

**Centro de Investigación Científica y de Educación  
Superior de Ensenada, Baja California**



---

**Maestría en Ciencias  
en Ciencias de la vida con  
orientación en Biotecnología Marina**

---

**Estudio de las comunidades microbianas asociadas a  
ambientes geotermales de la falla Wagner, golfo de California**

Tesis

para cubrir parcialmente los requisitos necesarios para obtener el grado de  
Maestro en Ciencias

Presenta:

**Israel Mateo Sánchez Leyva**

Ensenada, Baja California, México

2019

Tesis defendida por

**Israel Mateo Sánchez Leyva**

y aprobada por el siguiente Comité

---

**Dra. María Asunción Lago Lestón**  
Directora de tesis

**Dra. Meritxell Riquelme Pérez**

**Dr. Marcial Leonardo Lizárraga Partida**

**Dr. Thomas Gunter Kretzschmar**



---

**Dr. Clara Elizabeth Galindo Sánchez**  
Coordinadora del Posgrado en Ciencias de la Vida

---

**Dra. Rufina Hernández Martínez**  
Directora de Estudios de Posgrado

*Israel Mateo Sánchez Leyva © 2019*

*Queda prohibida la reproducción parcial o total de esta obra sin el permiso formal y explícito del autor y director de la tesis.*

Resumen de la tesis que presenta **Israel Mateo Sánchez Leyva**, como requisito parcial para la obtención del grado de Maestro en Ciencias de la Vida con orientación en Biotecnología Marina

**Estudio de las comunidades microbianas asociadas a ambientes geotermales de la falla Wagner, golfo de California**

Resumen aprobado por:

---

**Dra. María Asunción Lago Lestón**  
**Directora de tesis**

La falla Wagner es parte del sistema de fallas del golfo de California localizado en la zona norte del golfo. La falla presenta alta actividad termal y de emisión de gases en la frontera este de la cuenca Wagner, lo que la convierte en un sitio de interés para realizar estudios sobre la adaptación microbiana a condiciones extremas. En este trabajo se analizaron muestras de sedimento colectadas durante la campaña oceanográfica Geotermia II realizada de septiembre a octubre del 2015. Se estudiaron muestras tomadas hasta una profundidad de 2.2 metros bajo la superficie del suelo marino y que mostraron gradientes de temperatura a lo largo de la columna sedimentaria, siendo la temperatura máxima registrada de 99 °C en las secciones más profundas. Se analizaron las muestras provenientes de 3 estaciones de muestreo, donde se obtuvieron 2 núcleos por estación, los cuales presentaban diferencias en temperatura entre núcleos. Cada núcleo se seccionó cada 20 centímetros y cada sección fue liofilizada. Para el análisis metagenómico se extrajo el ADN total de 66 muestras independientes, se amplificó la región hipervariable V4 del gen de la subunidad 16S de ARNr y se secuenció utilizando la plataforma MiSeq de Illumina. En total se analizaron un total de 2.2 millones de secuencias, de las cuales se obtuvieron 1,706 OTUs distintos (1,336 bacterianos, 278 de arqueas y 92 sin asignar). Los resultados indicaron que las zonas más calientes presentan una baja diversidad de microorganismos, y que es la profundidad la característica que influye más en el patrón de distribución de la muestra. También se observan diferencias entre las comunidades de los núcleos fríos y los núcleos calientes y a lo largo del perfil de profundidad. Solo las secciones más superficiales, las de los 20 primeros centímetros, son las que mostraron unas comunidades muy parecidas entre todas las estaciones.

Palabras clave: **Sedimentos Geotermales, Metagenómica, ARNr 16S, Extremófilos, Secuenciación masiva.**

Abstract of the thesis presented by **Israel Mateo Sánchez Leyva** as a partial requirement to obtain the Master of Science degree in Life Sciences with orientation in Marine Biotechnology

**Study of the microbial communities associated to geothermal environments of the Wagner fault, Gulf of California**

Abstract approved by:

---

**Dra. María Asunción Lago Lestón**  
Thesis director

The Wagner fault is part of the Gulf of California fault system located in the Northern area. This fault has high thermal and gas seepage activities at the Eastern border of the Wagner basin, which makes it an interesting site for the study of microbial adaptation to extreme conditions. In this work, sediment samples collected during the Geotermia II oceanographic campaign (September-October, 2015) were analyzed. The studied samples were taken up to 2.2 meters below the seafloor and with different temperature gradients along the sedimentary column, with the highest temperature registered at 99 °C at the deepest sections. Samples from three sample stations were analyzed, from which were obtained two cores that presented differences in temperature. Each core was sectioned every 20 centimeters and each section was lyophilized. For metagenomic analysis, total DNA from 66 independent samples was extracted; the hypervariable region V4 was amplified from the 16S rRNA subunit and, the samples were sequenced using the Illumina MiSeq platform. A total of 2.2 million sequences were analyzed from which was obtained 1,706 OTUs (1,336 bacterial; 278 archaeal; and 92 unassigned). Results showed that the warmer regions have lower microorganism diversity and, that depth is crucial for the distribution pattern of the sample. It also was observed differences between microbial communities in the cold and hot core along the depth profile. Only the more superficial sections, i.e., the first 20 centimeters, showed very similar communities between all stations.

Key words: **Geothermal Sediments, Geothermal, Metagenomics, 16S rRNA, Extremophiles, High-throughput Sequencing.**

## Dedicatoria

A mi abuelo **Mateo Sánchez Quintal**.

Porque a pesar de tu ausencia siempre te encuentras presente en cada recuerdo.

## Agradecimientos

Al **Consejo Nacional de Ciencia y Tecnología** por el apoyo económico brindado durante estos dos años a través de programa de becas nacionales (No. 613344 ).

**Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California** por sus programas educativos de calidad y por haberme aceptado como estudiante en su institución.

Al programa de **Posgrado en Ciencias de la Vida** por el apoyo otorgado para la presentación de este trabajo de tesis en el INTERNATIONAL SYMPOSIUM "Extreme Ecosystems & Extremophile Organisms: Biodiversity, Physiology, Biochemistry & Biotechnology." que tuvo lugar en Cuenavaca del 17 al 20 de Septiembre de 2017.

A mi directora, la **Dra. María Asunción Lago Lestón** por haberme aceptado como su estudiante, y por toda la paciencia que ha gastado en mí en este proceso de aprendizaje, Muchas gracias.

A mis Sinodales **Dra. Meritxell Riquelme Pérez, Dr. Marcial Leonardo Lizárraga Partida** y el **Dr. Thomas Gunter Kretzschmar** por sus valiosas contribuciones a lo largo de la etapa experimental y sus valiosos comentarios en la escritura de esta tesis.

A la **Dra. Jennifers Chong Robles** por todo su apoyo para la conclusión de la fase experimental de este trabajo.

Al **Dr. Dante Alberto Magdaleno Moncayo** por todos sus consejos y enseñanzas en los análisis bioinformáticos.

Al **grupo de trabajo del Laboratorio de Metagenómica del CICESE**, por todos sus ánimos en los días malos y su apoyo en la elaboración de esta tesis.

Al **M.C. Luis Miguel Molina Martínez** por su vital ayuda en el procesamiento de las muestras utilizadas en este trabajo.

A mis Padres **María Asunción Leyva Aguilar** y **Mateo Sánchez Ross** por todo su apoyo y motivación para superarme cada día.

A mi **familia** por siempre estar ahí y preocuparse por mí.

A la **B.I. Ana Cecilia Flores Rendón** por todo este tiempo que hemos compartido.

# Tabla de contenido

<b>Resumen en español</b> .....	ii
<b>Resumen en inglés</b> .....	iii
<b>Dedicatoria</b> .....	iv
<b>Agradecimientos</b> .....	v
<b>Lista de figuras</b> .....	viii
<b>Lista de tablas</b> .....	x
<b>Capítulo 1. Introducción</b> .....	1
<b>1.1. Antecedentes</b> .....	1
<b>1.1.1. Sedimentos marinos</b> .....	1
<b>1.1.2. Interacciones biogeoquímicas</b> .....	1
<b>1.1.3. Microorganismos del suelo marino</b> .....	1
<b>1.1.4. Técnicas de secuenciación masiva</b> .....	2
<b>1.1.5. Golfo de California</b> .....	3
<b>1.1.6. Cuenca Wagner</b> .....	5
<b>1.2. Justificación</b> .....	5
<b>1.3. Hipótesis</b> .....	6
<b>1.4. Objetivos</b> .....	6
<b>1.4.1. Objetivo General</b> .....	6
<b>1.4.2. Objetivos Específicos</b> .....	6
<b>Capítulo 2. Metodología</b> .....	7
<b>2.1. Muestreo</b> .....	7
<b>2.2. Extracción de ADN</b> .....	9
<b>2.3. Concentración y purificación del ADN</b> .....	9
<b>2.4. Construcción de las bibliotecas de secuenciación</b> .....	10
<b>2.5. Normalización de los productos de PCR</b> .....	11

2.6. Análisis bioinformático .....	11
<b>Capítulo 3. Resultados.....</b>	<b>14</b>
3.1. Extracción de ADN .....	14
3.2. Amplificación de ADN .....	15
3.3. Preparación de las bibliotecas metagenómicas .....	15
3.4. Control de calidad de las secuencias.....	15
3.5. Descripción de la composición de la comunidad procariota.....	18
3.5.1. Comunidad procariota de la Estación 4 .....	19
3.5.2. Comunidad procariota de la Estación 9 .....	21
3.5.3. Comunidad procariota de la Estación 2 .....	22
3.6. Comparación entre estaciones mediante análisis de coordenadas principales .....	24
3.6.1. Análisis por estación.....	24
3.6.2. Análisis por temperatura .....	25
3.6.3. Análisis por profundidad .....	26
3.7. Análisis de las familias más abundantes .....	27
3.7.1. Principales familias de la Estación 4 .....	27
3.7.2. Principales familias de la Estación 9 .....	29
3.7.3. Principales familias de la Estación 2 .....	32
3.8. Predicción funcional utilizando herramientas bioinformáticas .....	34
<b>Capítulo 4. Discusión .....</b>	<b>37</b>
4.1 Relaciones observadas entre las comunidades .....	38
4.2 Análisis de las principales familias .....	40
<b>Capítulo 5. Conclusiones .....</b>	<b>42</b>
<b>Literatura citada .....</b>	<b>43</b>



