	22
Potencial funcional, capacidades metabólicas energéticas	27
Genomas potenciales recuperados	34
Discusión	35
Análisis de diversidad	35
Potencial funcional y capacidades de metabolismo energético	37
Binning	40
Conclusiones	42
Literatura citada	43
Anexo 1	49
Anexo 2	50
Anexo 3	52

Resumen

'McMurdo Dry Valleys' (MDV) es la región libre de hielo más grande en el continente antártico, es un área especial de importancia ambiental debido a sus características únicas derivadas del clima extremo, topografía y posición geográfica, es considerado un desierto polar extremo. Debido al cambio climático de las últimas décadas, la temperatura ha aumentado en la región, este aumento tiene efectos significativos en el paisaje, principalmente en la desglaciación, modificando las condiciones naturales y alterando el ecosistema. El panorama general es el aumento de zonas libres de hielo que serán colonizadas por comunidades microbianas, consideradas componente fundamental en el ecosistema antártico, sin embargo, las comunidades contemporáneas están amenazadas debido a que los cambios ocurren en un lapso corto de tiempo y no se ajustan al tiempo de adaptación de las comunidades. El objetivo de este trabajo fue hacer un análisis taxonómico y funcional de tapetes microbianos en diferentes valles de MDV desde una perspectiva metagenómica por secuenciación *shotgun* (diversidad genética) para realizar una descripción del potencial sobre diversidad y el funcionamiento. Los resultados señalan que los tapetes microbianos de MDV son diversos, con un potencial taxonómico por arriba de 700 géneros y un índice de Shannon H' mayor a 3, donde dominan los procariontes, principalmente Bacteroidetes, Proteobacteria y Cyanobacteria. El potencial funcional está dirigido a las funciones principales sobre metabolismo en general, donde destaca la fotosíntesis oxigénica y anoxigénica, fijación de carbono y nitrógeno. Se recuperaron 5 genomas ensamblados a partir de metagenomas principalmente del grupo Bacteroidetes. A pesar de que no se tienen más de 5 muestras, se pudo realizar una descripción sobre la composición, diversidad y el funcionamiento de estas comunidades fundamentales en los ecosistemas antárticos.

Abstract

"McMurdo Dry Valleys" (MDV) is the largest ice-free region in the Antarctic continent, it is a special area of environmental importance due to its unique characteristics derived from extreme weather, topography and geographical position, it is an extreme polar desert. Due to the climatic change of the last decades, the temperature has increased in the region, this increase has specific effects on the landscape, mainly in deglaciation, modifying the natural conditions and altering the ecosystem. The general panorama is the increase of ice-free zones that will be colonized by microbial communities, fundamental components considered in the Antarctic ecosystem, however, contemporary communities are affected due to changes occurring in a short period of time and do not conform to time of adaptation of the communities. The objective of this work was to make a taxonomic and functional analysis of microbial mats in different MDV valleys from a metagenomic perspective by shotgun sequencing (genetic diversity) to describe the potential for diversity and functioning. The results indicate that MDV microbial mats are diverse, with a taxonomic potential above 700 genera, where prokaryotes, mainly Bacteroidetes, Proteobacteria and Cyanobacteria, dominate. The functional potential is aimed at the main functions on the metabolism in general, which highlights oxygenic and anoxygenic photosynthesis, carbon and nitrogen fixation. Five assembled genomes were recovered from metagenomes mainly from the Bacteroidetes group. Although there are no more than 5 samples, a description could be made about the composition, diversity and functioning of these fundamental communities in the Antarctic ecosystems.

Introducción

La comunidad es un ensamble de poblaciones de distintas especies que comparten las condiciones bióticas y abióticas que permiten interactúen en el espacio-tiempo (Begon, Townsend y Harper, 2005). El resultado de la interacción entre los organismos que componen a la comunidad, mismos que están limitados por los filtros ambientales, son las propiedades emergentes de las comunidades, definidas por su diversidad, redundancia funcional y la estabilidad de la comunidad en el hábitat, por mencionar algunas. Algunas preguntas fundamentales que surgen al analizar una comunidad son ¿cómo se explica su diversidad? y ¿qué condiciona esa diversidad?

Los microorganismos son parte intrínseca de las comunidades siendo los componentes que albergan la mayor diversidad genética en las mismas. Igualmente, la mayor biomasa de los ecosistemas se concentra en los microorganismos (Wilkins *et al.*, 2013; Yau *et al.*, 2013; Cavicchioli, 2015). Uno de los objetivos del estudio de comunidades biológicas es conocer la interacción de la comunidad con su entorno, conocer las vías funcionales para la obtención de biomasa, conocer los flujos de energía, y comprender la interacción entre las poblaciones (tanto microbianas como macro) con su entorno (Konopka, 2009).

Un tapete microbiano es un sistema de modelo de estudio en donde la mayor parte de los componentes son microorganismos. Son considerados ecosistemas bentónicos, autosustentables, capaces de reciclar nutrientes y formar sus propias redes de interacción y niveles tróficos (Bolhuis, Cretoiu y Stal, 2014).

La ecología microbiana busca entender los procesos que estructuran a la diversidad microbiana y parte de dos preguntas ¿quiénes están? y ¿qué es lo que pueden hacer? La primera pregunta se ha respondido empleando técnicas dependientes e independientes de cultivo. Cultivar a los microorganismos es una técnica y un arte que nos ha llevado a comprender de forma íntima la evolución de historias de vida de algunos microorganismos; pero que, desafortunadamente no ha sido exitosa para recuperar la gran diversidad de microorganismos que se desarrollan en el planeta. Las técnicas independientes de cultivo se han basado en un acercamiento molecular, a partir del aporte de Carl Richard Woese sobre la taxonomía molecular basada en 16S ARNr (Woese y Fox, 1977). El conocimiento profundo de los microorganismos a partir de sus secuencias, no se ha basado solamente en los análisis filogenéticos de la subunidad pequeña del ADN ribosomal, sino que se han incorporado una gran diversidad de marcadores moleculares de regiones genéticas asociadas a diferentes enzimas. Estos estudios dirigidos de amplificación de ciertas regiones genéticas nos han permitido acercarnos a los estudios funcionales del papel de los microorganismos en el ambiente. Sin embargo, los métodos basados en amplificación de una región genómica (PCR, por las siglas en inglés de reacción en cadena de la polimerasa), se basan en el diseño de "oligos" "cebadores" o "*primers*" a partir del conocimiento de la señal genética proveniente de algunos organismos; y por ende, representan un sesgo a nuestro conocimiento de la diversidad. De estos conceptos y conjuntos de métodos se desprenden los primeros análisis filogenéticos de organismos no cultivables, como por ejemplo los realizados en *Yellowstone* en comunidades de organismos extremófilos formadores de tapetes

microbianos (Lane *et al.*, 1985). Igualmente, el desarrollo integral de la ecología microbiana incorpora la segunda pregunta ¿qué es lo que pueden hacer? y tiene su inicio en 2005 al ser acuñado el término *Metagenómica* por Jo Handelsman: "*La metagenómica (también conocida como genómica ambiental y comunitaria) es el análisis genómico de microorganismos mediante extracción directa y secuenciación del ADN*" (Handelsman, 2004). En este caso, el objetivo es acercarse al 99% de los microorganismos que aún no son cultivables (Rinke *et al.*, 2013), además de saltarse la limitante y sesgo de las comunidades por el empleo de cebadores dirigidos; y así poder responder quiénes las componen y cuales funciones podrán realizar.

El análisis metagenómico es una estrategia para analizar de manera conjunta a la mayoría de los organismos que componen a una comunidad. Es el análisis genómico de los ensamblajes de organismos sin el sesgo inicial de *PCR*. Se analizan fragmentos de genomas, que son el conjunto de genes de un organismo. El gen es la unidad mínima de herencia y visto desde el punto estructural es una secuencia de ADN que puede ser codificante (ARN mensajero) o no codificante (ARN ribosomal o ARN transferencia) (Watson, 2014; van der Walt *et al.*, 2017). Estas secuencias son comparadas contra bases de datos para averiguar si son similares, indicando relaciones funcionales o evolutivas (homología) si proceden de un ancestro común. Éste es un concepto clave en genómica (Koonin, 2005) y necesario para poder hacer inferencias sobre la información depositada en dichas bases de datos (anotación). En otras palabras, es ponerle 'nombre y apellido' a las secuencias. Con la mejora constante en los costos y calidad de las tecnologías de secuenciación, el enfoque de

secuenciación tipo '*shotgun*' especializado para la metagenómica, es el método de elección para estudiar las comunidades de microorganismos de diversos hábitats (Vollmers, Wiegand y Kaster, 2017).

Generalidades del sitio de estudio

La región antártica se caracteriza por su aislamiento geográfico y climático. Es uno de los ambientes terrestres física y químicamente más extremos para ser habitado por seres vivos. La mayor parte del continente está cubierto por hielo glaciar; sin embargo, existen áreas libres de hielo que constituyen ~10% de la masa terrestre continental que se distribuye de forma irregular alrededor de los márgenes costeros (Cary *et al.*, 2010).

Las capas de hielo polares y los gradientes de temperatura entre el ecuador y los polos modulan el clima del planeta. Las regiones polares están sufriendo reducciones importantes en la superficie cubierta por el hielo como consecuencia de los cambios climáticos y globales que experimenta el planeta (Turner, Overland y Walsh, 2007). Las mayores tasas del calentamiento atmosférico que ocurren en el hemisferio sur, provienen de la región Antártica Occidental. Particularmente en la Península Antártica, dichas tasas son seis veces mayores a la media global, con un aumento aproximado en la temperatura de 0.55 °C por década (Turner, Overland y Walsh, 2007). Las regiones que sufren este fuerte calentamiento presentan retracción glacial y derretimiento de hielo (Cary *et al.*, 2010; Cowan *et al.*, 2014a), exponiendo terrenos prístinos que están siendo colonizados rápidamente por comunidades de microorganismos que forman biofilms y tapetes microbianos.

'McMurdo Dry Valleys' ("MDV") es la región libre de hielo más grande en el continente antártico, con área de ~4000 km², localizada en el margen costero entre los ~77°-79°S y ~160°-164°E. El paisaje se compone por glaciares, valles, lagos y corrientes efímeras originadas por el deshielo de los glaciares. La temperatura promedio es de -18°C, con oscuridad continua a mitad del invierno austral y luz solar continua a mitad del verano austral. Debido a la acción de los vientos catabáticos no hay acumulación de nieve ni precipitación, al provocar que la humedad se sublime y no pueda acumularse. La zona presenta suelo arenoso-rocoso y depende de los glaciares para mantener la biogeoquímica del suelo, las corrientes y lagos que lo componen. Es un área de especial importancia ambiental debido a las características únicas derivadas del clima extremo, topografía y posición geográfica, por lo que se considera un desierto polar extremo (Fountain *et al.*, 2010; Chan *et al.*, 2013; Cavicchioli, 2015; Wei *et al.*, 2016) (Fig. 1).

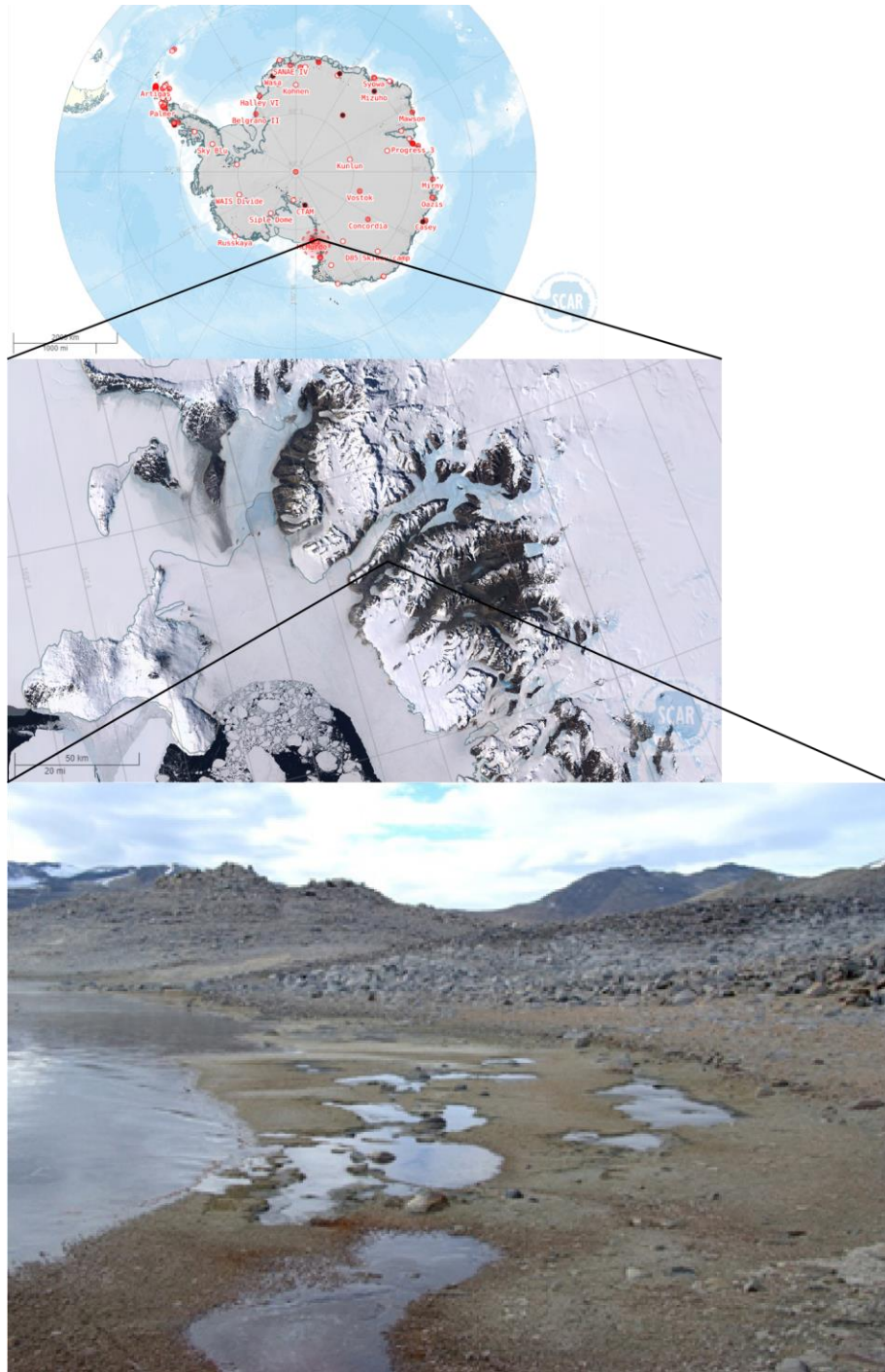


Figura 1. Posición geográfica de McMurdo Dry Valley, imagen satelital en mosaico LIMA (Landsat Image Mosaic of Antarctica) de la zona de McMurdo Dry Valley (Antarctic Digital Database Map Viewer SCAR) y paisaje de las corrientes efímeras, formadas en el verano austral por el deshielo de los glaciares, donde están asociados tapetes microbianos en Dry Valley (Don A. Cowan Creative Commons Attribution License (CC BY)).