## Lista de figuras

- Figura 1** Localización de las cuencas Wagner y Consag, y sus fallas. (Fuente: Prol-Ledesma, 2013).....4
- Figura 2** Localización de las estaciones de muestreo con respecto a la falla Wagner. A la izquierda se indica la zona del golfo de California donde se colectaron los núcleos y a la derecha se puede ver la posición de los puntos de colecta junto a la batimetría de la zona.....7
- Figura 3** Perfiles de temperaturas de las muestras colectadas. A la izquierda se muestra la ubicación de los núcleos colectados y la temperatura superficial del núcleo más caliente por estación. A la derecha se muestran los perfiles de temperatura de los núcleos analizados. En azul se indica la estación 4 (E4), en amarillo la estación 9 (E9) y en rojo la estación 2 (E2) .....8
- Figura 4** Amplicón etiquetado para secuenciación, se muestran los componentes del primer de amplificación, los adaptadores P5 y P7 (33pb), los índices forward y reverse (8pb), la secuencia pad (10pb), la secuencia linker (2pb), los primers gen específicos, modificado de Kozich *et al.* (2010). .....10
- Figura 5** Gráficas de calidad de las secuencias. a) Calidad de las secuencias obtenidas donde se muestran todas las muestras en un gráfico proporcionado por el programa MultiQC, b) distribución de la calidad de las secuencias obtenidas de todas las muestras. ....16
- Figura 6** Control de calidad de las secuencias obtenidas y secuencias utilizadas para el análisis en QIIME. ....16
- Figura 7** Número de OTUs observados por muestra. En la muestra se comparan el número de los OTUs observados por para cada uno de los núcleos por estación. ....17
- Figura 8** a) Curva de rarefacción que muestra el esfuerzo de secuenciación en cada una de las muestras. En el eje de las X se describe el número de secuencias por muestra y el número de OTUs observados. b) Gráfica de correlación entre el número de OTUs observados y OTUs esperados. Los puntos rojos indican el valor Chao 1 que representa los valores estimados o esperados y los puntos azules los valores observados. ....18
- Figura 9** Abundancia relativa de los dominios Arquea (en naranja) y Bacteria (en morado). En verde se muestran aquellos OTUs que no lograron asignarse a ninguno de los dos dominios.....19
- Figura 10** Abundancias relativas de los organismos procariontes, a nivel filo, a lo largo del perfil de profundidad de la estación 4. A izquierda se observa la abundancia relativa del núcleo frío. En el centro se muestra el perfil de temperatura de ambos núcleos. A la derecha, se puede observar la abundancia relativa el núcleo caliente.....20

- Figura 11** Abundancias relativas de los organismos procariontes, a nivel filo, a lo largo del perfil de profundidad de la estación 9. A izquierda se observa la abundancia relativa del núcleo frío. En el centro se muestra el perfil de temperatura de ambos núcleos. A la derecha, se puede observar la abundancia relativa el núcleo caliente.....22
- Figura 12** Abundancias relativas de los organismos procariontes, a nivel filo, a lo largo del perfil de profundidad de la estación 2. A izquierda se observa la abundancia relativa del núcleo frío. En el centro se muestra el perfil de temperatura de ambos núcleos, con línea azul el núcleo frío y con línea roja el núcleo caliente. A la derecha, se puede observar la abundancia relativa el núcleo caliente.....24
- Figura 13** Análisis de coordenadas principales. A la izquierda se puede ver el análisis no ponderado. A la izquierda se puede ver el análisis ponderado. Los diferentes colores indican la estación de la que proviene cada núcleo. ....25
- Figura 14** Análisis de coordenadas principales. A la izquierda se puede ver el análisis no ponderado, A la derecha se puede ver el análisis ponderado. Las muestras se agruparon por intervalos de 20 °C de temperatura. ....26
- Figura 15** Análisis de coordenadas principales indicando las profundidades en la columna sedimentaria. A la izquierda se puede ver el análisis no ponderado y a la derecha el análisis ponderado. ....27
- Figura 16** Gráfica de líneas que muestra el perfil de profundidad de las familias más abundantes de los núcleos de la estación 4. ....29
- Figura 17.** Gráfica de líneas que muestra el perfil de profundidad de las familias más abundantes de los núcleos de la estación 9. A la izquierda se muestra el perfil del núcleo frío y a la derecha el del núcleo caliente. ....31
- Figura 18** Gráfica de líneas que muestra el perfil de profundidad de las familias más abundantes de los núcleos de la estación 2. A la izquierda se muestra el perfil del núcleo frío y a la derecha el del núcleo caliente. ....34
- Figura 19** Heatmap de las funciones predichas por el programa PICRUSt .....35
- Figura 20.** Heatmap de las funciones más abundantes predichas por Tax4Fun .....36

## Lista de tablas

<b>Tabla 1.</b> Localización geográfica de los puntos de donde se recolectaron los núcleos analizados y sus temperaturas promedio. F (núcleo frío); C (núcleo caliente).....	8
<b>Tabla 2</b> Mezcla de reacción para PCR .....	10
<b>Tabla 3</b> Concentración de ADN (ng/ $\mu$ L) en las muestras de sedimento. En amarillo se resaltan las extracciones que tuvieron dificultad para ser extraídas y debieron hacerse varias extracciones y varias amplificaciones de la misma muestra. En verde más claro la que presentó menor concentración.....	14

# Capítulo 1. Introducción

---

## 1.1. Antecedentes

### 1.1.1. Sedimentos marinos

Los sedimentos marinos son un hábitat que comprende  $4.5 \times 10^{17} \text{ m}^3$  en todo el mundo representando, en volumen, el segundo hábitat más vasto del océano, después de la corteza oceánica (Lever M., 2013). Presentan un amplio rango de temperaturas que varía desde zonas muy frías (0-3 °C) a zonas muy calientes (hasta 400 °C) cuando están cerca de áreas geo e hidrotermales y en las cuales se ha observado la presencia de microorganismos capaces de tolerar tales temperaturas (Orcutt *et al.*, 2011). Los sedimentos pueden clasificarse de dos maneras: de acuerdo a su localización en relación al nivel del mar y de acuerdo a la profundidad de la columna sedimentaria. Según su localización en la columna de agua existen dos clasificaciones: los sedimentos profundos y los sedimentos someros. Esta clasificación se debe a la capacidad de penetración de la luz en el agua, así los más someros se encuentran en la zona fótica (generalmente <200 metros bajo la superficie del mar (mbsm)), mientras que los profundos se encuentran en la zona afótica del océano, sin embargo, esto varía dependiendo de las características de la penetración lumínica relacionada con la turbidez en cada región (D'Hondt *et al.*, 2004). La segunda clasificación está en relación a la profundidad de la columna sedimentaria, ésta puede variar desde centímetros hasta kilómetros dependiendo de la tasa de sedimentación característica de cada región, y muchos de los autores los clasifican en superficiales o subsuperficiales. Los primeros se encuentran a menos de 1.5 metros bajo el fondo marino (mbfm) y generalmente presentan una mayor productividad biológica y reciben una mayor aportación de materia orgánica, mientras que los sedimentos subsuperficiales se encuentran a profundidades mayores, generalmente son anóxicos, y se ha observado que presentan una gran abundancia microbiana, así como alta diversidad de actividades metabólicas (Orcutt *et al.*, 2011).

Las comunidades microbianas en los sedimentos subsuperficiales contienen taxones que habitan un amplio espectro de ambientes redox, sin embargo algunos taxones parecen exhibir una preferencia por ambientes subsuperficiales específicos, por lo cual se cree que podría existir una estratificación de microambientes a lo largo de la columna sedimentaria (D'Hondt *et al.*, 2004).

### **1.1.2. Interacciones biogeoquímicas**

Los microorganismos son los mediadores de los ciclos biogeoquímicos del carbono, azufre, hierro y nitrógeno, entre otros. Su actividad metabólica se lleva a cabo por el acoplamiento de reacciones de óxido-reducción la cual depende de la especiación de donadores y aceptores de electrones. Los donadores de electrones más comunes son la materia orgánica, el hidrógeno, el metano, los compuestos reducidos del azufre, el hierro, el manganeso reducido y el amonio. Los aceptores de electrones más comunes en estos ambientes son el oxígeno, el nitrato, el nitrito, los óxidos de manganeso y el hierro, también los compuestos oxidados del azufre, el azufre elemental, el tiosulfato, el dióxido de carbono y los compuestos orgánicos halogenados (Orcutt *et al.*, 2011). En los sedimentos la disponibilidad y concentración de los nutrientes depende de procesos diagenéticos de difusión y advección, que pueden ser mediados por factores bióticos o abióticos, como es la bioturbación, la bioirrigación, la compactación o la pirolisis, estos provocan una zonación espacial y temporal de los sustratos (Boudreau, 1997). En sedimentos marinos esta zonación está muy bien caracterizada, y se ha descrito una cascada de uso preferente de donadores de electrones en un gradiente de tasas metabólicas. El intercambio químico se da en su mayoría por difusión molecular, sin embargo este se puede ver afectado por procesos advectivos como la bioturbación y bioirrigación producidos por los animales en la superficie y por la migración de fluidos, como los lodos fluidizados y los gases, procedentes de reservorios profundos producidos por fallas y fracturas en el sedimento (Boudreau, 1997).

### **1.1.3. Microorganismos del suelo marino**

La diversidad de microorganismos que habitan ciertos estratos del subsuelo marino está condicionada por varios factores, entre ellos la disponibilidad de nutrientes. La concentración de estos nutrientes en el suelo puede variar en pocos centímetros y a su vez, al ser consumidos por los microorganismos circundantes, desencadena un gradiente de actividad microbiana a medida que son metabolizados. Los microorganismos poseen diversos mecanismos de obtención de energía a partir de los sustratos que estén a su disposición. Los microorganismos fotosintéticos pueden aprovechar la luz para catalizar diversos procesos celulares y generar polisacáridos para el almacenamiento de esta energía, como es el caso de las micro y macroalgas, bacterias verde-azules y púrpura. Por otra parte, los organismos quimiosintéticos aprovechan la energía que se encuentra en los compuestos químicos como el sulfato y

el metano, y algunos de ellos son capaces de utilizar el ácido sulfhídrico que emana de ventilas hidrotermales en el fondo marino, y otros pueden metabolizar el hidrógeno proveniente de áreas geotermales.

Además de la disponibilidad de los componentes químicos presentes en los sedimentos, la temperatura también es un factor determinante en la estructura de las comunidades de microorganismos. Los microorganismos que son capaces de vivir a temperaturas elevadas se conocen como termófilos. Wagner and Wiegel (2008), los clasificaron en varias categorías según la temperatura del área geotérmica natural en donde se encuentran, como las ventilas hidrotermales del mar profundo y fuentes hidrotermales. Estas categorías fueron: termófilos, cuando presentan un crecimiento óptimo a 50 °C o más, termófilos extremos (con un crecimiento preferente entre los 65 y 79 °C) y los hipertermófilos (cuando son capaces de crecer por encima de 80 °C).

El estudio de organismos adaptados a estos ambientes tan duros nos da información sobre su flexibilidad metabólica (Badhai *et al.*, 2015) y convierte a los microorganismos termófilos y a sus proteínas termoestables en interesantes desde el punto de vista biotecnológico e industrial.

#### **1.1.4. Técnicas de secuenciación masiva**

Existe una diversidad de técnicas para el estudio de las comunidades microbianas. Tradicionalmente se han venido utilizando técnicas de cultivo, donde las muestras provenientes de los diversos ambientes se enriquecían en medios que contienen los nutrientes necesarios para el desarrollo de los microorganismos. Sin embargo, solo un pequeño porcentaje de los microorganismos son cultivables en condiciones de laboratorio. Los recientes avances en técnicas independientes de cultivo, como la secuenciación masiva, ha proporcionado información relevante en torno a la biodiversidad y las interacciones ecológicas de los microorganismos en diferentes nichos, así como de su participación en los ciclos biogeoquímicos a escala global (Anantharaman *et al.*, 2016). Gracias al estudio de los ácidos nucleicos se han desvelado múltiples aspectos de la vida microbiana, desde su taxonomía mediante la secuenciación de genes del ARN ribosomal (ARNr), su metabolismo potencial mediante técnicas de secuenciación fragmentada, e incluso determinar que genes se están expresando en determinado momento mediante la secuenciación del ARN.

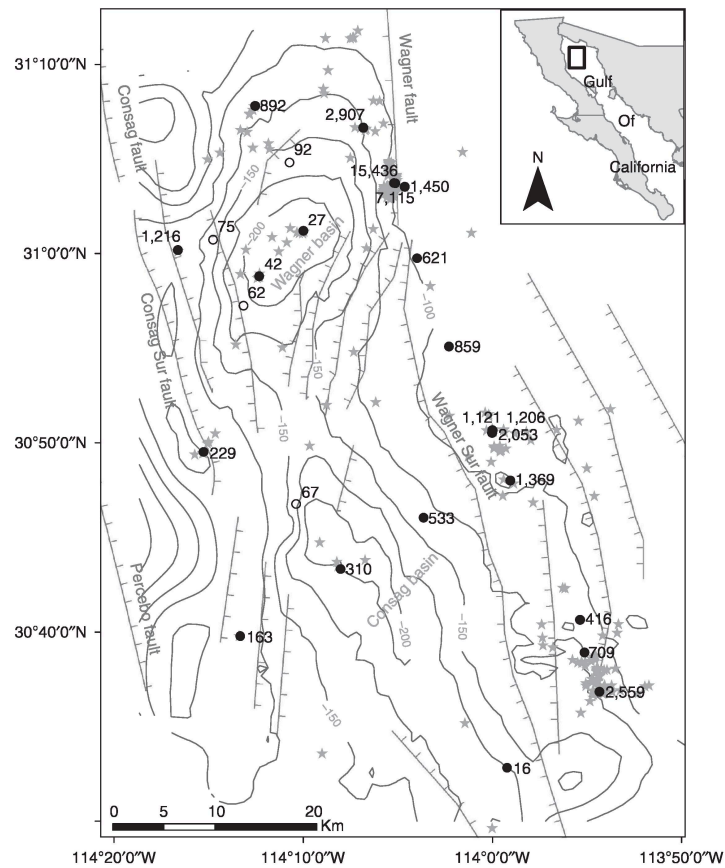
El “*metabarcoding*” o secuenciación del gen del ARNr es una de las técnicas más eficientes para determinar cuáles son las poblaciones microbianas que existen en determinado ambiente. Esta técnica

consiste en la extracción del ADN total presente en una muestra, seguido por la amplificación del gen o región de interés, y su posterior secuenciación. Las regiones de interés más usadas y que son taxonómicamente informativas son: la región de la subunidad 16S del gen del ARNr, en el caso de bacterias y arqueas, la 18S en el caso de eucariotas y la región ITS, la más usada para el estudio de hongos. Hoy en día el desarrollo de las tecnologías de secuenciación nos permiten obtener hasta 400 Gigabases por corrida (Illumina, 2018), lecturas de hasta 10,000 pares de bases por secuencia (PacBio, 2018) con una alta calidad de las secuencias obtenidas ( $Q30 > 85\%$ , 1 error cada 1000 bases), dependiendo del tipo de tecnología empleada. El uso de estas tecnologías se puede optimizar para obtener un mejor rendimiento mediante el uso de secuenciación por extremos pareados, donde un extremo de la secuencia es leído en un primer ciclo y posteriormente se lee el otro extremo en un segundo ciclo (Caporaso, 2012). También se ha utilizado la técnica de índices pareados para incrementar el número de muestras que pueden incorporarse en una sola corrida de secuenciación, esto se refiere al uso de identificadores únicos en ambos extremos de la secuencia por muestra, con solo tan solo 22 índices diferentes se pueden incorporar 117 muestras diferentes en una sola corrida de secuenciación (Kozich, 2013). Sin embargo, existen diferentes retos a superar cuando se utilizan este tipo de tecnologías, uno de ellos es la formación de quimeras, introducidas durante la etapa de amplificación, son aquellas secuencias que pertenecen a dos unidades taxonómicas operacionales (OTUs, por sus siglas en inglés) diferentes y solo pueden ser detectados mediante el uso de herramientas bioinformáticas. Durante la etapa de secuenciación se ha detectado que en ocasiones la diversidad en las muestras estudiadas no es suficiente para cubrir la celda de secuenciación, lo que ocasiona que la secuenciación no pueda ser completada debido a errores en la lectura de la emisión lumínica, para esto se ha incorporado el uso de bibliotecas sintéticas, como una comunidad simulada o el uso de bibliotecas generadas a partir del genoma del virus PhiX (Illumina, 2018).

### **1.1.5. Golfo de California**

El golfo de California (GdeC), está localizado entre la península de California y los estados de Sonora y Sinaloa, México. El GdeC se formó como resultado del desplazamiento de placas tectónicas, dando como resultado la formación de cuencas submarinas del “tipo rift”, las cuales están unidas por fallas de transformación largas y centros de dispersión cortos, hace aproximadamente 6 Ma (Lonsdale, 1989; Aragón-Arreola *et al.*, 2005). A lo largo del GdeC existen diferentes elementos tectónicos que presentan actividad hidrotermal en torno a las diferentes cuencas. Entre ellas se encuentran las cuencas de Guaymas y Yaqui, en el Sur y las cuencas Delfín, Wagner y Consag en el norte (Lonsdale *et al.* 1989; Persuad *et al.* 2000).

Dentro del GdeC la cuenca más estudiada en términos de diversidad microbiana es la cuenca de Guaymas. El estudio de los microorganismos presentes en esta cuenca ha sido sobre todo mediante el uso de técnicas microbiológicas y moleculares. Gracias al análisis de las secuencias de los genes de las subunidades 16S y 18S del ARN ribosomal (Karl *et al.*, 1988; Edgecomb *et al.*, 2002; Teske *et al.*, 2002), se han encontrado una gran variedad de organismos dentro de los tres dominios de la vida, incluyendo arqueas, metanógenas hipertermófilos reductoras de azufre del género *Archaeoglobus*, *Methanopyrus* y *Methanococcus*, algunas bacterias anaerobias reductoras de azufre con la capacidad de degradar alcanos, compuestos aromáticos y ácidos grasos, algunas bacterias aerobias degradadoras de hidrocarburos, especialmente ácidos carboxílicos aromáticos además de algunos simbioses de bivalvos oxidantes de azufre, y eucariotas especialistas de ambientes hidrotermales que carecen de peroxisomas y mitocondrias (Karl *et al.*, 1988; Bazylinski *et al.*, 1989; Edgecomb *et al.*, 2002; Teske *et al.*, 2002)



**Figura 1** Localización de las cuencas Wagner y Consag, y sus fallas. (Fuente: Prol-Ledesma, 2013)



### 1.1.6. Cuenca Wagner

Las cuencas Wagner y Consag se han estudiado por su posible aprovechamiento en la producción de energía geotérmica. En 1986 se realizaron estudios de geomorfología, oceanografía, geofísica, geología, geoquímica y sedimentología en la cuenca Wagner, donde se observaron montículos de lodo, descargas de gases (butano) e hidrocarburos y altas temperaturas en la columna de agua (hasta 200 °C a los 180 m). Además, se detectó azufre, hierro y níquel en las muestras de sedimento (Grijalva, 1986). En investigaciones más recientes se han detectado flujos de calor entre los 16 y 15,436 MWm<sup>2</sup> (promedio 1,875 MWm<sup>2</sup>), lo que sugiere, aunando la intensa actividad sísmica de la zona, la presencia de sistemas geotermales (Prol-Ledesma, 2013). Se ha reportado la presencia de emanaciones de CH<sub>4</sub> y CO<sub>2</sub>, que a su vez provocan acidificación en las cercanías de las emanaciones (pH de 7.55 – 7.85, 8.29 superficial) de 20 a 50 m del fondo submarino en un área aproximada de 2,000 km<sup>2</sup>. Cabe destacar la presencia de hidrocarburos aromáticos policíclicos (PAHs, por sus siglas en inglés) en estos sitios, así como presencia de compuestos que se han relacionado con la presencia de microorganismos pertenecientes al filo Crenarcheota (Prol-Ledesma, 2013). Un estudio presentado en una conferencia sobre las comunidades de arqueas en diferentes puntos de la cuenca Wagner realizado por Perez-Villalatoro *et al.* (2017), identificó que las clases *Methanobacteria* y *Thermoprotei* dominaban en estos ambientes (20% y 24%), además de la presencia de secuencias que correspondían al grupo *Lokiarchaeota*, el cual es un grupo ancestral que está relacionado filogenéticamente con las eucariotas y que se ha descrito como un eslabón importante para el estudio de los mecanismos por los cuales se llevó a cabo la eukaryogenesis (Spang *et al.*, 2015).

## 1.2. Justificación

La abundancia y diversidad microbiana procariota que habita en los sedimentos geotermales de la cuenca Wagner ha sido muy poco estudiada. Estudiar los microorganismos que están adaptados a estos ambientes tan extremos implica la oportunidad de descubrir nuevas bacterias y arqueas que posean nuevas enzimas o metabolitos de interés biotecnológico y, al mismo tiempo, poder entender los mecanismos para la adaptación a condiciones tan extremas para la vida.

### **1.3. Hipótesis**

La abundancia y diversidad microbiana procariota en los sedimentos geotermales varía en función de la profundidad y de los gradientes de temperatura que se presentan a lo largo de la columna sedimentaria .

### **1.4. Objetivos**

#### **1.4.1. Objetivo General**

Caracterizar las comunidades microbianas procariontes de los sedimentos procedentes de áreas geotermales de la falla Wagner, a diferentes profundidades de la columna sedimentaria y con diferentes gradientes de temperatura, mediante técnicas de secuenciación masiva.