El sustrato rocoso es el mayor hábitat en "MDV", ha sido caracterizado por una alta riqueza bacteriana en las comunidades microbianas, pero limitada en diversidad eucarionte, donde se incluyen nematodos, hongos, rotíferos y tardígrados (Buelow *et al.*, 2016). La composición trófica recae principalmente en las comunidades microbianas, siendo la biota dominante. Las comunidades microbianas son vitales, ya que representan las principales fuentes de entrada de nutrientes. También son bioindicadores de las fluctuaciones ambientales y son mediadores de los principales ciclos biogeoquímicos, particularmente del carbono y nitrógeno (Cowan *et al.*, 2014a; de Scally *et al.*, 2016). La mayoría de la biomasa en "MDV" se produce en rocas y suelos, aunque también se da en afluentes y pozas temporales formadas a partir del deshielo en el verano austral (Niederberger, Sohm, T. E. Gunderson, *et al.*, 2015; Wei *et al.*, 2016). Se tiene la hipótesis que las comunidades microbianas de "MDV" son relictos de antiguos lagos formados aproximadamente 40000 años (Fountain *et al.*, 1999; Zaikova *et al.*, 2019).

Generalidades de los tapetes microbianos

Los tapetes microbianos son comunidades estratificadas verticalmente, compuestas por una gran diversidad de grupos funcionales de microorganismos que están embebidos en una matriz orgánica. Crecen en el fondo de cuerpos de agua, cerca de la superficie. El mayor componente de biomasa de los tapetes microbianos está constituido por cianobacterias filamentosas y unicelulares. La materia orgánica generada a partir de las cianobacterias es la base de la estructura trófica de la comunidad microbiana. Generalmente se consideran

análogos no litificados de los estromatolitos, estructuras fósiles datadas en 3500 millones de años, por lo que representan a los ecosistema más antiguos conocidos (Stal, 2012; Bolhuis, Cretoiu y Stal, 2014).

Las cianobacterias son los organismos productores primarios clave en la región Antártica, debido a su capacidad de fijar CO₂ y N₂, y que soportan ciclos de congelación-descongelación. Tienen una latencia prolongada en invierno, resisten temperaturas extremadamente bajas aproximadamente de -60°C y soportan la alta radiación solar en el verano austral. Están presentes en diversos hábitats: suelos, glaciares, corrientes efímeras y lagos cubiertos de hielo. Son el componente principal para la formación de tapetes microbianos (Taton *et al.*, 2003; Komárek, Elster y Komárek, 2008; Michaud, Šabacká y Priscu, 2012), los cuales están también compuestos por una alta diversidad de Proteobacteria, Chloroflexi, Bacteroidetes, entre otros.

Justificación.

En general el continente Antártico es de los más prístinos y aislados por el clima y geografía (Cavicchioli, 2015). Sin embargo debido al cambio climático de las últimas décadas, la temperatura ha aumentado alrededor de 3°C en la parte occidental del continente Antártico (Turner, Overland y Walsh, 2007). Este aumento tiene efectos en la biodiversidad, dominada por microorganismos, debido al fenómeno de amplificación polar, el aumento en la temperatura tiene efectos significativos en el paisaje, principalmente en la desglaciación, modificando las condiciones naturales y alterando el ecosistema (Barrett *et al.*, 2008; Buelow *et al.*, 2016). El panorama general es el aumento de zonas libres de hielo que serán colonizadas por comunidades microbianas, sin embargo, las

comunidades contemporáneas están amenazadas debido a que los cambios ocurren en un lapso corto de tiempo y no se ajustan al tiempo de adaptación de las comunidades. Bajo este contexto, este trabajo es un aporte sobre el conocimiento de las comunidades microbianas formadoras de tapetes microbianos en "McMurdo Dry Valley's" desde una perspectiva independiente de cultivo y aislamiento de cepas como lo es la metagenómica.

Hipótesis

Los tapetes microbianos que se desarrollan en diferentes valles de la región de "McMurdo Dry Valleys" mantendrán una estructura taxonómica y funcional similar sin importar su posición geográfica.

Objetivo

Llevar a cabo un análisis de la composición taxonómica y funcional de los tapetes microbianos de diferentes valles en "McMurdo Dry Valleys" a partir de una estrategia metagenómica, enfocada en rutas metabólicas de formación de biomasa a partir de la fijación del N y C.

Objetivos particulares

- Analizar la diversidad alfa y beta de los tapetes microbianos a partir de la anotación de secuencias codificantes contra la base RefSeq a nivel género.
- Conocer la composición taxonómica de los tapetes microbianos a nivel phylum.
- Analizar el potencial genético de los taxa en la fijación de C y N a partir del ensamble de los metagenomas.
- Recuperar los 'MAGs' ('metagenome-assembled genomes') a partir del ensamble de los metagenomas.

Antecedentes

Los estudios pioneros sobre las comunidades de tapetes microbianos en "MDV" se hicieron basados en observaciones macro y microscópicas. "MDV", se describió como un hábitat 'estéril' y con muy poca diversidad biológica hasta después de los años 70s. Con el avance de las técnicas de cultivo, microscopía y ecología molecular, se sugirió que las comunidades microbianas que forman los tapetes microbianos son en realidad muy diversas (Taton *et al.*, 2003; De Los Ríos *et al.*, 2004; Cary *et al.*, 2010).

Durante el verano austral el deshielo de los glaciares en "MDV" mantiene corrientes y arroyos que son puntos importantes de actividad biológica, en éstos se desarrollan las comunidades formadoras de tapetes microbianos. Las corrientes efímeras son la única fuente de agua líquida en "MDV", forman zonas húmedas a lo largo y ancho de las corrientes llamadas zonas hiporreicas. Una característica principal es la presencia de cianobacterias y biofilms en estas zonas. La biota de estas comunidades está compuesta por bacterias, microalgas, hongos y microeucariontes heterótrofos, siendo los procariontes los que dominan en número y biomasa (Niederberger, Sohm, T. E. Gunderson, *et al.*, 2015). Durante el invierno austral estas comunidades permanecen secas/congeladas e inactivas. Se piensa que son una fuente de carbono y nitrógeno de origen orgánico en el ecosistema por dispersión eólica como detritos, siendo la mayor fuente de nutrientes en un ecosistema oligotrófico como "MDV" (Niederberger, Sohm, T. Gunderson, *et al.*, 2015; Kohler *et al.*, 2016).

Varios estudios de comunidades microbianas en "MDV" están basados en la reconstrucción taxonómica de secuencias no codificantes -como el 16S ARNr-

(White *et al.*, 2015; Wei *et al.*, 2016) y análisis funcionales metabólicos implicados en el ciclo del carbono (*ppc*, *rbcl*) y nitrógeno (*nifH*, *nirK*) (Rojas *et al.*, 2009; Niederberger, Sohm, T. Gunderson, *et al.*, 2015). Gran parte de estas aproximaciones están enfocadas a la diversidad bacteriana de sustratos (Smith *et al.*, 2006; Babalola *et al.*, 2009; Cary *et al.*, 2010; Wei *et al.*, 2015) y lagos antárticos (Van Trappen *et al.*, 2002; Yau *et al.*, 2013; Tytgat *et al.*, 2014; Jungblut *et al.*, 2016). Por lo tanto, existen pocos estudios que han analizado la composición taxonómica y funcional de los tapetes microbianos que se desarrollan en “MDV” (Taton *et al.*, 2003; Chan *et al.*, 2013). Estos trabajos indican que hay una plasticidad fenotípica en estrategias de autotrofia, heterotrofia y diazotrofia que dan soporte a las comunidades, y que el grupo taxonómico principal son las cianobacterias. Sin embargo, tienen un enfoque relacionado a métodos clásicos como la microscopía, el cultivo y aislamiento de cepas. Por lo que hace falta estudios desde una perspectiva metagenómica, sobre la diversidad genética de las comunidades que forman tapetes microbianos en “MDV”.

Métodos

Obtención de secuencias metagenómicas

Para este estudio, se analizaron cinco metagenomas derivados de tapetes microbianos de “MDV” encontrados en las corrientes de deshielo generadas en el verano austral. En el verano austral de 2015, en 1m² fueron colectadas 5 submuestras por sitio, en ‘Taylor Valley’ (1 muestra) y ‘Garwood Valley’ (4 muestras) (Figura 2). El ADN ambiental total de la comunidad de cada muestra fue extraído con el kit *MoBio PowerSoil DNA isolation QIAGEN* y el protocolo

adaptado CTAB (bromuro de cetil trimetil amonio) y secuenciado por el método 'Shotgun metagenomic sequencing' en la plataforma 'Illumina HiSeq 2500' con el servicio 'Yale Center for Genome Analysis (YCGA)', de donde se obtuvieron lecturas pares ('forward' y 'reverse') en crudo ('raw reads') para cada muestra: S010, S011, S012, S013 y S014 (Tabla 1).

Tabla 1. Información de los metagenomas recuperados después de la secuenciación.

Muestra	Lecturas pares obtenidas en crudo
S010	45,302,832
S011	43,073,298
S012	60,291,510
S013	49,993,630
S014	36,888,162

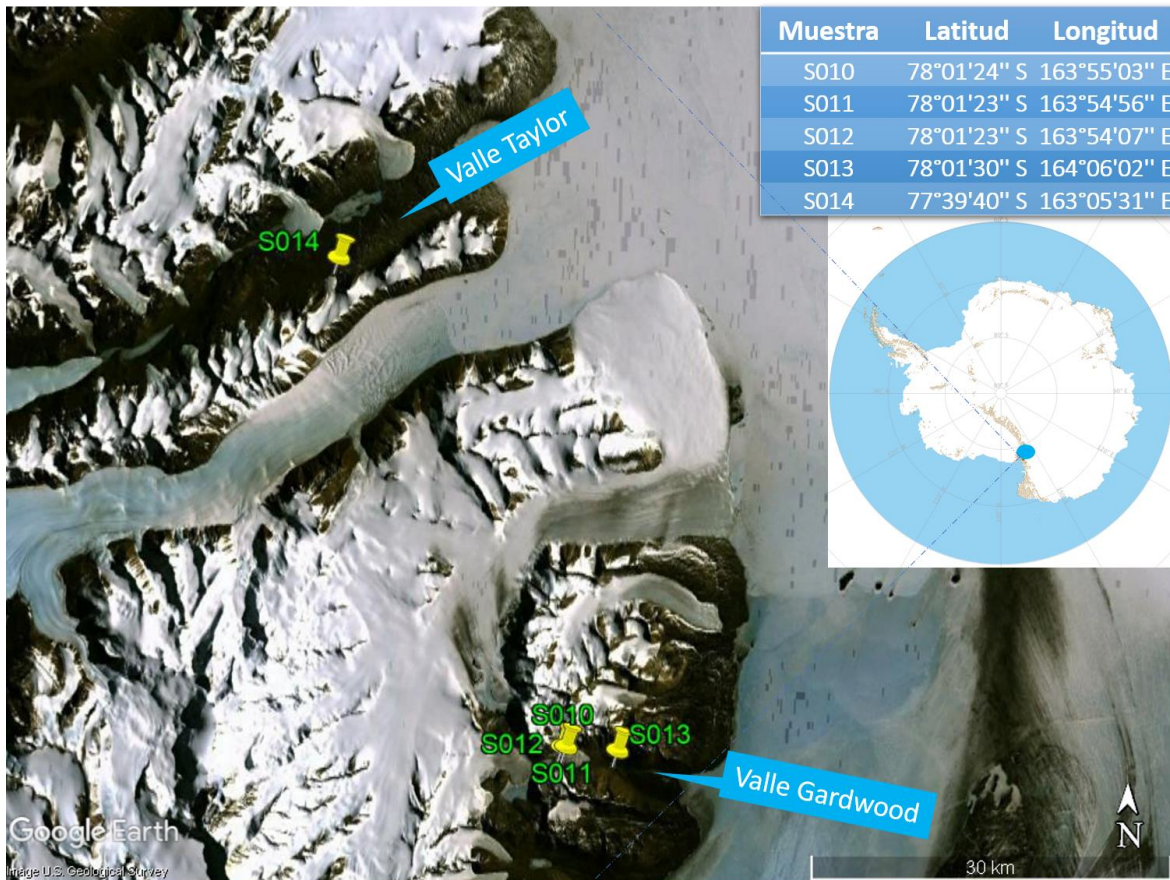


Figura 2. Ubicación de "MDV" en el continente Antártico (derecha, círculo azul) e imagen satelital Landsat/Copernicus Google Earth® de los lugares donde fueron colectados los tapetes microbianos.