#### **1.4.2. Objetivos Específicos**

Identificar las comunidades microbianas procariontes principales de las estaciones de muestreo, mediante secuenciación masiva de la región 16S del gen del ARNr.

Comparar las comunidades microbianas procariontes entre las diferentes estaciones de muestreo y a lo largo de la columna sedimentaria.

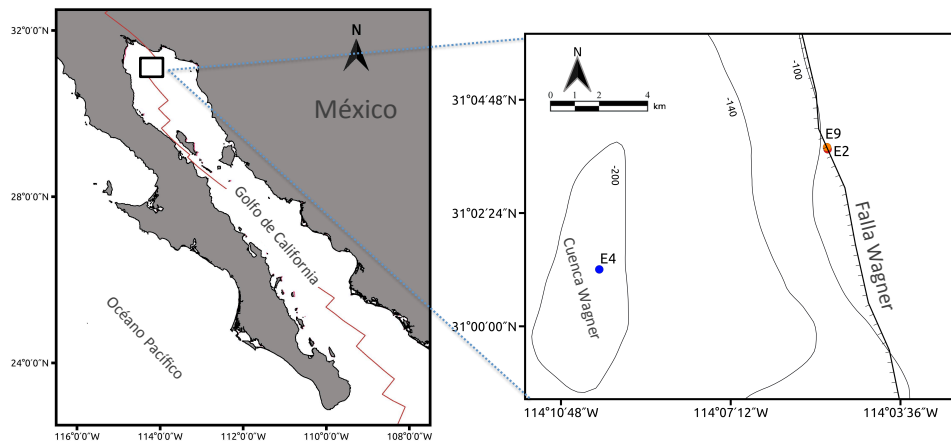
Realizar una inferencia funcional de las comunidades microbianas procariontes a partir de las secuencias de la región hipervariable V4 de la subunidad 16S del ARNr mediante el uso de herramientas bioinformáticas

## Capítulo 2. Metodología

### 2.1. Muestreo

Las muestras de sedimento se obtuvieron durante el crucero Geotermia II realizado del 24 de septiembre al 8 de octubre del 2015, a bordo del buque oceanográfico Alpha Helix. Las muestras se colectaron utilizando nucleadores de gravedad en un total de 13 estaciones de muestreo con una profundidad de penetración variable. Una vez recuperados los núcleos, estos se seccionaron y se congelaron a  $-80\text{ }^{\circ}\text{C}$ . Los núcleos colectados se mantuvieron congelados durante su transporte y posteriormente en el laboratorio cortaron en secciones de 20 cm que posteriormente fueron liofilizadas. Una vez liofilizadas, las muestras se preservaron en una cámara libre de humedad hasta su procesamiento.

Para este trabajo se eligieron muestras de tres estaciones: dos estaciones que presentaban altas temperaturas entre  $35$  y  $50\text{ }^{\circ}\text{C}$  en la sección superficial y hasta  $92$  y  $99\text{ }^{\circ}\text{C}$  en la parte más profunda del núcleo y que se encontraban en la frontera entre la cuenca Wagner y la falla Wagner. También se seleccionó una estación control, cuya temperatura fluctuaba entre los  $14$  y  $16\text{ }^{\circ}\text{C}$  a lo largo de toda la columna sedimentaria, y que estaba ubicada en el centro de la cuenca Wagner. La posición y batimetría de los núcleos de las estaciones utilizadas se muestran en la Figura 2 y sus coordenadas geográficas y temperaturas medias se indican en la Tabla 1.

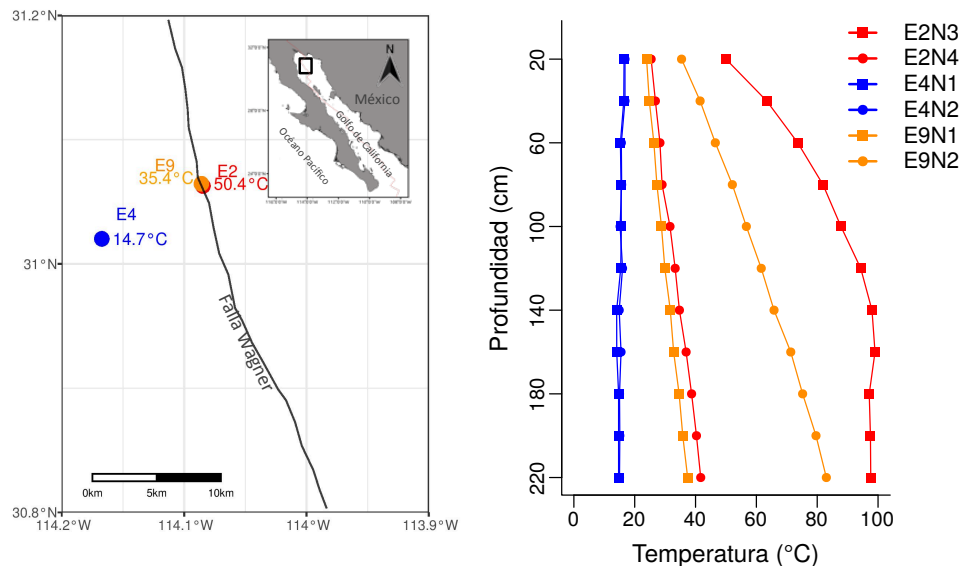


**Figura 2** Localización de las estaciones de muestreo con respecto a la falla Wagner. A la izquierda se indica la zona del golfo de California donde se colectaron los núcleos y a la derecha se puede ver la posición de los puntos de colecta junto a la batimetría de la zona.

**Tabla 1.** Localización geográfica de los puntos de donde se recolectaron los núcleos analizados y sus temperaturas promedio. F (núcleo frío); C (núcleo caliente)

Núcleo	Coordenadas		Temperatura (°C)		
	Longitud	Latitud	Máxima	Mínima	Promedio
E4N1 (F)	31° 01.2105' N	114° 10.041' W	16.5	14.7	15.1
E4N2 (F)	31° 01.2007' N	114° 09.9344' W	16.8	14.9	15.6
E9N1 (F)	31° 03.724' N	114° 05.159' W	42.1	23.9	32.5
E9N2 (C)	31° 03.868' N	114° 05.155' W	92.1	35.4	66.93
E2N4 (F)	31° 03.742' N	114° 05.204' W	46.7	25.4	35.8
E2N3 (C)	31° 03.7627' N	114° 05.086' W	99	50	94.8

Por cada estación se utilizaron dos núcleos que presentaban diferencias en sus temperaturas, uno de ellos presentaba una temperatura más alta, desde 50 a 99 °C en la estación 2 (E2N3) y de 35 a 83 °C en la estación 9 (E9N2) (núcleo “caliente”) y otro mostraba una temperatura más baja (núcleo “frío”), desde 25 a 42 °C en estación 2 (E2N4) y de 24 a 37 °C en la estación (E9N1). En la estación control, estación 4 (E4), ambos núcleos presentaban la misma temperatura (ambos eran núcleos fríos) . En todos los núcleos utilizados la profundidad analizada en la columna sedimentaria fue hasta los 2.2 metros. La posición, batimetría y temperaturas de los núcleos de las estaciones utilizadas se muestran en la Figura 2. Asimismo, en la Tabla 1 se indican los puntos georeferenciados de donde se extrajeron los núcleos utilizados en esta tesis.



**Figura 3** Perfiles de temperaturas de las muestras colectadas. A la izquierda se muestra la ubicación de los núcleos colectados y la temperatura superficial del núcleo más caliente por estación. A la derecha se muestran los perfiles de temperatura de los núcleos analizados. En azul se indica la estación 4 (E4), en amarillo la estación 9 (E9) y en rojo la estación 2 (E2)

## 2.2. Extracción de ADN

Se extrajo el ADN total a partir de 0.15 g de sedimento liofilizado por extracción, utilizando el kit de extracción PowerSoil™ (QIAGEN®) siguiendo el protocolo modificado que indicaba el fabricante. Esta modificación consistía en añadir 200 µL de fenol/cloroformo/alcohol isoamílico (25:24:1) en la etapa de lisis. También se aumentó el tiempo de lisis a 10 min, se añadió alcohol etílico absoluto en la etapa de unión a la columna y se añadió un lavado adicional con alcohol etílico absoluto antes del lavado final con solución C6. La elución final se separó en dos etapas una de 50 µL y una segunda de 30 µL, teniendo un volumen de elución final de 80 µL.

La concentración de ADN se cuantificó mediante espectrofotometría utilizando un NanoDrop™ Lite (Thermo Scientific), a una longitud de onda de 260 nm y se midió la pureza mediante el índice 260nm/280nm. Se repitieron las mediciones en NanoDrop™ 2000 (Thermo Scientific), donde se verificó el espectro de absorción desde 220 nm hasta 350 nm para descartar la presencia de contaminantes. Se utilizaron 2 µL de la extracción de ADN por cada medición. Se verificó la calidad y la cantidad de las extracciones mediante electroforesis en gel de agarosa al 0.8%, utilizando 10 µL de la extracción de ADN teñido con GelRed®, se añadió una escalera HyperLadder™, con un tamaño de fragmentos que va desde los 10 000 pares de bases hasta 200 pares de bases, en uno de los pozos vacíos para poder discernir el tamaño de nuestros fragmentos de ácidos nucleicos (Bioline), la electroforesis se llevó a cabo a 80V por 40 min, permitiendo que las bandas recorrieran aproximadamente 3 cm del pozo.

## 2.3. Concentración y purificación del ADN

Las muestras más profundas (100-220 cm) del núcleo caliente de la estación 2 (E2N3) mostraron un bajo rendimiento en la extracción de ácidos nucleicos, lo que dificultaba su posterior amplificación, por lo cual se procedió a realizar extracciones por triplicado, estas extracciones se juntaron y posteriormente se concentraron utilizando el kit de limpieza y purificación Genomic DNA clean & concentrator™ (Zymo Research Corporation), siguiendo las recomendaciones del fabricante.

## 2.4. Construcción de las bibliotecas de secuenciación

Para la construcción de las bibliotecas se utilizó un enfoque de secuenciación de extremos pareados (*paired-end*) y de doble indexado propuesto por Kozich *et al.* (2013), en el cual la amplificación y la ligación de los adaptadores se realizan en una única reacción de PCR, ya que los *primers* incluyen los adaptadores para la plataforma de secuenciación, los índices, las secuencias PAD-linker, y los *primers* específicos para el gen de la región V4 de la subunidad 16S del ARNr como se muestra en la Figura 4.



**Figura 4** Amplicón etiquetado para secuenciación, se muestran los componentes del primer de amplificación, los adaptadores P5 y P7 (33pb), los índices forward y reverse (8pb), la secuencia pad (10pb), la secuencia linker (2pb), los primers gen específicos, modificado de Kozich *et al.* (2010).

Se amplificaron 20 ng de ADN utilizando una polimerasa MyTaq<sup>TM</sup> (Bioline) y los *primers* gen específicos propuestos por Caporaso *et al.* (2011) para la región V4: 515F (5' GTGCCAGCMGCCGCGGTAA 3') y 806R (5' GGACTACHVGGGTWTCTAAT 3') utilizando una mezcla de reacción en un volumen final de reacción de 20  $\mu$ L, como se muestra en la Tabla 2.

**Tabla 2** Mezcla de reacción para PCR

### Reacción de PCR

Buffer	4 $\mu$ L
MyTaq <sup>TM</sup> (Bioline)	0.125 $\mu$ L
Primers (515F y 806R)	1.6 $\mu$ L
Templado (5ng/ $\mu$ L)	2 $\mu$ L
H <sub>2</sub> O	12.275 $\mu$ L
Total	20 $\mu$ L

El programa para la amplificación constó de 3 etapas: una etapa de desnaturalización inicial a 95 °C por 2 min, seguido de una etapa de 35 ciclos que consistía en un paso de desnaturalización a 95 °C por 20 seg, otro de alineamiento a 55 °C por 15 seg y uno de extensión a 72 °C por 5 min, y una última etapa de extensión final a 72 °C por 7 min. Los productos de la reacción se mantuvieron a 4 °C y se realizó una electroforesis en gel de agarosa al 1% para observar si la amplificación se realizó exitosamente. En el caso de ser exitosa se obtiene una banda de aproximadamente 400 pb.

Las muestras más profundas (100-220 cm) del núcleo caliente de la estación 2 mostraron un bajo rendimiento tanto en la extracción de ADN como en la amplificación, incluso cuando se aumentó a 35 ciclos el programa de amplificación. Para intentar obtener suficiente cantidad de producto, estas muestras se amplificaron por sextuplicado y se concentraron estos productos de amplificación utilizando el *kit QIAquick PCR Purification* (QIAGEN®) siguiendo las recomendaciones del fabricante.

## 2.5. Normalización de los productos de PCR

Debido a que los productos de la amplificación se encontraban en diferentes concentraciones se realizó una normalización de los amplicones con la finalidad de llevarlos a una misma concentración, utilizando el kit de normalización en placa SequalPrep (Invitrogen). La concentración final de la biblioteca se cuantificó mediante un ensayo de fluorometría utilizando un fluorómetro Qubit™ (Thermo Fisher Scientific) y el *kit labeling detection* (Thermo Fisher Scientific) siguiendo el protocolo del fabricante.

Se utilizó una biblioteca control generada a partir del virus PhiX, la cual se mezcló con la biblioteca generada a partir de las muestras de sedimento siguiendo el protocolo recomendado por el fabricante. La secuenciación de las bibliotecas se realizó en la plataforma MiSeq de Illumina que se tiene en CICESE.

## 2.6. Análisis bioinformático

Después de la secuenciación se revisaron los índices de calidad proporcionados por el equipo, y se determinó la calidad de las secuencias utilizando la herramienta FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Las secuencias obtenidas se analizaron utilizando un *pipeline* basado en el uso de la suite QIIME (Caporaso, 2010) (<http://qiime.org/>), que realiza el filtrado de las secuencias con baja calidad (USEARCH, <https://www.drive5.com/usearch/>), el

agrupamiento de las secuencias *de novo* (UCLUST, [https://drive5.com/usearch/manual/uclust\\_algo.html](https://drive5.com/usearch/manual/uclust_algo.html)), la selección de OTUs, la asignación taxonómica utilizando la base de datos SILVA v.128 (Quast, 2013), la reconstrucción filogenética, el análisis de diversidad y la generación de gráficos visuales.

El análisis consistió, de manera más detallada, en un preprocesamiento utilizando el programa USEARCH (Edgar, 2010) y su procesamiento en la suite QIIME. En este pipeline se unieron los extremos de cada una de las secuencias pares (*Forward* y *Reverse*) utilizando el comando *-mergепairs*, seguido del filtrado de calidad de aquellas secuencias que obtuvieron un *quality score* menor a 30 (es decir 1 error cada 1000 bases) utilizando el comando *-fastq\_filter* y a continuación se realizó el filtrado de las secuencias quiméricas utilizando el comando *-uchime*.

Después del preprocesamiento, las secuencias se agruparon utilizando el programa UCLUST (Edgar, 2010) utilizando dos aproximaciones: una por referencia donde las secuencias se agrupan utilizando secuencias de una base de datos de referencia como el centro del grupo; la otra por agrupación *de novo*, que no utiliza secuencias de referencia, sino que construye los grupos comparando todas las secuencias entre si y agrupando las que sean más similares.

La selección de OTUs o asignación taxonómica se realizó utilizando una base de datos de referencia para ambos casos. Para la generación de la tabla de OTUs, los análisis de diversidad y la generación de gráficos visuales se utilizan comandos incluidos también en la suite QIIME (<http://qiime.org/>). Para obtener las unidades taxonómicas operacionales (OTUs, por sus siglas en inglés) se compararon las secuencias obtenidas contra la base de datos SILVA128 16S, mediante un algoritmo basado en la formación de clusters (UCLUST) cuyo centroide es una secuencia de referencia (*closed reference*) o bien una secuencia al azar (*de novo*) y el radio el porcentaje de identidad, en este caso 97%. Posteriormente se construyó una tabla a partir de los OTUs encontrados en las secuencias de cada muestra. Se filtraron los OTUs etiquetados como cloroplastos y mitocondrias, además de aquellos OTUs que solo aparecieron una sola vez "*singletons*". Se calculó la diversidad alfa y beta de las muestras, y posteriormente se elaboraron las gráficas de abundancia relativa, rarefacción y coordenadas principales.

Para el cálculo de la diversidad beta se utilizó la métrica UniFrac. UniFrac nos permite para calcular las diferencias cualitativas (análisis no ponderado) y cuantitativas (análisis ponderado) entre las muestras analizadas, en base a sus relaciones filogenéticas (Lozupone *et al.*, 2007). Para ello, se generó una matriz basada en las distancias filogenéticas y los resultados obtenidos se visualizaron mediante un análisis de



coordenadas principales (PCoA) en donde se eligieron las coordenadas que tenían un mayor peso en la distribución de las muestras.

Adicionalmente se realizó una inferencia funcional de las secuencias mediante el uso de PICRUSt (<http://picrust.github.io/picrust/>). Este es un software que utiliza un algoritmo de modelación evolutiva para la predicción de metagenomas a partir de secuencias de 16S y la comparación contra genomas de referencia basándose en la base de datos GreenGenes (Langille., 2013, <http://greengenes.lbl.gov/>). También se utilizó la paquetería Tax4Fun para obtener perfiles funcionales de las comunidades microbianas encontradas en las muestras de sedimento. Este programa realiza una transformación de los perfiles taxonómicos obtenidos a partir de la base de datos SILVA v.123 (Quast C., 2013) a perfiles funcionales a partir de la base de datos KEGG (<https://www.genome.jp/kegg/>), mediante el uso de matrices de asociación previamente calculadas (Aßhauer K.P. *et al.*, 2015).

Brevemente, el programa PICRUSt permite realizar una reconstrucción ancestral de las actividades microbianas mediante la comparación de los OTUs observados en la secuenciación del 16S y el OTU secuenciado más cercano con base a relaciones del árbol filogenético construido a partir de la base de datos de GreenGenes. Tax4Fun, en cambio, se basa en la asignación de funciones hipotéticas, mediante la similitud de las secuencias, para encontrar los organismos ya secuenciados más cercanos a los obtenidos mediante la secuenciación de la región 16S. Este programa utiliza la base de datos SILVA versión 123 para realizar el análisis. En ambos casos los OTUs se eligieron utilizando la base de datos correspondiente, Green Genes v.13\_5 para PICRUSt y SILVA v.123 para Tax4Fun.

## Capítulo 3. Resultados

### 3.1. Extracción de ADN

Se realizaron extracciones de ADN de un total de 66 muestras de sedimento liofilizado, las cuales se corresponden a las 11 profundidades definidas en la columna sedimentaria. Estas profundidades iban desde la capa superficial, los primeros 20 cm, hasta los 220 cm, divididas en segmentos de 20 cm para cada una de las tres estaciones de muestreo y de los dos núcleos por estación (Tabla 2). El promedio de concentración fue 10 ng/ $\mu$ L siendo la máxima concentración 42.1 ng/ $\mu$ L y la mínima de 1,35 ng/ $\mu$ L. Las muestras que presentaron problemas durante la extracción fueron las muestras del núcleo E2N3 a partir de los 100 cm, que era la zona donde se registraron las temperaturas más altas. Las muestras extraídas y la concentración de ADN (ng/ $\mu$ L) se muestran en la Tabla 3.

**Tabla 3** Concentración de ADN (ng/ $\mu$ L) en las muestras de sedimento. En amarillo se resaltan las extracciones que tuvieron dificultad para ser extraídas y debieron hacerse varias extracciones y varias amplificaciones de la misma muestra. En verde más claro la que presentó menor concentración.

Estación/ Profundidad(cm)	E2N3	E2N4	E4N1	E4N2	E9N1	E9N2
0-20	7.85	9.05	20.35	42.1	8.75	9.25
20-40	6.05	1.9	14.4	20.3	9.5	8.2
40-60	5.15	4.2	16.6	13.45	9.15	8.65
60-80	3.8	1.35	19.4	10.95	9.45	9.45
80-100	3.45	3	12.15	8.7	7.6	11
100-120	6.65	9.8	15	19.15	8.9	9.15
120-140	6.85	6.95	11.6	8.25	7.5	6.35
140-160	6.4	3.45	7.05	4.5	1.7	5.55
160-180	3.75	6.4	7.7	3.55	5.55	5.6
180-200	1.9	5.15	5.4	5.65	5.55	5.9
200-220	9.2	8.6	14.6	15.45	11.05	9.45

### 3.2. Amplificación de ADN

Se realizaron reacciones de amplificación a partir del ADN extraído de las muestras de sedimento, utilizando de 2.5 a 10 ng de ácidos nucleicos (según las mediciones realizadas en Nanodrop) para la amplificación. La mezcla de reacción fue de 20  $\mu$ L de volumen final. En total se lograron amplificar 60 de las 66 seleccionadas inicialmente (resaltadas en verde en la Tabla 3) sin demasiada dificultad y en las condiciones previamente optimizadas. Las otras seis muestras mostraron problemas de amplificación, y coincidió que fueron las mismas muestras que presentaron problemas durante la extracción de ADN. Para resolver este problema, se procedió a realizar varias amplificaciones y a concentrar los amplicones obtenidos de estas muestras mediante el uso del kit QIAquick® PCR purification kit, (QIAGEN®) siguiendo las recomendaciones del fabricante. Con esta metodología se logró preparar la biblioteca metagenómica para su secuenciación con todas las muestras.