Análisis de diversidad basado en lecturas

Para el análisis basado en lecturas ('reads'), éstas fueron subidas en formato *fastq* y bajo la compresión *gzip*, con el procedimiento operativo estándar para su anotación y recuperación en una cuenta personal de MG-RAST ('Metagenomic Rapid Annotations using Subsystems Technology') version 4.0.2 (Meyer et al., 2008) en mayo de 2018 usando el 'API' de MG-RAST (<http://api.metagenomics.anl.gov/api.html>):

```
curl -X POST -H "auth: tu_webkey_de_mgrast" -F
"upload=@tus_secuencias.fq.gz" "http://api.mg-rast.org/inbox".
```

En la cuenta personal de MG-RAST se eligieron las opciones: 'remove artificial replicate sequences, remove any host specific species sequences = *H. sapiens*, NCBI v36, remove low quality sequences, lowest phread score ≥ 15 , trimmed low phread score bases < 5 '. Después de ser procesados en los servidores de MG-RAST, se recuperaron los conjuntos de datos anotados contra la base de datos 'RefSeq' (Pruitt, Tatusova y Maglott, 2007).

Los archivos de anotación fueron cargados a partir de la página de análisis de MG-RAST (<https://www.mg-rast.org/mgmain.html?mgpage=analysis>), seleccionando la base de datos 'RefSeq' para el análisis taxonómico. Se eligieron las opciones *e-value* $\geq 1e-05$, *%ident* ≥ 70 , *length* ≥ 50 , *besthit* y el filtro a nivel género; se exportaron las tablas de datos de cada metagenoma anotado con la opción *TSV detailed*, que es un archivo con las columnas de los niveles taxonómicos, las abundancias, *e-value*, longitud de alineamiento y porcentaje de identidad.

En R (R Core Team, 2019) con el paquete 'tidyverse' (Wickham, 2017), se enlistó cada archivo de anotación en un objeto, se importó en una lista cada archivo y se unieron en un *dataframe* único. Se seleccionaron las columnas de los nombres de los metagenomas, los niveles taxonómicos y las abundancias, se hizo un filtrado de los niveles taxonómicos que no corresponden a los organismos que componen las comunidades de tapetes microbianos ("Annelida | Arthropoda | Brachiopoda | Chaetognatha | Chordata | Cnidaria | Echinodermata | Hemichordata | Onychophora | Placozoa | Platyhelminthes | Porifera | Priapulida"). A partir del *dataframe* filtrado se generó una matriz de abundancia

expandiendo por metagenoma los niveles taxonómicos y llenando con 0 las filas no compartidas de los niveles taxonómicos entre los metagenomas.

La estimación de la diversidad se realizó a partir de la matriz a nivel género con la función `estimateR()` del paquete 'vegan' (Oksanen *et al.*, 2019), usa la estimación basada en los conteos por sitio, se obtienen géneros observados, géneros potenciales a partir del estimador Chao y su error estándar.

$$\text{Chao } S_P = S_O + a1*(a1-1)/(2*(a2+1))$$

S_P diversidad potencial, S_O número observado de géneros, $a1$ y $a2$ número de géneros que están solo una o dos veces en cada metagenoma.

Se realizó la estimación de la diversidad en cada metagenoma a partir del índice de Shannon con la función `diversity()` del paquete 'vegan' (Oksanen *et al.*, 2019). El índice de esta definido como:

$$H = -\sum p_i \log(b) p_i$$

Donde p_i es la abundancia proporcional de los géneros i y b es la base del logaritmo.

Se hizo una transformación de la matriz de géneros para obtener la abundancia relativa y presencia/ausencia entre metagenomas con la función `decostand()` (`method = "total"` y `method = "pa"`) para hacer la estimación de la diversidad beta (matrices de distancia entre sitios) a partir de la distancia (índice de disimilitud) entre las abundancias relativas por el método 'bray' y la presencia/ausencia de géneros a partir del método 'jaccard' con la función `vegdist()` del paquete 'vegan' (Oksanen *et al.*, 2019).

Distancia "bray" $d[jk] = (\text{sumabs}(x[ij] - x[ik])) / (\text{sum}(x[ij] + x[ik]))$; $d[jk]$ la disimilitud entre sitios; $x[ij]$ y $x[ik]$ son la cantidad de géneros (columna) i en los sitios (filas) j y k .

Distancia "jaccard" $2B / (1 + B)$; donde B es la disimilitud Bray-Curtis y se basa en la desviación binomial bajo una hipótesis nula de que las dos comunidades comparadas son iguales.

Se realizó un ordenamiento de escalamiento multidimensional no métrico NMDS para ver si existe algún patrón de agrupamiento entre las comunidades con la función `metaMDS()` del paquete 'vegan' (Oksanen *et al.*, 2019).

La composición de la comunidad fue obtenida a partir del *dataframe* filtrado, seleccionando las columnas del nombre de metagenoma, nivel taxonómico phylum y la abundancia; se agruparon por metagenoma y phylum; se sumaron el total de abundancias y se generó una nueva columna para obtener la frecuencia relativa dividiendo las abundancias por phylum entre el total de las abundancias de cada metagenoma; se filtraron los valores ≥ 0.01 de frecuencia relativa por phylum; fue representada en un gráfico de barras apiladas utilizando el paquete 'tidyverse' (Wickham, 2017) para observar la composición de cada metagenoma a nivel phylum mayores al 1% de frecuencia relativa.

Análisis basado en anotación de secuencias (ensamble)

Para el análisis con un enfoque en el ensamble de las lecturas, se verificó la calidad las lecturas en crudo antes y después de ser filtradas, recortadas y fusionadas con Fastqc 0.11.8

(<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), se removieron adaptadores, fueron recortadas las primeras 10 bases y filtradas a una longitud \geq 75 bases y calidad Phred \geq 20 utilizando el programa BBDuk del conjunto de herramientas BBTools version 38.22 (<https://jgi.doe.gov/data-and-tools/bbtools/>). Las lecturas pares filtradas y recortadas fueron fusionadas con FLASH v1.2.11 (Magoč y Salzberg, 2011).

Las lecturas que pasaron el control de calidad fueron ensambladas con MEGAHIT v1.1.3 (Li *et al.*, 2015) con los siguientes parámetros; --presets meta-sensitive --min-contig-len 1000. Se predijeron secuencias codificantes con PRODIGAL v2.6.3 (Hyatt *et al.*, 2010) en las secuencias contiguas o 'contigs' con los parámetros -c -p meta. Las secuencias codificantes se anotaron utilizando eggNOG-mapper v2.0 contra la base de datos eggNOG v5.0 (Huerta-Cepas *et al.*, 2017, 2019). Se extrajeron los ortólogos funcionales KEGG y las predicciones de los grupos taxonómicos de las salidas del archivo de anotación. Éstos fueron mapeados en 'KEGG Mapper' v4.0 (Kanehisa *et al.*, 2016) para obtener una reconstrucción de las rutas metabólicas referentes a metabolismo, procesamiento información ambiental, procesamiento información genética y procesos celulares; también se mapearon las secuencias codificantes anotadas como ortólogos KEGG contra los módulos funcionales KEGG sobre metabolismo energético y la tabla de jerarquía 'Brite' br01620 sobre capacidades fotosintéticas y quimiosíntesis, para tener una perspectiva del potencial funcional entre los metagenomas. Se hizo un análisis comparativo de los ortólogos KEGG que son únicos y compartidos entre los metagenomas, se construyó un diagrama de Venn listando los KO en la página (<http://bioinformatics.psb.ugent.be/webtools/Venn/>)

y gráficas de barras de la asignación funcional compartida y única entre los metagenomas, para visualizar el core funcional y las diferencias entre los metagenomas. Se hizo una gráfica de la relación taxonómica y la asignación funcional basada en el mapeo de los KO en los módulos KEGG referentes al metabolismo energético: fijación carbono, metabolismo metano, metabolismo nitrógeno, metabolismo azufre y fotosíntesis (Anexo 1).

Se buscaron secuencias homólogas de los modelos ocultos de Markov 'hmm', basados en Peura, et al (Peura *et al.*, 2015) (Anexo 2). Los 'hmm' fueron descargados de la base de datos Pfam 31.0 (El-Gebali *et al.*, 2019). Cada modelo se utilizó para buscar homólogos con HMMER 3.1b2 (Eddy, 1998), con el comando `hmmsearch` en los *contigs* ensamblados de cada metagenoma, para buscar cualidades sobre la transformación de carbono y nitrógeno. Cada secuencia homóloga predicha fue mapeada en los *contigs* ensamblados para anotar y asignar taxonomía contra la base de datos '*RefSeq non-redundants proteins*' con CAT (*Contig Annotation Tool*) (von Meijenfeldt *et al.*, 2019).

Se predijeron las secuencias 18S de ARN ribosomal en los '*contigs*' ensamblados de cada metagenoma, para buscar genes eucariontes a partir de un modelo oculto de Markov utilizando Barrnap 0.9 (Seeman, 2019) con las opciones `--kingdom euk --reject 0.9`. Los 18S ARNr predichos se '*blastearon*' contra la base de datos 'nr' de 'NCBI' para identificar las secuencias biológicas similares y su taxonomía con *blastn* 2.9.0+ (Altschul *et al.*, 1990) con las opciones: '`-remote -db nr -evaluate 1e-50 -word_size 64 -perc_identity 0.9 -penalty -3 -reward 1 -outfmt 7 -max_hsps 10 -qcov_hsp_perc 0.7`'. Después de ubicar y anotar la taxonomía de las secuencias se alinearon con MAFFT v7.271 (Rozewicki *et al.*, 2019) donde se

obtuvo un archivo de alineamiento múltiple multifasta *archivo.msa* del que se infirió una filogenia con *FastTree Version 2.1.10 No SSE3* (Price, Dehal y Arkin, 2010) con las opciones estándar. Con el archivo de inferencia filogenética resultante en formato 'newick' se dibujó un árbol filogenético con iTOL v5 (Letunic y Bork, 2019) para visualizar las secuencias predichas 18S ARN ribosomal en todos los metagenomas y caracterizar algunos de los organismos eucariontes dentro de las comunidades (Figura 9).

Recuperación de 'MAG's' ('Binning')

Para recuperar genomas a partir del ensamble de metagenomas, las lecturas que pasaron el control de calidad se ensamblaron con SPADES 3.13.0 (Bankevich *et al.*, 2012) con las opciones `--only-assembler -k 21,31,41,51,61,71`. Las secuencias contiguas o 'contigs' fueron recortadas a ≥ 1000 pb utilizando el script `filter_fasta_by_length.sh` *akutils-v1.2.4* (Andrews, 2018). Las secuencias contiguas fueron alineadas contra las lecturas con Bowtie2 v2.3.5 (Langmead *et al.*, 2009) y procesadas con *samtools-1.9* (Li *et al.*, 2009) para obtener archivos `bam.sort`. La agrupación de las secuencias contiguas o 'binning' se hizo con el script `runMetaBat.sh` 'contigs' 'bam.sort' con MetaBAT2 (Kang *et al.*, 2015). Las secuencias contiguas agrupadas o 'bins' fueron evaluados con CheckM v1.0.12 (Parks *et al.*, 2015) y se filtraron a una completitud ≥ 90 , contaminación ≤ 10 y heterogeneidad ≤ 10 . Se predijeron secuencias no codificantes 5S, 16S, 23S ARN ribosomal en los 'bins' filtrados con *barrnap* v0.9 (<https://github.com/tseemann/barrnap>) y ARN de transferencia con *tRNAscan-SE* (Chan y Lowe, 2019). Se consideraron como genomas potenciales a los 'bins'

filtrados con presencia 23S, 16S y 5S rARN y al menos 18 ARNt, completitud ≥ 90 y contaminación < 5 (Bowers *et al.*, 2017).

El proceso de calidad de lecturas, ensamble, agrupamiento, predicción y 'binning' fue realizado en un servidor GNU/Linux x86_64 con 32 núcleos y 256 GB RAM.