### 3.3. Preparación de las bibliotecas metagenómicas

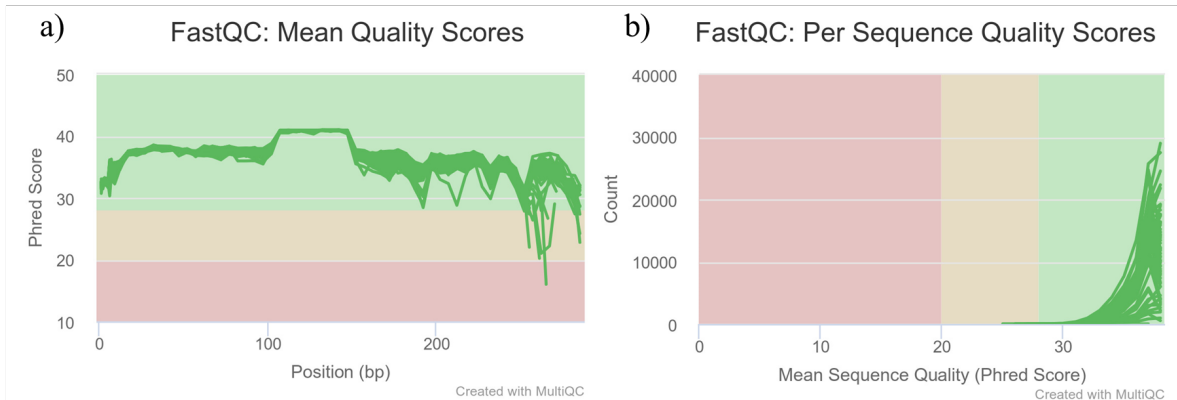
Los productos de PCR obtenidos, fueron normalizados para que todas las librerías tuvieran la misma concentración utilizando el kit SequalPrep™ (Invitrogen, Carlsbad, CA). Para la normalización se utilizaron 10  $\mu$ L de los productos de PCR en cada pocillo. Posteriormente se cuantificó la concentración de ADN final para elaborar los *metapools* para la secuenciación, y llevarlos a la concentración de trabajo (6 pM). Finalmente se preparó la solución de carga utilizando 50  $\mu$ L de una biblioteca PhiX (6 pM) previamente preparada y 950  $\mu$ L de nuestra biblioteca metagenómica (6 pM). Estas soluciones se cargaron en la placa de secuenciación, y se procedió a realizar la secuenciación en la plataforma MiSeq de Illumina del CICESE, se obtuvieron un total de 2 863 531 secuencias.

### 3.4. Control de calidad de las secuencias

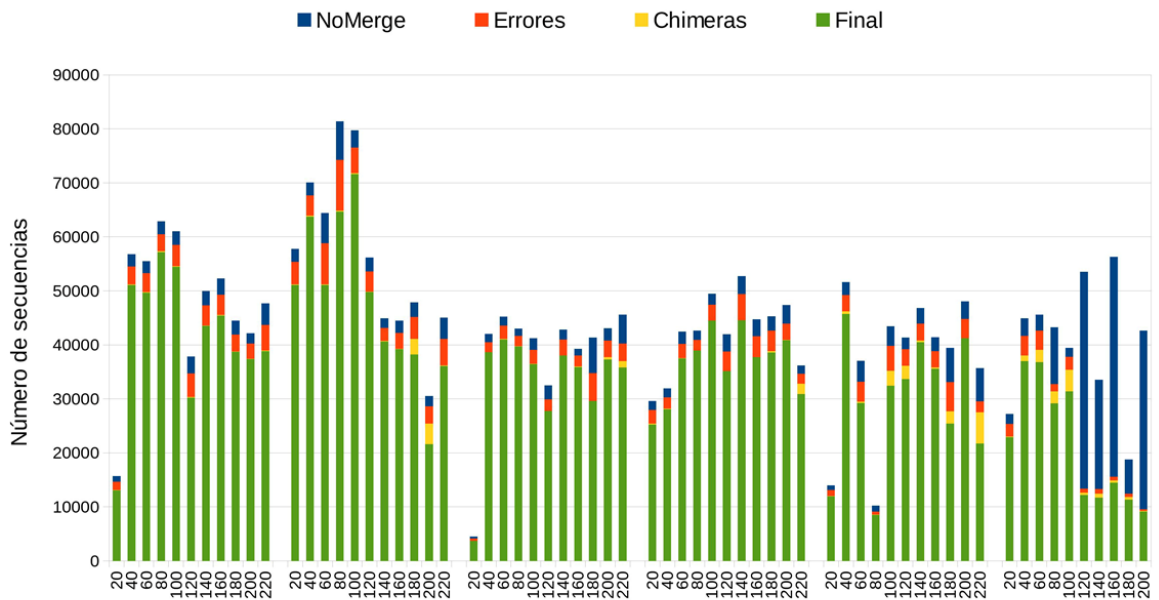
La calidad de las secuencias obtenidas se verificó utilizando el programa FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Las secuencias obtenidas se descargaron demultiplexadas del portal BaseSpace® de Illumina. y verificaron los parámetros de la secuenciación. Se

observó una buena calidad de las secuencias obtenidas (Quality score >30) en la mayoría de ellas, sin embargo se observó una caída en la calidad al final de la secuencia como se muestra en la Figura 5, pero este comportamiento es normal en este tipo de secuenciación (*paired-end*).

La relación entre las lecturas filtradas en cada etapa y las lecturas utilizadas en el análisis final se muestran en la Figura 6, incluyeron un total de 2 261 113 de secuencias.

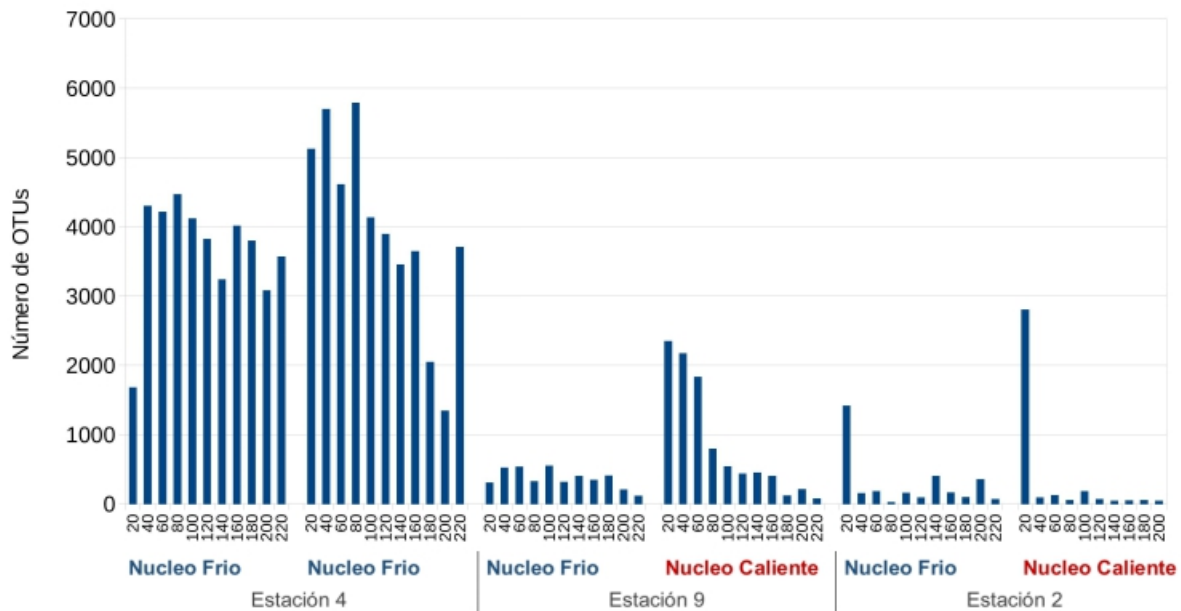


**Figura 5** Gráficas de calidad de las secuencias. a) Calidad de las secuencias obtenidas donde se muestran todas las muestras en un gráfico proporcionado por el programa MultiQC, b) distribución de la calidad de las secuencias obtenidas de todas las muestras.



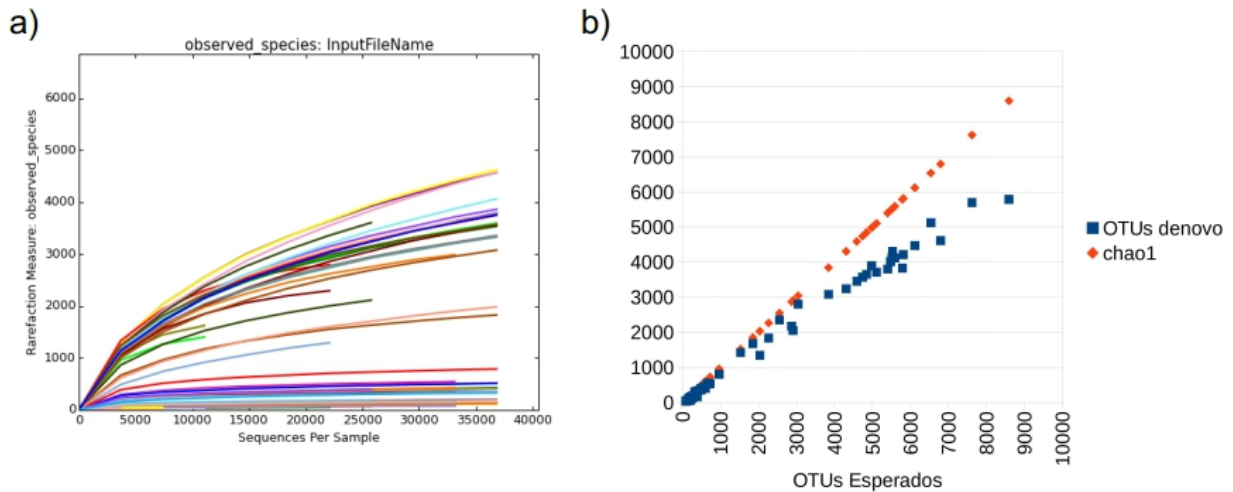
**Figura 6** Control de calidad de las secuencias obtenidas y secuencias utilizadas para el análisis en QIIME.

Una vez que se revisó la calidad de las secuencias, estas se incluyeron en el pipeline de análisis bioinformático observándose que la cantidad de OTUs entre muestras fue muy variable, desde los 27 en la muestra E2NP20 hasta los 5787 OTUs en la muestra E4N2P80. De manera general la mayoría de los núcleos presenta su máxima cantidad de OTUs en las secciones superficiales y va disminuyendo a medida que aumenta la profundidad en los núcleos, como se muestra en la Figura 7.