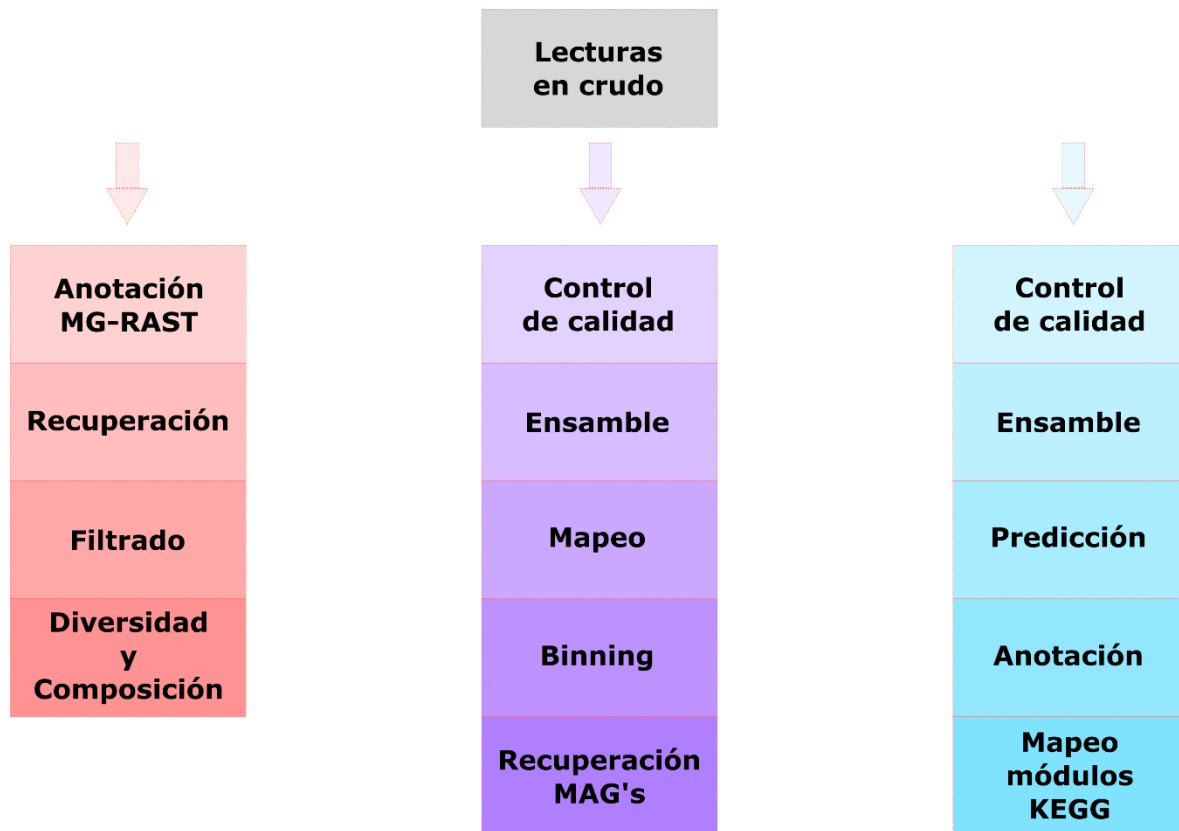


Figura 3. Procedimiento general para el análisis de los metagenomas.

Resultados

Análisis de diversidad

La composición taxonómica potencial a nivel genético mayor a 95% a partir de la frecuencia relativa en todos los metagenomas está representada por los phylum: Actinobacteria, Bacillariophyta, Bacteroidetes, Chloroflexi, Cyanobacteria, Firmicutes y Proteobacteria. La diversidad potencial de los tapetes microbianos de

la región Antártica de "MDV" indica que los phyla más abundantes son Cyanobacteria, Bacteroidetes y Proteobacteria en conjunto representan más del 75% de la composición de las comunidades formadoras de tapetes microbianos, pero no hay una composición definida entre las muestras; por ejemplo, el phylum Cyanobacteria representa solo un 7% de la muestra S010 y un 63% en contraste en la muestra S013 (Fig. 4).

De manera general, el número de géneros potenciales reportados está por arriba de 770, y a partir de las asignaciones únicas o en dos ocasiones reportadas con el estimador Chao1, hay un estimado de hasta 920 géneros. La diversidad alfa presenta un valor promedio de 4.4 respecto al índice de Shannon H' (Tabla 2). Para la diversidad beta se obtuvieron matrices de disimilitud por abundancia relativa y presencia/ausencia entre las muestras (Tabla 3 y 4).

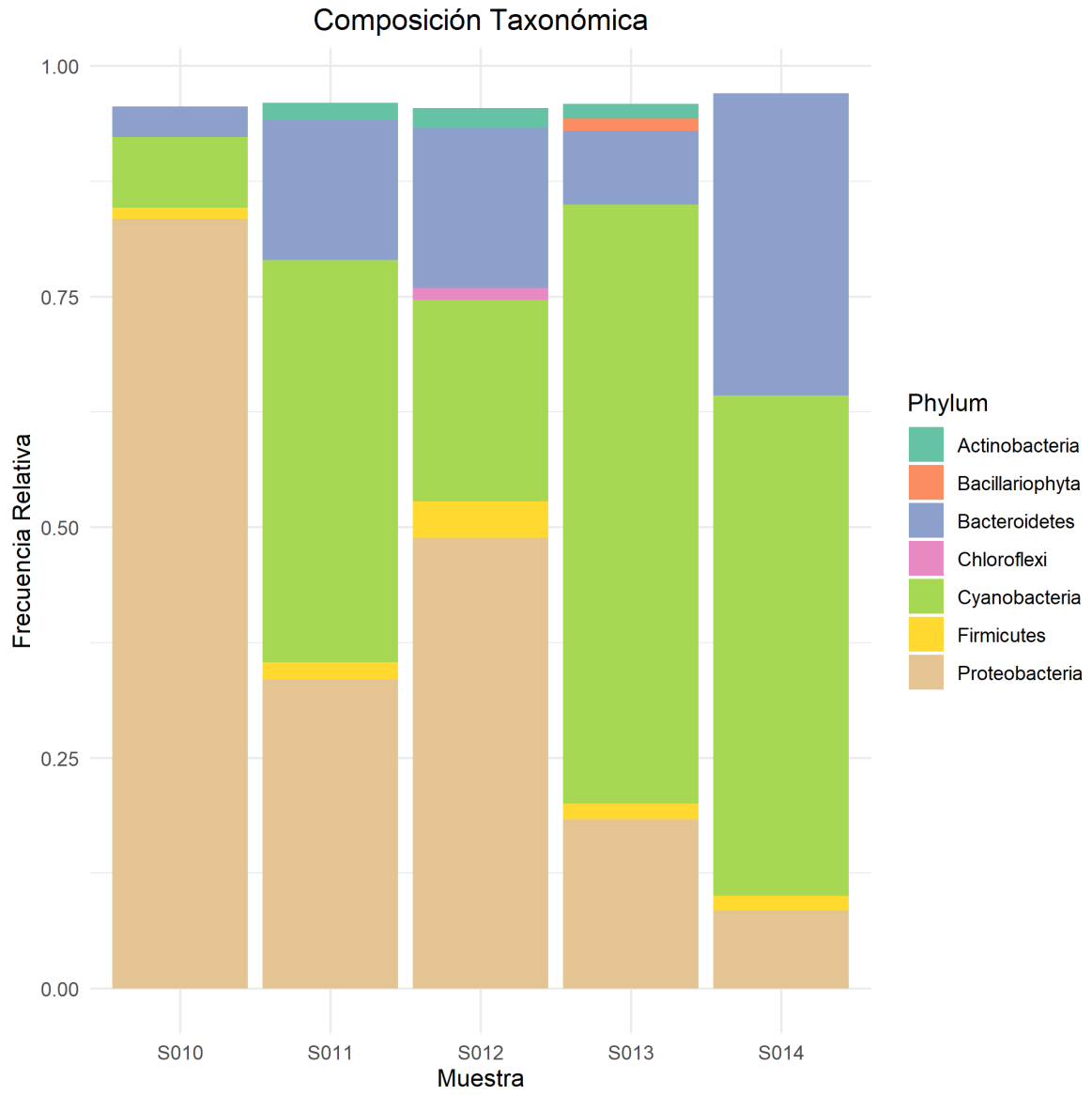


Figura 4. Gráfica de composición taxonómica entre los tapetes microbianos a nivel phylum

Tabla 2. Valores de diversidad de los tapetes microbianos a partir del nivel género.

	S010	S011	S012	S013	S014
Observado	830	859	869	847	770
Chao1	853.33	889.52	894.50	920.50	854.00
se.Chao1	10.02	12.09	11.51	28.98	33.76
H'	4.25	4.46	4.99	4.13	3.88

Tabla 3. Matriz de disimilitud de abundancia relativa de géneros potenciales entre los tapetes microbianos

	S010	S011	S012	S013
S011	0.55			
S012	0.45	0.34		
S013	0.68	0.36	0.46	
S014	0.79	0.30	0.52	0.42

Tabla 4. Matriz de disimilitud de presencia/ausencia de géneros potenciales entre los tapetes microbianos

	S010	S011	S012	S013
S011	0.12			
S012	0.12	0.11		
S013	0.12	0.11	0.12	
S014	0.15	0.14	0.15	0.16

Se obtuvo una filogenia basada en los homólogos encontrados en las secuencias de los ensamblajes de los metagenomas utilizando un modelo oculto de Markov 18S ARN ribosomal para mostrar la variedad de algunos eucariontes que forman parte de las comunidades de los tapetes microbianos (Figura 9). Entre los distintos

eucariontes hay ciliados, microalgas, rotíferos, tardígrados y nemátodos. Algunas de las secuencias no fueron ubicadas y otras fueron anotadas como muestras ambientales.

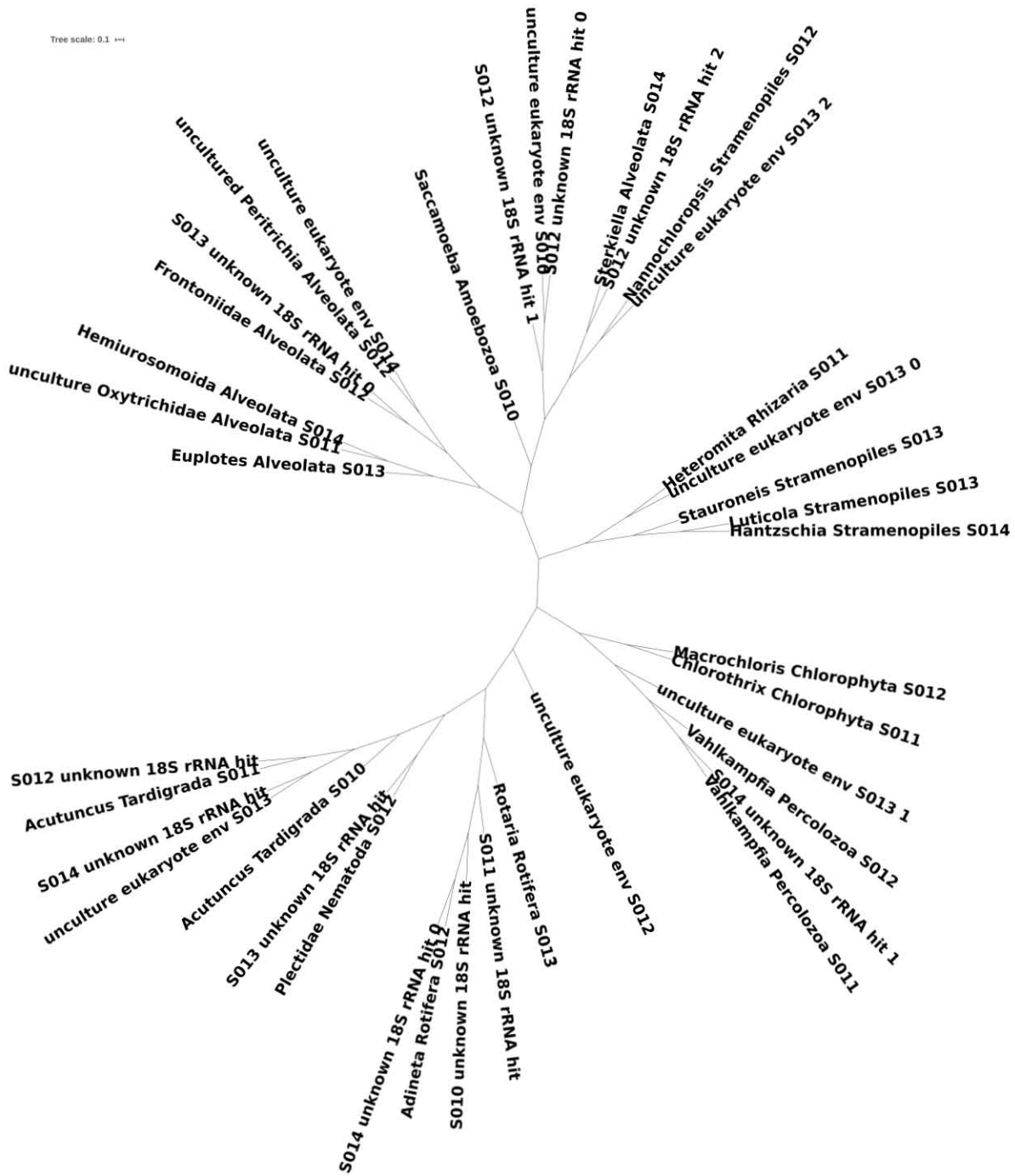


Figura 5. Filogenia basada en secuencias predichas 18S ARNr en los metagenomas MDV (S010,S011,S012,S013,S014).

Potencial funcional, capacidades metabólicas energéticas

A partir del ensamble y la predicción de secuencias codificantes, se recuperaron los ortólogos KEGG (KO) (Tabla 5) y su asignación taxonómica. Se realizó un diagrama de Venn para visualizar los KO compartidos y únicos entre los metagenomas (Figura 5). Los metagenomas comparten 4379 KO, el metagenoma S014 tiene más KO únicos 1174. En promedio se pudieron anotar aproximadamente un 75% de las secuencias codificantes predichas (Tabla 5).

Tabla 5. Resumen de la anotación contra los ortólogos KEGG (KO) de secuencias codificantes predichas entre los metagenomas.

Muestra	Secuencias codificantes predichas	Ortólogos asignados	Ortólogos no repetidos
S010	114457	95920	5724
S011	134367	110128	5682
S012	208001	176228	6287
S013	136875	113409	5856
S014	126554	106117	6587
		<i>Ortólogos únicos totales</i>	<i>8621</i>
		<i>Ortólogos compartidos</i>	<i>4379</i>

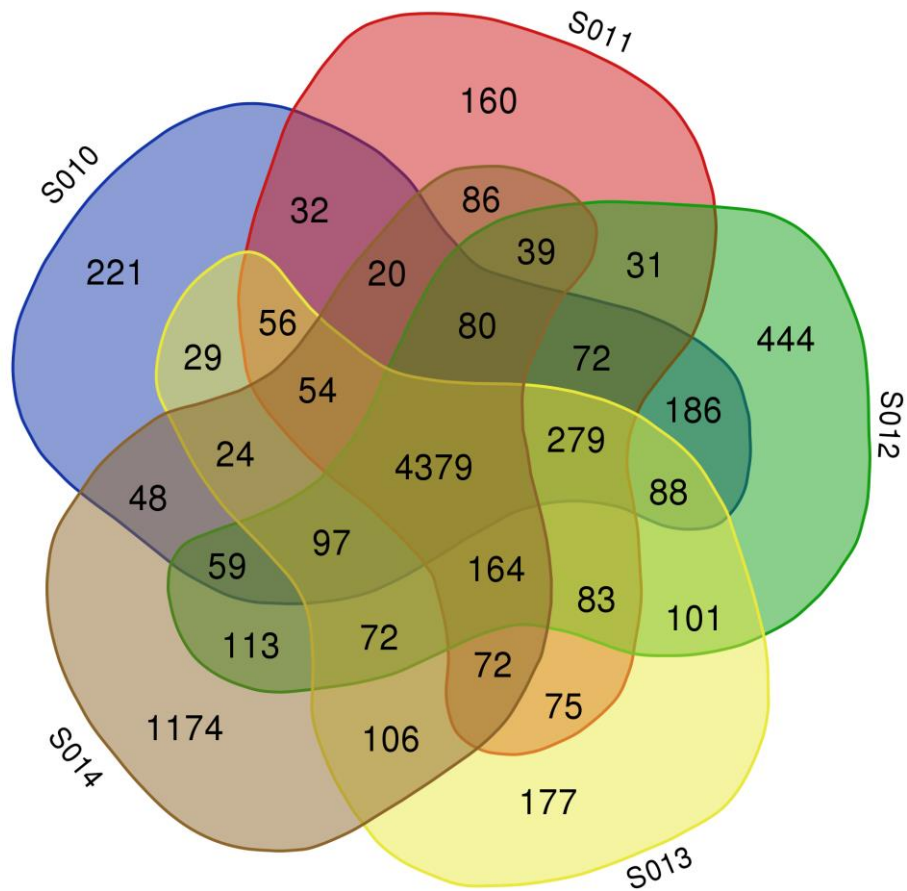


Figura 6. Diagrama de Venn comparativo de los KO no repetidos anotados entre los metagenomas.

Se graficaron las asignaciones funcionales de los KO únicos compartidos (core funcional) entre los metagenomas (Figura 6). El metabolismo tiene mayor asignación de KO, seguido de procesamiento de información ambiental, procesamiento de información genética y procesos celulares. Cada categoría tiene subcategorías (Figura 6).

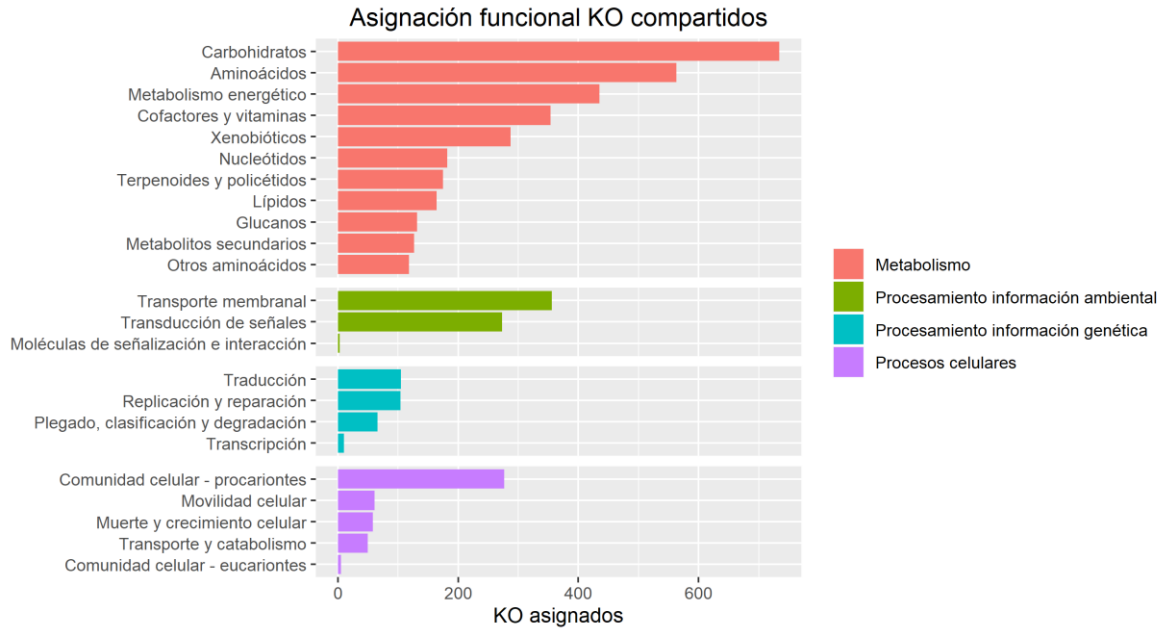


Figura 7. Gráfico de barras de la asignación funcional de los KO compartidos entre los metagenomas.

Se graficó la asignación funcional de los KO únicos de cada metagenoma (Figura 7). Los KO únicos del metagenoma S014 tiene una mayor asignación de funciones relacionadas con el procesamiento información ambiental en la subcategoría transducción de señales por arriba de 400 y también en el procesamiento de información genética y procesos celulares (Figura 7)

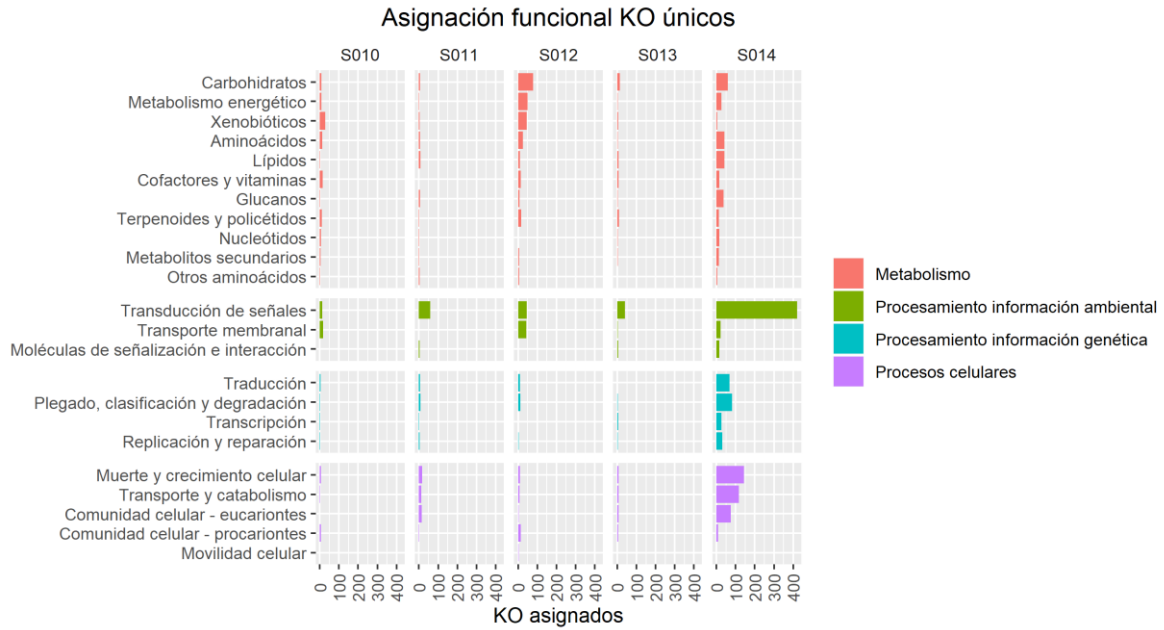


Figura 8. Gráfico de barras comparativo de la asignación funcional de los KO únicos de cada metagenoma.

La búsqueda de las capacidades metabólicas potenciales de fotosíntesis y quimiosíntesis se hizo a partir del mapeo contra los módulos KEGG completos, donde se pueden inferir el potencial funcional. Las capacidades fotosintéticas, fotosistemas I y II, fotosistema anoxigénico II y la fijación de carbono a partir del ciclo de Calvin, son compartidas en todas las muestras. Las capacidades de oxidación de azufre (por SOX) son compartidas en las muestras S010 a la S013; y metilótrofos por dos rutas distintas en las muestras S010, S011 y S012 (Tabla 6). Las capacidades fotosintéticas oxigénica y anoxigénica se comparten en todas las muestras (Tabla 7).

Tabla 6. Presencia/Ausencia de las capacidades fotosintéticas y de quimiosíntesis entre los metagenomas a partir del mapeo de las secuencias codificantes predichas en los módulos KEGG. verde = presencia

Nombre	Módulo	S010	S011	S012	S013	S014	Fenotipo
Fotosistema II	M00161	■	■	■	■	■	Organismos fotosintéticos
Fotosistema I	M00163						
Fotosistema anoxigénico II	M00597						
Fotosistema anoxigénico I	M00598						
Ciclo reductivo de las pentosas fosfato (Ciclo de Calvin)	M00165	■	■	■	■	■	Fijadores de carbono
Ciclo del ácido cítrico inverso (Ciclo Arnon-Buchanan)	M00173						
Ruta reductora de la acetil-CoA (Ruta Wood_Ljungdahl)	M00377						
Ciclo del 3-hidroxiacetil-CoA	M00376						
Ciclo del dicarboxilato-hidroxiacetil-CoA	M00374						
Ciclo hidroxipropionato-hidroxiacetil-CoA	M00375						
Oxidación tiosulfato por SOX, tiosulfato => sulfato	M00595	■	■	■	■	■	Oxidantes del azufre
Oxigenasa/reductasa azufre	K16952						
Reducción desasimilatoria del sulfato, sulfato => H ₂ S	M00596						Sulfato reductores
Metanogénesis, CO ₂ => metano	M00567						Metanógenos
Metanogénesis, metanol => metano	M00356						
Metanogénesis, acetato => metano	M00357						
Oxidación del metano, metano => formaldehído	M00174						Metanótrofos
Asimilación de formaldehído, Ciclo xilulosa-monofosfato	M00344						Metilótrofos
Asimilación de formaldehído, Ciclo ribulosa-monofosfato	M00345			■			
Asimilación de formaldehído, Ruta de la serina	M00346	■	■	■			
Nitrificación, amonio => nitrito	M00528						Oxidantes de amonio

Tabla 7. Presencia/Ausencia de la combinación de asignación de módulos KEGG entre los metagenomas. verde = presencia

Combinación de módulos	Definición	Fenotipo	S010	S011	S012	S013	S014
M00611	(M00161, M00163) M00165	Fotosíntesis oxigénica en cianobacterias					
M00612	M00597 M00165	Fotosíntesis anoxygenica en bacterias púrpura					
M00613	M00597 M00376	Fotosíntesis anoxygenica en bacterias verdes no del azufre					
M00614	M00598 M00173	Fotosíntesis anoxygenica en bacterias verdes del azufre					
M00617	(M00567, M00357, M00356, M00563)	Metanógenos					
M00618	M00377 M00579	Acetógenos					

Se graficó en un mapa de calor la relación entre la taxonomía y los KO anotados y asignados de manera completa y parcial, en los módulos KEGG referentes al metabolismo energético: fijación carbono, metabolismo metano, metabolismo nitrógeno, metabolismo azufre y fotosíntesis (Figura 8). La mayor parte de la asignación taxonómica está en bacterias, seguido de eucariontes y arqueas. El metabolismo referente a la fijación de carbono es donde hay más diversidad de phylum asociados. La asociación taxonómica relacionada con fotosíntesis solo esta asignada en 5 phylum: Bacteroidetes, Cyanobacteria, Proteobacteria en bacterias y Bacillariophyta y Chlorophyta en eucariontes.