**Figura 7** Número de OTUs observados por muestra. En la muestra se comparan el número de los OTUs observados por para cada uno de los núcleos por estación.

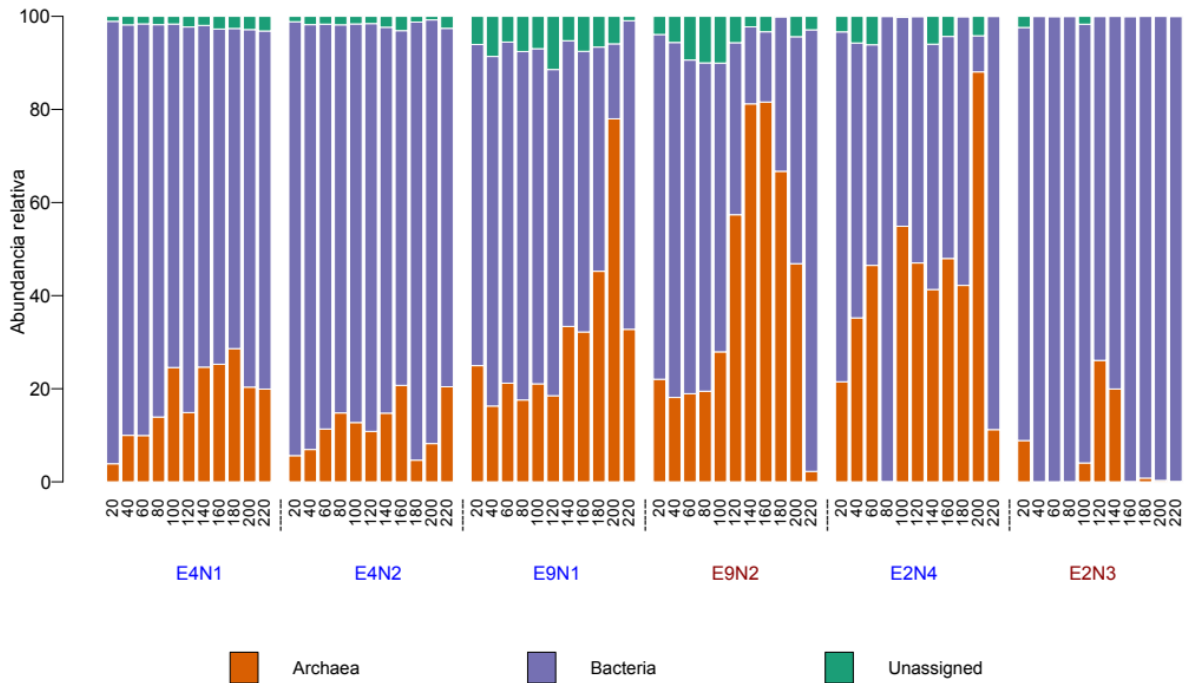
En la mayoría de las muestras la profundidad de secuenciación fue la suficiente para alcanzar el punto de inflexión de las curvas de rarefacción, lo que indica que prácticamente todo lo que había en la muestra fue secuenciado. Sin embargo, otras resultaron ser muy diversas y no se alcanzó la asíntota, aún en muestras con más de 35 000 secuencias, como se puede observar en la Figura 8a. En general y para la mayoría de las muestras, a partir de las 5000 secuencias ya prácticamente se había cubierto toda la diversidad procariota presente en esa muestra. El mismo resultado se observa al analizar la correlación entre los OTUs observados y los OTUs esperados según el índice Chao1 (Figura 8b), donde las muestras con alto número de OTUs observados fueron menos que los OTUs esperados. Aunque esto es algo esperado en aquellas muestras que presentan una alta diversidad. Sin embargo, en el peor de los casos se alcanzó a cubrir hasta un 75% de los OTUs esperados.



**Figura 8** a) Curva de rarefacción que muestra el esfuerzo de secuenciación en cada una de las muestras. En el eje de las X se describe el número de secuencias por muestra y el número de OTUs observados. b) Gráfica de correlación entre el número de OTUs observados y OTUs esperados. Los puntos rojos indican el valor Chao 1 que representa los valores estimados o esperados y los puntos azules los valores observados.

### 3.5. Descripción de la composición de la comunidad procariota.

Al analizar la primera categoría taxonómica, Dominio, se observó una mayor abundancia de OTUs bacterianos con respecto a los de arqueas en la mayoría de las muestras, como se observa en la Figura 9. A medida que se aumenta la profundidad, hay una tendencia al aumento de las arqueas, sin embargo éstas disminuyen en el último segmento de algunos de los núcleos analizados y, en particular, en la estación E2N3, aunque esto puede estar relacionado con el bajo rendimiento de la muestra en término de cantidad de ADN y este estar relacionado con la alta temperatura. Por otro lado, los OTUs no asignados fueron muy pocos siendo la muestra de la E9N1P120 la que presentó mayor número de OTUs no asignados (con un máximo 11.5% ).



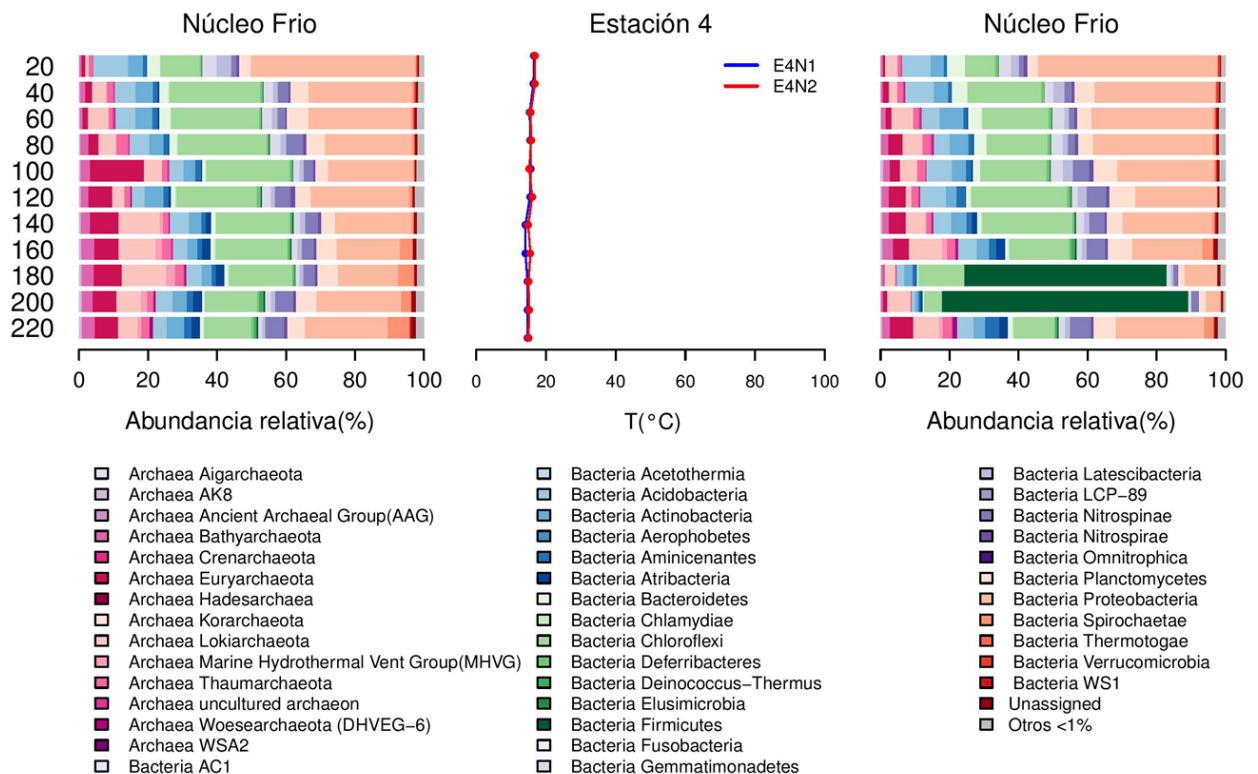
**Figura 9** Abundancia relativa de los dominios Arquea (en naranja) y Bacteria (en morado). En verde se muestran aquellos OTUs que no lograron asignarse a ninguno de los dos dominios.

A nivel de Filo, se analizaron los perfiles de profundidad entre los núcleos de cada estación y se observaron los cambios de las comunidades microbianas procariotas en relación a la profundidad y a las variaciones en la temperatura, sobre todo en aquellos núcleos donde se presentaban temperaturas muy altas. En general, en la mayoría de las muestras de todas las estaciones los filos predominantes fueron *Proteobacteria*, *Chloroflexi* y en menor medida el *Hadesarcheota*, *Acidobacteria* y *Aerophobetes*. Sin embargo, para tener una mayor idea de cómo la temperatura y la profundidad afecta a la distribución de la comunidad microbiana se realizó una primera comparación de los filos por estaciones.

### 3.5.1. Comunidad procariota de la Estación 4

Como se puede ver en la Figura 10, los núcleos de la estación 4 no presentaban muchas diferencias entre ellos y tampoco a lo largo de todo el perfil de profundidad. En general los filos más abundantes fueron *Proteobacteria* (43 y 47%), *Acidobacteria* (10.6 y 9%), *Chloroflexi* (10.4 y 7.8%) y *Planctomycetes* (6.5 y 5.4%), seguido por *Latescibacteria* (4.5%) en el primer núcleo y *Bacteroidetes* (5.2%) en el

segundo núcleo. En las secciones más profundas esta proporción se vio alterada siendo los de mayor abundancia *Proteobacteria*. Al aumentar la profundidad el filo *Acidobacteria* se vio desplazado por *Chloroflexi*, que pasó a ser el filo dominante junto a *Proteobacteria* hasta los 220 cm en el núcleo 1 y hasta los 160 centímetros en el núcleo 2 (12.6-23.6% y 16.15--28.1% respectivamente), como se muestra en la Figura 10. Las arqueas mostraron un incremento a medida que también se incrementaba la profundidad, destacando las *Lokiarchaeotas* como grupo más abundante en la mayoría de las secciones a partir de los 140 cm. Uno de los resultados más llamativos de esta estación es fue la alta abundancia del filo *Firmicutes* (57.4-70%) en las secciones de los 180 cm y de los 200 cm de profundidad en el núcleo 2 (E4N2). Aunque inicialmente se pensó que esta dominancia de *Firmicutes* podría deberse a una contaminación, realmente los resultados indicaron que eran grupos que sí se encuentran presentes en los sedimentos de manera natural y que es la familia *Bacilliaceae* la que predomina, como se describirá en la sección 3.7.