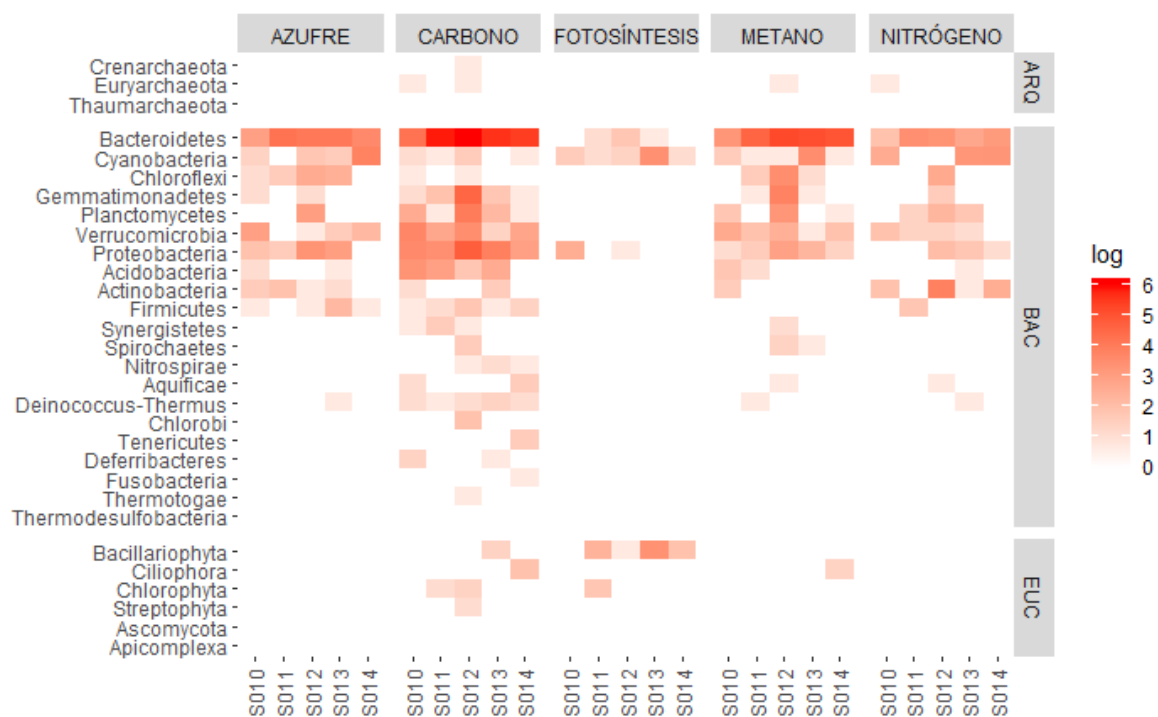


Figura 9. Mapa de calor de la relación entre la taxonomía (ARQ=arqueas, BAC=bacterias, EUC=eucariontes) a nivel phylum y la asignación de funciones en los metabolismos generales indicados, entre los metagenomas. La intensidad de color indica un mayor número de KO asignados al metabolismo indicado.

Genomas potenciales recuperados

Se lograron recuperar 19 'bins' de alta calidad de todos los ensamblajes y 5 de ellos son candidatos a ser publicables como 'MAG' metagenome assembled genome (Tabla 8). La mayoría de bins pertenece al phylum Bacteroidetes, 2 no tienen asignación más allá del nivel Bacteria, por lo que podrían resultar en alguna descripción nueva si se hace una curación exhaustiva.

Tabla 8. 'Bins' recuperados y asignados taxonómicamente, en verde los potenciales genomas a recuperar.

meta	Bin Id	linaje taxonómico	Complejidad	Contaminación	Heterogeneidad	Tamaño (pb)	# contigs	GC%	# genes predichos	ARNr	ARNt
S010	bin.10	k__Bacteria	95.54	2.23	0	3926501	212	51.5	3957	5S	50
S010	bin.12	f__Flavobacteriaceae	98.59	0.47	0	3013210	100	42.2	2732	5S	28
S010	bin.13	o__Burkholderiales	90.05	4.32	0	4412948	306	69.9	4303	NA	42
S010	bin.16	p__Bacteroidetes	96.49	1.89	0	5026436	357	41.6	4735	16S	46
S011	bin.13	c__Deltaproteobacteria	90.32	8.63	6.25	11095002	579	68.9	9162	NA	67
S011	bin.24	o__Cytophagales	94.79	3.24	0	4388336	249	34.3	4002	16S,23S,5S	44
S012	bin.1	k__Bacteria	98.39	3.49	0	3889065	92	37.6	3258	16S,23S,5S	37
S012	bin.25	k__Bacteria	97.7	8.8	0	5309075	432	40.2	5236	16S,23S,5S	36
S013	bin.17	p__Bacteroidetes	93.59	6.69	8.33	3223837	329	31.8	3043	NA	49
S013	bin.22	p__Bacteroidetes	94.93	4.19	10	4143511	341	41.7	3831	5S	24
S013	bin.23	c__Deltaproteobacteria	94.19	5.99	0	11343740	459	68.9	9197	5S	65
S013	bin.30	p__Bacteroidetes	97.04	3.69	10	6490409	279	38.1	5312	16S,23S,5S	41
S013	bin.38	o__Burkholderiales	98.6	2.45	10	4147935	170	64.3	3855	16S,5S	58
S014	bin.23	p__Bacteroidetes	98.65	1.59	0	3552431	78	30.2	3037	NA	25
S014	bin.24	p__Bacteroidetes	98.77	1.48	0	3272443	116	30.3	3125	16S	31
S014	bin.25	p__Bacteroidetes	95.59	1.77	0	4778365	215	35.5	4126	16S	32
S014	bin.29	o__Burkholderiales	96.59	1.81	0	4202460	275	57.8	4239	5S	51
S014	bin.34	o__Cytophagales	93.88	0.15	0	3496725	292	38.4	2702	NA	35
S014	bin.40	p__Bacteroidetes	95.24	0	0	3083273	162	31	2727	16S,23S,5S	33

Discusión

En este estudio presentamos una descripción del potencial taxonómico y funcional, posiblemente de los primeros que se realizan en "MDV", desde una perspectiva metagenómica (a nivel genético) entre las comunidades que forman tapetes microbianos en arroyos temporales (zona hiporreica) producto del deshielo en el verano austral en un sitio de condiciones extremas considerado un desierto polar. A pesar de que el ecosistema impone restricciones severas para el desarrollo de la vida, las comunidades se abren paso entre los espacios disponibles para colonizarlos, siendo caracterizadas como parte fundamental del ecosistema por ser la biota principal del ecosistema en MDV.

Como consecuencia del cambio climático registrado en las últimas décadas, en especial en la zona antártica (Rignot *et al.*, 2002; Amesbury *et al.*, 2017), el hábitat se ha transformado rápidamente debido al deshielo acelerado. Por lo que es necesario conocer el funcionamiento de las comunidades contemporáneas que están presentes en MDV. Estudios en MDV indican que existe una estructura comunitaria que consiste en productores primarios (fotoautótrofos) y degradadores/consumidores de no más de 3 niveles tróficos (Cary *et al.*, 2010; Cowan *et al.*, 2014b) por lo que el hábitat puede ser vulnerable al cambio climático.

Análisis de diversidad

Los índices de diversidad se basaron en los géneros que componen a los cinco metagenomas de tapetes microbianos (Fig. 4, Tabla 2). En conjunto hay en promedio 835 géneros en las cinco muestras. Algo interesante es que el promedio del índice de Shannon-Weaver (H') se mantiene por encima de tres en todas las

muestras, lo que es indicativo que las muestras son diversas (Jost, 2006; Kim *et al.*, 2017). Otro aspecto interesante es que la diversidad estimada por el índice Chao1 indica una cantidad mayor potencial en la diversidad de géneros de las muestras (Tabla 2) por encima de 47 ± 19 géneros en promedio. La estimación de la diversidad en estudios previos donde se han utilizado amplicones de 16S ARN para analizar tapetes microbianos de la zona de "MDV" indican valores similares en lo observado a nivel género (Varin *et al.*, 2010, 2012; Niederberger, Sohm, T. E. Gunderson, *et al.*, 2015).

La composición de la comunidad mayor a 1% de frecuencia relativa (Fig. 4) indica que los phyla con mayor proporción son Bacteroidetes, Cyanobacteria, Proteobacteria, los tres son mayores al 75% de abundancia relativa en todos los tapetes y en menor proporción: Actinobacteria, Bacillariophyta, Chloroflexi y Firmicutes. Esta composición es parecida a estudios previos de comunidades microbianas de muestras de sustrato rocoso Antártico basadas en 16S ARNr (Niederberger, Sohm, T. E. Gunderson, *et al.*, 2015; Wei *et al.*, 2016), donde no hay una abundancia relativa fija entre muestras, pero los taxa abundantes en las muestras son los mismos.

Otro punto importante es la diversidad beta, ya que la disimilitud entre muestras a partir de la abundancia relativa es muy desigual entre pares. Hay 0.48 de diferencia promedio. En contraste, la disimilitud entre pares a partir de la presencia/ausencia de géneros es de 0.13 promedio, es decir comparten cerca del 87% de géneros (Tabla 3 y 4). La pregunta que surge es ¿qué géneros son los que modifican el cambio en abundancia relativa entre los tapetes?, ¿existirá algún patrón de correlación positiva/negativa que sea indicativo quizá de

exclusión mutua o facilitación? Se realizó un análisis de ordenación escalamiento multidimensional no métrico (NMDS) para ver si existe algún patrón de agrupamiento entre los tapetes, sin embargo, el valor *stress* es muy cercano a cero, lo que indica que no hay la suficiente variabilidad entre las muestras y por lo tanto hay datos insuficientes. Podría ser muy tentador utilizar datos metagenómicos de otras investigaciones para completar el análisis NMDS, pero podría resultar no válida cualquier interpretación estadística al no haber sido obtenidas las muestras en circunstancias similares. La limitante es que para hacer las pruebas estadísticas pertinentes se requieren un mayor número de datos para obtener mayor variabilidad en las pruebas y tener algún estadístico que corrobore las afirmaciones.

Se realizó una filogenia basada en 18S ARN para visualizar a los eucariontes que existen en los tapetes microbianos (Fig. 9). Hay ciliados, microalgas, tardígrados, nematodos y rotíferos, es destacable que varias ramas se alinearon con muestras ambientales desconocidas que podrían resolverse debido a la cercanía con taxones ya identificados.

Potencial funcional y capacidades de metabolismo energético

Los tapetes microbianos presentan una distribución proporcional aproximada sobre capacidades metabólicas compartidas, basadas en las rutas KEGG, referentes al metabolismo 80%, procesamiento de información ambiental 10%, procesamiento de información genética 5% y procesos celulares 5% (Fig. 6). Lo interesante son las capacidades únicas entre los tapetes microbianos (Fig. 7), donde destaca la muestra S014, en donde hay una gran cantidad de asignaciones KO referentes al procesamiento de información ambiental, de

manera concreta a los sistemas de transducción de señal de dos componentes. Estos sistemas les permiten a los microorganismos censar y responder a estímulos ambientales (Tiwari *et al.*, 2017), sería interesante relacionar las variables ambientales del lugar de muestreo debido a que es el tapete muestreado separado por kilómetros de distancia (Fig. 2).

En la Antártida las comunidades microbianas son la base para la biodisponibilidad de los bloques fundamentales de la vida como el carbono y nitrógeno. Las rutas fundamentales para acumular biomasa son la fotosíntesis, fijación carbono y la fijación de nitrógeno (Cowan *et al.*, 2011; Niederberger, Sohm, T. Gunderson, *et al.*, 2015; Kohler *et al.*, 2016). Para obtener una interpretación cualitativa de estas capacidades en el metabolismo energético se mapearon las secuencias codificantes predichas de los metagenomas en los módulos KEGG (Tabla 6 y 7). Basados en los resultados, los tapetes microbianos de "MDV" poseen el potencial funcional para formar biomasa a partir de fotosíntesis, tanto oxigénica como anoxigénica, al igual que oxidar azufre y usar compuestos reducidos de un carbono como fuente de carbono (Tabla 6). Sin embargo, no todos los tapetes de "MDV" comparten, a partir del mapeo, la capacidad de oxidar azufre o la metilotrofia. Se buscaron cualidades sobre el metabolismo energético (Anexo 2 y 3) para ubicar a los grupos taxonómicos específicos que participan, sin embargo al obtener la clasificación de los *contigs* de los modelos ocultos de Markov, la gran mayoría no tiene una asignación *not classified* mas allá del nivel phylum.

A pesar de compartir el 82% de géneros en promedio entre pares; se esperaría una redundancia funcional debido a la diversidad taxonómica (Louca *et al.*, 2018). Esto puede deberse a la profundidad de secuenciación, en este caso el

promedio fue de 7X. En estudios donde el promedio de secuenciación supera 100X (Lee *et al.*, 2017; Stewart *et al.*, 2018) es posible observar un mayor número de características funcionales, inclusive la recuperación de taxa no abundantes en la comunidad.

A pesar de esta limitación, se puede obtener la asociación taxonómica y funcional a partir de los ortólogos funcionales (KO) clave en las rutas metabólicas energéticas (Anexo 1). En el mapa de calor (Fig. 8) se observa la comparación del potencial funcional de diversos microorganismos en distintos niveles taxonómicos. Destacan a nivel phylum Cyanobacteria, Proteobacteria y Bacteroidetes (Fig. 8 y Anexo 3). El grupo de Cyanobacteria es fundamental al proveer los elementos fundamentales de un sistema biológico en la comunidad (C y N); además, son el grupo principal considerado como formador de tapetes microbianos en sistemas fóticos (Bolhuis, Cretoiu y Stal, 2014). Por esta razón, es muy activa su participación en la fotosíntesis, la fijación de carbono y nitrógeno (Fig. 5). El grupo de los Bacteroidetes son un grupo de gran abundancia entre los tapetes microbianos, se encuentran asociados a los tapetes microbianos como organismos heterótrofos degradadores de EPS (*extracellular polymeric substances*) y producto de esta degradación liberan subproductos que sirven como nutrientes en la comunidad (Rojas *et al.*, 2009; Niederberger, Sohm, T. E. Gunderson, *et al.*, 2015).

Las proteobacterias son un grupo diverso y abundante en los tapetes microbianos, y en este estudio están asociadas a la fijación de carbono y a la fotosíntesis anoxigénica (Fig. 8). El grupo en general se le asocia con el ciclaje de azufre a través de la fotosíntesis anoxigénica (Rojas *et al.*, 2009; Varin *et al.*, 2010). Todas estas características revelan el potencial funcional que poseen los tapetes

microbianos como estrategias de colonización del hábitat, y en este caso, en uno totalmente adverso al ser seco y oligotrófico (de Scally *et al.*, 2016; Kohler *et al.*, 2016).

Binning

Los 'MAG' se recuperaron con herramientas de agrupamiento de secuencias contiguas ensambladas (contigs) y evaluados a partir de un árbol filogenético para ubicarlos con la mayor certeza posible (Parks *et al.*, 2015; Bowers *et al.*, 2017). A menudo quedan fragmentados al provenir de datos metagenómicos y es posible obtener 'contigs' quiméricos debido a la heterogeneidad de los organismos que componen a la comunidad. Lo interesante es que, a pesar de tener una mezcla por encima de 800 géneros, se lograron recuperar 'bins'. La gran mayoría pertenecen al grupo de los Bacteroidetes, una posible explicación puede ser por dimensiones y estructura de los organismos que componen la comunidad. Por ejemplo, *Flavobacterium* (Bacteroidetes) puede medir entre 3 y 5 micrómetros (<https://microbewiki.kenyon.edu/index.php/Flavobacterium>). En cambio, *Phormidium* (Cyanobacteria) puede extenderse más de 200 micrómetros (Loza *et al.*, 2013); además de que, poseen una densa capa de exopolisacáridos lo que dificulta la extracción de material genético. Solo cinco 'bins' son candidatos para reanalizar, reensamblar y a través de genómica comparativa verificar que sean ubicados como MAG (metagenomic assembled genome). Dos de ellos S012.bin.1 y S012.bin.25 (tabla 8); los cuales posiblemente sean algún microorganismo nuevo al no poder ubicarlos a un nivel taxonómico particular en la filogenia de más de 1000 organismos de genomas completos (Parks *et al.*, 2015).

Después de analizar de forma cuantitativa (basado en lecturas) y cuantitativa (basado en ensamble) no se puede considerar que la hipótesis se cumple debido a no existe la variabilidad suficiente en las muestras para obtener un resultado estadísticamente significativo. En el análisis metagenómico por secuenciación *shotgun* no hay un protocolo específico que permita hacer comparación en distintos proyectos como los realizados con amplicones basados en *earth microbiome project* (<http://www.earthmicrobiome.org/>), lo que si existe son buenas practicas para obtener resultados confiables, control de calidad, ensamble, predicción, anotación, nivel de cortes y filtrado de información.

Los tapetes microbianos de MDV poseen el potencial funcional para colonizar un ambiente tan extremo y son muy diversos, sin embargo, aún queda pendiente conocer si realmente este potencial funcional y taxonómico esta activo, saber cual es la dinámica espacio-temporal en la zona, realizar ensayos sobre fisiología y registro de variables ambientales para correlacionar la información y alguno relacionado con la transición de gradientes para saber si existe alguna partición de nicho o si son una metacomunidad.

Conclusiones

La secuenciación masiva "metagenomic shotgun DNA" ha permitido, especialmente en el campo de la ecología microbiana, aplicar técnicas de identificación y medición, basadas en el material genético, a toda la comunidad microbiana sin necesidad de separar o aislar a los miembros de la comunidad. El análisis por 'metagenomic shotgun DNA' permite describir la taxonomía, diversidad, genes y capacidades funcionales de una comunidad de organismos.

Los cinco tapetes microbianos estudiados de la región Antártica de McMurdo Dry Valleys son diversos y tienen similitudes taxonómicas y funcionales. Es interesante observar a todos los metabolismos necesarios para formar biomasa (fotosíntesis oxi y anoxigénica; fijación de carbono y fijación de nitrógeno) al igual que aquellos necesarios para la remineralización. Este estudio contribuye al entendimiento de las comunidades que forman tapetes microbianos, indicando que sus componentes biológicos les permiten subsistir en uno de los ecosistemas más extremos en el planeta. Igualmente, nos permite comprender cómo se forman y estructuran comunidades microbianas. Ante escenarios de cambio global, cada vez existe mayor tasa de pérdida de masa glaciar, lo que podría favorecer el crecimiento de comunidades microbianas formadoras de tapetes microbianos. A la fecha existe información acerca de su composición taxonómica y funcional, pero aún desconocemos la contribución de estas comunidades a los ciclos globales del C y N.

Literatura citada

Altschul, S. F. *et al.* (1990) "Basic local alignment search tool", *Journal of Molecular Biology*, 215(3), pp. 403–410. doi: 10.1016/S0022-2836(05)80360-2.

Amesbury, M. J. *et al.* (2017) "Widespread Biological Response to Rapid Warming on the Antarctic Peninsula.", *Current biology : CB*. Elsevier, 27(11), pp. 1616-1622.e2. doi: 10.1016/j.cub.2017.04.034.

Andrews, L. V. (2018) "akutils-v1.2: Facilitating analyses of microbial communities through QIIME.", *Zenodo*. 10.5281/zenodo.120579. doi: 10.5281/ZENODO.1205079.

Babalola, O. O. *et al.* (2009) "Phylogenetic analysis of actinobacterial populations associated with Antarctic Dry Valley mineral soils", *Environmental Microbiology*. Blackwell Publishing Ltd, 11(3), pp. 566–576. doi: 10.1111/j.1462-2920.2008.01809.x.

Bankevich, A. *et al.* (2012) "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing", *Journal of Computational Biology*. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, 19(5), pp. 455–477. doi: 10.1089/cmb.2012.0021.

Barrett, J. E. *et al.* (2008) "Persistent effects of a discrete warming event on a polar desert ecosystem", *Global Change Biology*, 14(10), pp. 2249–2261. doi: 10.1111/j.1365-2486.2008.01641.x.

Begon, M., Townsend, C. R. y Harper, J. L. (2005) *Ecology : from individuals to ecosystems*. Wiley InterScience.

Bolhuis, H., Cretoiu, M. S. y Stal, L. J. (2014) "Molecular ecology of microbial mats", *FEMS Microbiology Ecology*. Oxford University Press, 90(2), pp. 335–350. doi: 10.1111/1574-6941.12408.

Bowers, R. M. *et al.* (2017) "Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea", *Nature Biotechnology*. Nature Publishing Group, 35(8), pp. 725–731. doi: 10.1038/nbt.3893.

Buelow, H. N. *et al.* (2016) "Microbial community responses to increased water and organic matter in the arid soils of the mcmurdo dry valleys, antarctica", *Frontiers in Microbiology*. Frontiers Media S.A., 7(JUL). doi: 10.3389/fmicb.2016.01040.

Cary, S. C. *et al.* (2010) "On the rocks: the microbiology of Antarctic Dry Valley soils", *Nature Reviews Microbiology*. Nature Publishing Group, 8(2), pp. 129–138. doi: 10.1038/nrmicro2281.

Cavicchioli, R. (2015) "Microbial ecology of Antarctic aquatic systems", *Nature Reviews Microbiology*, 13(11), pp. 691–706. doi: 10.1038/nrmicro3549.

Chan, P. P. y Lowe, T. M. (2019) "tRNAscan-SE: Searching for tRNA genes in genomic sequences", en *Methods in Molecular Biology*. doi: 10.1007/978-1-4939-9173-0_1.

Chan, Y. *et al.* (2013) "Functional ecology of an Antarctic Dry Valley", *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 110(22), pp. 8990–8995. doi: 10.1073/pnas.1300643110.

Cowan, D. A. *et al.* (2011) "Hypolithic communities: Important nitrogen sources in Antarctic desert soils", *Environmental Microbiology Reports*, 3(5), pp. 581–586. doi: 10.1111/j.1758-2229.2011.00266.x.

Cowan, D. A. *et al.* (2014) "Microbial ecology and biogeochemistry of continental Antarctic soils", *Frontiers in Microbiology*, 5(APR), pp. 1–10. doi: 10.3389/fmicb.2014.00154.

Eddy, S. R. (1998) "Profile hidden Markov models", *Bioinformatics*, 14(9), pp. 755–763. doi: 10.1093/bioinformatics/14.9.755.

El-Gebali, S. *et al.* (2019) "The Pfam protein families database in 2019", *Nucleic Acids Research*, 47(D1), pp. D427–D432. doi: 10.1093/nar/gky995.

Fountain, A. G. *et al.* (1999) "Physical controls on the Taylor Valley ecosystem, Antarctica", *BioScience*, 49(12), pp. 961–971. doi: 10.1525/bisi.1999.49.12.961.

Fountain, A. G. *et al.* (2010) "Snow in the McMurdo Dry Valleys, Antarctica", *INTERNATIONAL JOURNAL OF CLIMATOLOGY Int. J. Climatol.* Wiley InterScience, 30, pp. 633–642. doi: 10.1002/joc.1933.

Handelsman, J. (2004) "Metagenomics: Application of Genomics to Uncultured Microorganisms", *Microbiology and Molecular Biology Reviews*, 68(4), pp. 669–685. doi: 10.1128/MMBR.68.4.669-685.2004.

Huerta-Cepas, J. *et al.* (2017) "Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper", *Molecular Biology and Evolution*. doi: 10.1093/molbev/msx148.

Huerta-Cepas, J. *et al.* (2019) "EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses", *Nucleic Acids Research*. doi: 10.1093/nar/gky1085.

Hyatt, D. *et al.* (2010) "Prodigal: Prokaryotic gene recognition and translation initiation site identification", *BMC Bioinformatics*. BioMed Central, 11(1), p. 119. doi: 10.1186/1471-2105-11-119.

Jost, L. (2006) "Entropy and diversity", *Oikos*, 113(2), pp. 363–375. doi: 10.1111/j.2006.0030-1299.14714.x.

Jungblut, A. D. *et al.* (2016) "Microbial Mat Communities along an Oxygen Gradient in a Perennially Ice-Covered Antarctic Lake.", *Applied and environmental microbiology*. American Society for Microbiology, 82(2), pp. 620–30. doi: 10.1128/AEM.02699-15.

Kanehisa, M. *et al.* (2016) "KEGG as a reference resource for gene and protein annotation", *Nucleic Acids Research*. doi: 10.1093/nar/gkv1070.

Kang, D. D. *et al.* (2015) "MetaBAT, an efficient tool for accurately reconstructing

- single genomes from complex microbial communities", *PeerJ*. PeerJ Inc., 3, p. e1165. doi: 10.7717/peerj.1165.
- Kim, B. R. *et al.* (2017) "Deciphering diversity indices for a better understanding of microbial communities", *Journal of Microbiology and Biotechnology*, 27(12), pp. 2089–2093. doi: 10.4014/jmb.1709.09027.
- Kohler, T. J. *et al.* (2016) "Nutrient treatments alter microbial mat colonization in two glacial meltwater streams from the McMurdo Dry Valleys, Antarctica", *FEMS Microbiology Ecology*. Editado por D. Wagner. Oxford University Press, 92(4), p. fiw049. doi: 10.1093/femsec/fiw049.
- Komárek, J., Elster, J. y Komárek, O. (2008) "Diversity of the cyanobacterial microflora of the northern part of James Ross Island, NW Weddell Sea, Antarctica", *Polar Biology*. Springer-Verlag, 31(7), pp. 853–865. doi: 10.1007/s00300-008-0424-1.
- Konopka, A. (2009) "What is microbial community ecology", *ISME Journal*, 3(11), pp. 1223–1230. doi: 10.1038/ismej.2009.88.
- Koonin, E. V. (2005) "Orthologs, Paralogs, and Evolutionary Genomics", *Annual review of genetics*, 39(1), pp. 309–338. doi: 10.1146/annurev.genet.39.073003.114725.
- Lane, D. J. *et al.* (1985) "Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses", *Proceedings of the National Academy of Sciences of the United States of America*, 82(20), pp. 6955–6959. doi: 10.1073/pnas.82.20.6955.
- Langmead, B. *et al.* (2009) "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome", *Genome Biology*. doi: 10.1186/gb-2009-10-3-r25.
- Lee, J. Z. *et al.* (2017) "Metagenomics reveals niche partitioning within the phototrophic zone of a microbial mat", *bioRxiv*. Cold Spring Harbor Laboratory, p. 151704. doi: 10.1101/151704.
- Letunic, I. y Bork, P. (2019) "Interactive Tree Of Life (iTOL) v4: recent updates and new developments", *Nucleic Acids Research*, 47(W1), pp. W256–W259. doi: 10.1093/nar/gkz239.
- Li, D. *et al.* (2015) "MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph", *Bioinformatics*. Oxford University Press, 31(10), pp. 1674–1676. doi: 10.1093/bioinformatics/btv033.
- Li, H. *et al.* (2009) "The Sequence Alignment/Map format and SAMtools", *Bioinformatics*. doi: 10.1093/bioinformatics/btp352.
- De Los Ríos, A. *et al.* (2004) "Microstructural Characterization of Cyanobacterial Mats from the McMurdo Ice Shelf, Antarctica", *Applied and Environmental Microbiology*. American Society for Microbiology, 70(1), pp. 569–580. doi: 10.1128/AEM.70.1.569-580.2004.
- Louca, S. *et al.* (2018) "Function and functional redundancy in microbial systems", *Nature Ecology & Evolution*. Nature Publishing Group, p. 1. doi: 10.1038/s41559-018-

0519-1.

Loza, V. *et al.* (2013) "Phenotypic and genotypic characteristics of Phormidium-like cyanobacteria inhabiting microbial mats are correlated with the trophic status of running waters", *European Journal of Phycology*, 48(2), pp. 235–252. doi: 10.1080/09670262.2013.799715.

Magoč, T. y Salzberg, S. L. (2011) "FLASH: Fast length adjustment of short reads to improve genome assemblies", *Bioinformatics*, 27(21), pp. 2957–2963. doi: 10.1093/bioinformatics/btr507.

von Meijenfeldt, F. A. B. *et al.* (2019) "Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT", *Genome Biology*, 20(1), p. 217. doi: 10.1186/s13059-019-1817-x.

Meyer, F. *et al.* (2008) "The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes", *BMC Bioinformatics*. BioMed Central, 9(1), p. 386. doi: 10.1186/1471-2105-9-386.

Michaud, A. B., Šabacká, M. y Priscu, J. C. (2012) "Cyanobacterial diversity across landscape units in a polar desert: Taylor Valley, Antarctica", *FEMS Microbiology Ecology*. Oxford University Press, 82(2), pp. 268–278. doi: 10.1111/j.1574-6941.2012.01297.x.

Niederberger, T. D., Sohm, J. A., Gunderson, T., *et al.* (2015) "Carbon-Fixation Rates and Associated Microbial Communities Residing in Arid and Ephemeral Wet Antarctic Dry Valley Soils", *Frontiers in Microbiology*. Frontiers, 6(DEC), p. 1347. doi: 10.3389/fmicb.2015.01347.

Niederberger, T. D., Sohm, J. A., Gunderson, T. E., *et al.* (2015) "Microbial community composition of transiently wetted Antarctic Dry Valley soils", *Frontiers in Microbiology*. Frontiers, 6(JAN), p. 9. doi: 10.3389/fmicb.2015.00009.

Oksanen, J. *et al.* (2019) "vegan: Community Ecology Package". R package.

Parks, D. H. *et al.* (2015) "CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes.", *Genome research*. Cold Spring Harbor Laboratory Press, 25(7), pp. 1043–55. doi: 10.1101/gr.186072.114.

Peura, S. *et al.* (2015) "Metagenomic insights into strategies of aerobic and anaerobic carbon and nitrogen transformation in boreal lakes", *Scientific Reports*. Nature Publishing Group, 5. doi: 10.1038/srep12102.

Price, M. N., Dehal, P. S. y Arkin, A. P. (2010) "FastTree 2 - Approximately maximum-likelihood trees for large alignments", *PLoS ONE*, 5(3). doi: 10.1371/journal.pone.0009490.

Pruitt, K. D., Tatusova, T. y Maglott, D. R. (2007) "NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins", *Nucleic Acids Research*, 35(SUPPL. 1), pp. D61–D65. doi: 10.1093/nar/gkl842.

R Core Team (2019) "R: A language and environment for statistical computing".

Vienna, Austria., Austria.: R Foundation for Statistical Computing.

Rignot, Eric *et al.* (2002) "Rapid bottom melting widespread near Antarctic Ice Sheet grounding lines.", *Science (New York, N.Y.)*. American Association for the Advancement of Science, 296(5575), pp. 2020–3. doi: 10.1126/science.1070942.

Rinke, C. *et al.* (2013) "Insights into the phylogeny and coding potential of microbial dark matter", *Nature*, 499(7459), pp. 431–437. doi: 10.1038/nature12352.

Rojas, J. L. *et al.* (2009) "Bacterial diversity from benthic mats of Antarctic lakes as a source of new bioactive metabolites", *Marine Genomics*, 2(1), pp. 33–41. doi: 10.1016/j.margen.2009.03.005.

Rozewicki, J. *et al.* (2019) "MAFFT-DASH: integrated protein sequence and structural alignment", *Nucleic Acids Research*. doi: 10.1093/nar/gkz342.

de Scally, S. Z. *et al.* (2016) "Antarctic microbial communities are functionally redundant, adapted and resistant to short term temperature perturbations", *Soil Biology and Biochemistry*, 103, pp. 160–170. doi: 10.1016/j.soilbio.2016.08.013.

Seeman, T. (2019) "barnap 0.9 : rapid ribosomal RNA prediction".

Smith, J. J. *et al.* (2006) "Bacterial diversity in three different antarctic cold desert mineral soils", *Microbial Ecology*, 51(4), pp. 413–421. doi: 10.1007/s00248-006-9022-3.

Stal, L. J. (2012) "Cyanobacterial Mats and Stromatolites", en *Ecology of Cyanobacteria II*. Dordrecht: Springer Netherlands, pp. 65–125. doi: 10.1007/978-94-007-3855-3_4.

Stewart, R. D. *et al.* (2018) "Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen", *Nature Communications*. Nature Publishing Group, 9(1), p. 870. doi: 10.1038/s41467-018-03317-6.

Taton, A. *et al.* (2003) "Cyanobacterial Diversity in Natural and Artificial Microbial Mats of Lake Fryxell (McMurdo Dry Valleys , Antarctica): a Morphological and Molecular Approach", *Society*, pp. 5157–5169. doi: 10.1128/AEM.69.9.5157.

Tiwari, S. *et al.* (2017) "Two-Component Signal Transduction Systems of Pathogenic Bacteria As Targets for Antimicrobial Therapy: An Overview", *Frontiers in Microbiology*, 8. doi: 10.3389/fmicb.2017.01878.

Van Trappen, S. *et al.* (2002) "Diversity of 746 Heterotrophic Bacteria Isolated from Microbial Mats from Ten Antarctic Lakes", *Systematic and Applied Microbiology*. Urban & Fischer, 25(4), pp. 603–610. doi: 10.1078/07232020260517742.

Turner, J., Overland, J. E. y Walsh, J. E. (2007) "An Arctic and antarctic perspective on recent climate change", *International Journal of Climatology*, 27(3), pp. 277–293. doi: 10.1002/joc.1406.

Tytgat, B. *et al.* (2014) "Bacterial Diversity Assessment in Antarctic Terrestrial and Aquatic Microbial Mats: A Comparison between Bidirectional Pyrosequencing and Cultivation", *PLoS ONE*. Editado por S. Bereswill. Public Library of Science, 9(6), p.

e97564. doi: 10.1371/journal.pone.0097564.

Varin, T. *et al.* (2010) "Metagenomic profiling of Arctic microbial mat communities as nutrient scavenging and recycling systems", *Limnology and Oceanography*, 55(5), pp. 1901–1911. doi: 10.4319/lo.2010.55.5.1901.

Varin, T. *et al.* (2012) "Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the high Arctic", *Applied and Environmental Microbiology*. American Society for Microbiology, 78(2), pp. 549–559. doi: 10.1128/AEM.06354-11.

Vollmers, J., Wiegand, S. y Kaster, A.-K. (2017) "Comparing and Evaluating Metagenome Assembly Tools from a Microbiologist's Perspective - Not Only Size Matters!", *PLOS ONE*. Editado por F. Rodriguez-Valera. Public Library of Science, 12(1), p. e0169662. doi: 10.1371/journal.pone.0169662.

van der Walt, A. J. *et al.* (2017) "Assembling metagenomes, one community at a time", *BMC Genomics*. BioMed Central, 18(1), p. 521. doi: 10.1186/s12864-017-3918-9.

Watson, J. D. (2014) *Molecular Biology of the Gene*. Pearson (Always learning).

Wei, S. T. S. *et al.* (2015) "Diverse metabolic and stress-tolerance pathways in chasmoendolithic and soil communities of Miers Valley, McMurdo Dry Valleys, Antarctica", *Polar Biology*. Springer Berlin Heidelberg, 38(4), pp. 433–443. doi: 10.1007/s00300-014-1598-3.

Wei, S. T. S. *et al.* (2016) "Taxonomic and Functional Diversity of Soil and Hypolithic Microbial Communities in Miers Valley, McMurdo Dry Valleys, Antarctica", *Frontiers in Microbiology*. Frontiers, 7(OCT), p. 1642. doi: 10.3389/fmicb.2016.01642.

White, R. A. *et al.* (2015) "Metagenomic analysis reveals that modern microbialites and polar microbial mats have similar taxonomic and functional potential", *Frontiers in Microbiology*. Frontiers, 6(SEP), p. 966. doi: 10.3389/fmicb.2015.00966.

Wickham, H. (2017) "tidyverse: Easily Install and Load the 'Tidyverse'". R package.

Wilkins, D. *et al.* (2013) "Key microbial drivers in Antarctic aquatic environments", *FEMS Microbiology Reviews*. Oxford University Press, 37(3), pp. 303–335. doi: 10.1111/1574-6976.12007.

Woese, C. R. y Fox, G. E. (1977) "Phylogenetic structure of the prokaryotic domain: The primary kingdoms", *Proceedings of the National Academy of Sciences of the United States of America*, 74(11), pp. 5088–5090. doi: 10.1073/pnas.74.11.5088.

Yau, S. *et al.* (2013) "Metagenomic insights into strategies of carbon conservation and unusual sulfur biogeochemistry in a hypersaline Antarctic lake", *The ISME Journal*. Nature Publishing Group, 7(10), pp. 1944–1961. doi: 10.1038/ismej.2013.69.

Zaikova, E. *et al.* (2019) "Antarctic relic microbial mat community revealed by metagenomics and metatranscriptomics", *Frontiers in Ecology and Evolution*, 7(JAN). doi: 10.3389/fevo.2019.00001.

Anexo 1

Lista de los módulos KEGG que representan a las rutas metabólicas relacionadas con el metabolismo energético

Metabolismo energético

Fijación de carbono

- M00165 Ciclo de Calvin
- M00173 Ciclo Arnon-Buchanan
- M00376 3-Hidroxipropionato
- M00375 Hidroxipropionato-hidroxiutilato
- M00374 Dicarboxilato-hidroxiutilato
- M00377 Ruta Wood-Ljungdahl

Metabolismo metano

- M00567 Metanogénesis, CO₂ => metano
- M00357 Metanogénesis, acetato => metano
- M00356 Metanogénesis, metanol => metano
- M00563 Metanogénesis, metil/dimetil/trimetilamina => metano
- M00174 Oxidación metano, metanótrofos, metano => formaldehído
- M00346 Asimilación formaldehído, ruta de la serina
- M00345 Asimilación formaldehído, ruta ribulosa monofosfato
- M00344 Asimilación formaldehído, ruta xilulosa monofosfato

Metabolismo nitrógeno

- M00175 Fijación nitrógeno, nitrógeno => amonio
- M00531 Reducción nitrato asimilativa, nitrato => amonio
- M00530 Reducción nitrato disimilativa, nitrato => amonio
- M00529 Desnitrificación, nitrato => nitrógeno
- M00528 Nitrificación, amonio => nitrito
- M00804 Nitrificación completa, comammox, amonio => nitrito => nitrato

Metabolismo azufre

- M00176 Reducción sulfato asimilativa, sulfato => H₂S
- M00596 Reducción sulfato disimilativa, sulfato => H₂S
- M00595 Oxidación tiosulfato por SOX, tiosulfato => sulfato

Fotosíntesis

- M00161 Fotosistema II
- M00163 Fotosistema I
- M00597 Fotosistema anoxigénico II
- M00598 Fotosistema anoxigénico I

Anexo 2

Lista de modelos ocultos de Markov y sus funciones utilizadas para caracterizar la transformación de carbono y nitrógeno.

ID	Función general
AMO	Nitrificación
AmoC	Nitrificación/oxidación de metano
APS-reductase_C	Sulfato reducción desasimilatoria
APS_kinase	Sulfato reducción asimilatoria
AmoA	Nitrificación
ATP-sulfurylase	Sulfato reducción asimilatoria
Bac_GDH	Mineralización/asimilación nitrógeno
BChI_A	Fotosíntesis bacteriana BVF
CdhC	Ruta Wood-Ljungdahl
CDO_I	Mineralización azufre
CitF	Ciclo reverso ácidos tricarbónicos
Citrate_synt	Ciclo de Krebs
COX1	Fosforilación oxidativa
COX3	Respiración
COXG	CO oxidación
Cytochrom_C	Fosforilación oxidativa
Cytochrom_C552	Nitrato reducción desasimilatoria
Cytochrome_CBB3	Respiración
DmsC	DMSO reducción
Fer4_NifH	Fijación nitrógeno
GDH_N	Mineralización/asimilación nitrógeno
Gln-synt_C	Asimilación nitrógeno
Gln-synt_N	Asimilación nitrógeno
Glu_syn_central	Asimilación nitrógeno

Glu_synthase Asimilación nitrógeno
MCR_alpha Metanogénesis y metanogénesis reversa
MCR_alpha_N Metanogénesis y metanogénesis reversa
MCR_beta Metanogénesis y metanogénesis reversa
MCR_beta_N Metanogénesis y metanogénesis reversa
MCR_gamma Metanogénesis y metanogénesis reversa
MeMO_Hyd_G Oxidación metano
Monooxygenase_B Nitrificación/oxidación de metano
NAD-GH Mineralización/asimilación nitrógeno
NIR_SIR Nitrato/sulfato reducción desasimilatoria
NIR_SIR_ferr Asimilación nitrógeno y azufre
Nitr_red_alph_N Desnitrificación
Nitrate_red_gam Desnitrificación
Nitro_FeMo-Co Fijación nitrógeno
NosL Desnitrificación
PFO_beta_C Fermentación
Photo_RC Fototrofía anoxigénica aeróbica
PRK Ruta pentosas fosfato
RC-P840_PscD Fotosíntesis BVA
RuBisCO_large Ciclo de Calvin
RuBisCO_small Ciclo de Calvin

Anexo 3

Gráfica de conteos sobre la relación entre los metabolismos relacionados con el ciclaje de nutrientes y la anotación taxonómica de *contigs*