**Figura 10** Abundancias relativas de los organismos procariontes, a nivel filo, a lo largo del perfil de profundidad de la estación 4. A izquierda se observa la abundancia relativa del núcleo frío. En el centro se muestra el perfil de temperatura de ambos núcleos. A la derecha, se puede observar la abundancia relativa el núcleo caliente.



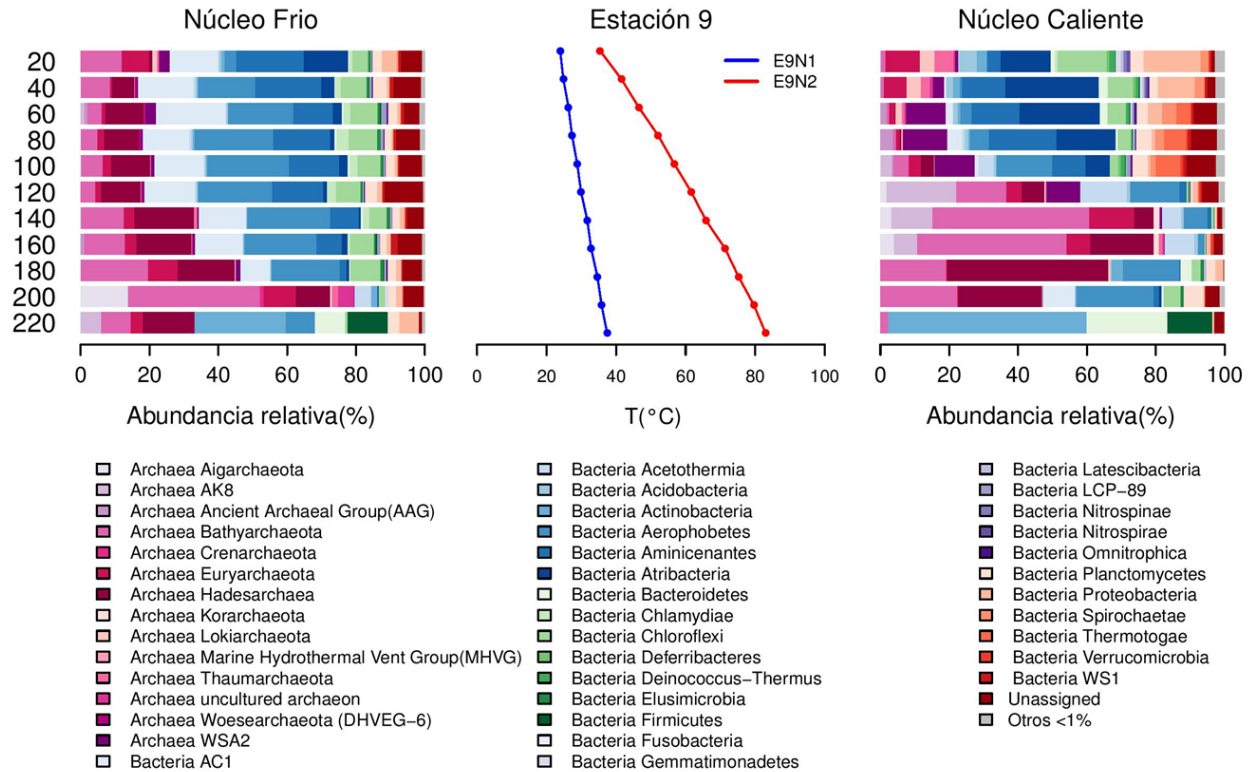
### 3.5.2. Comunidad procariota de la Estación 9

En esta estación la mayoría de los filos están presentes en ambos núcleos, aunque claramente en el núcleo caliente (E9N2) se observó un cambio más marcado en el patrón de las abundancias, a partir de la sección de 120 cm. Este cambio podría estar relacionado con la temperatura pues en el núcleo frío (E9N1) este cambio es más gradual, solo con la excepción de los 200 cm (Figura 11), además en el caliente el cambio se produce cuando la temperatura está por encima de los 60 °C. En ambos núcleos también se observó que hay un incremento de las arqueas a medida que aumenta la profundidad, pero en el núcleo caliente es mucho más marcado a partir de los 200 cm, pasando a ser el grupo dominante en la mayoría de las secciones por debajo de esa profundidad.

Ya más en detalle, en la primera sección superficial, se encontraron diferencias entre ambos núcleos. En el núcleo frío dominaban los filos *Aminicenantes* (17.3%), *AC1* (12.2%), *Atribacteria* (10.9%) y *Bathyarcheota* (7.8%), mientras que en el núcleo caliente fueron los filos *Proteobacteria* (15.3%), *Chloroflexi* (13%), *Atribacteria* (11.3%) y *Euryarcheota* (9.3%). También se presentaron diferencias en los filos dominantes entre los núcleos. En el núcleo frío se observó que *Aerophobetes* y *AC1* dominaban desde la sección subsuperficial hasta los 160 cm de profundidad (15.8 y 15.2 en subsuperficie y 20.4 y 13.6% a 160 cm respectivamente). En cambio, el filo *Aminicenantes* se encontró en mayor proporción en la sección que comprende desde los 80 cm hasta los 120 cm (14-16%). Dentro de las arqueas el filo *Hadesarcheota* fue el más abundante a lo largo de toda la columna desde los 60 cm (6.2-17%). Sin embargo, a los 200 cm *Bathyarcheota* (38%) reemplaza a *Hadesarcheota* (9.7%) como arquea más abundante y aparece *Aigarcheota* (13.4%), un filo perteneciente al clado de arqueas TACK, Finalmente, en la sección más profunda, se volvió a observar un patrón similar a la mayoría de las profundidades pero destacando la gran proporción de *Actinobacteria* (25.8%) y la aparición de *Firmicutes* (11.4 %) y *Bacteroidetes* (8.4%).

En el núcleo caliente (E9N2), desde los 40 hasta los 80 cm, se observó que *Atribacteria* (22-15.5%) junto con los *Aminicenantes* (17.6-10.4%) fueron los filos más abundantes. En la sección siguiente la comunidad estaba dominada por *Aerophobetes* (15.5%) y *WSA2* (11.1%) y a 120 cm de profundidad predominaba el filo *AK8* (19.2%), el cual se encuentra presente únicamente desde los 100 cm hasta los 160 cm. Desde los 140 hasta los 160 cm se observó que *Bathyarcheota* (44.7-42%) era uno de los filos más dominantes, así como *Euryarcheota* (12.6-6.7%) a 140 cm y *Hadesarcheota* (6-18.5%) a 160 cm, siendo este último el que dominaba la comunidad desde los 180 hasta los 200 cm (46.7-23.7%), solo seguido por *Aerophobetes* y *Bathyarcheota* (16.4-22.27 y 19.19-22% respectivamente). En la sección

más profunda (220 cm), al igual que ocurría en el núcleo frío se encontró una gran proporción de *Actinobacteria* (56.3%), *Bacteroidetes* (24.1%) y *Firmicutes* (13%).



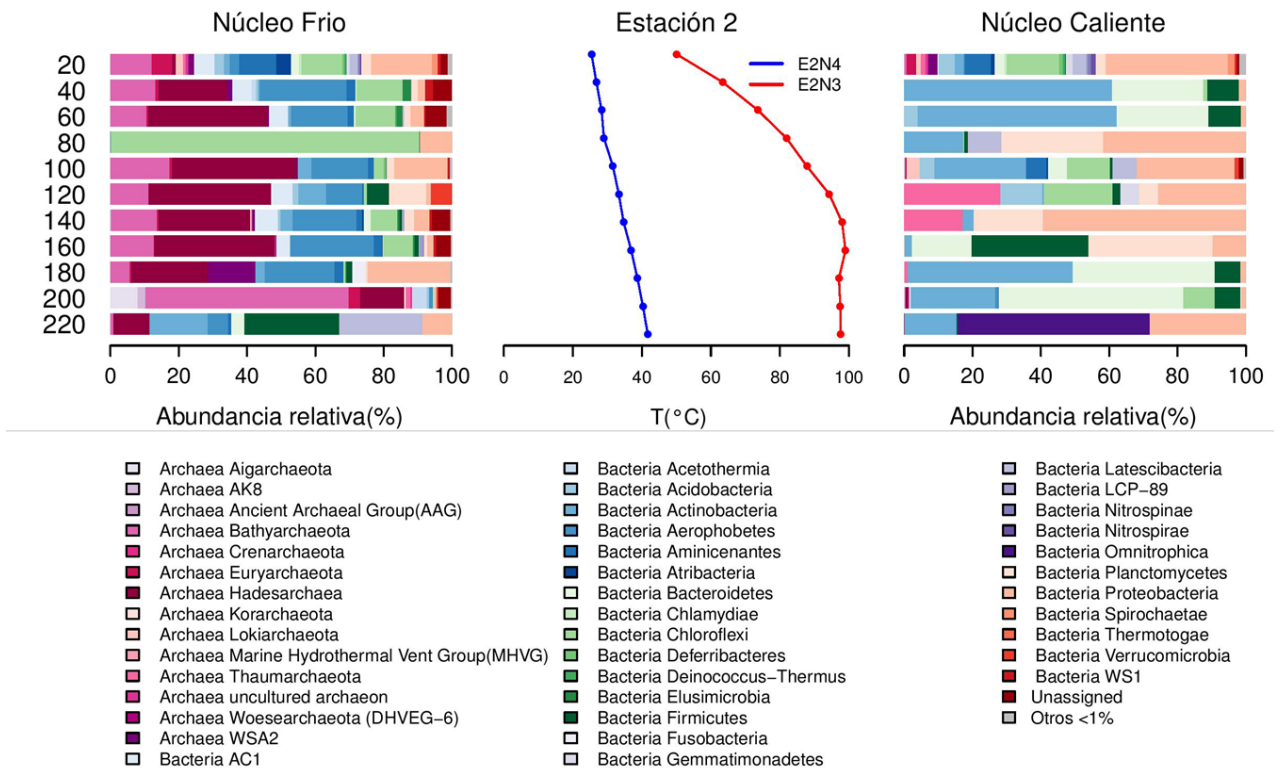
**Figura 11** Abundancias relativas de los organismos procariontes, a nivel filo, a lo largo del perfil de profundidad de la estación 9. A izquierda se observa la abundancia relativa del núcleo frío. En el centro se muestra el perfil de temperatura de ambos núcleos. A la derecha, se puede observar la abundancia relativa el núcleo caliente.

### 3.5.3. Comunidad procarionta de la Estación 2

LaFigura 12 muestra la distribución de los principales filos de la estación 2, y como se puede observar, las abundancias y composición de los filos en ambos núcleos fue claramente diferente, siendo el núcleo caliente (E2N3) el que presentó una mayor zonación. Analizando por secciones se observó que en la sección superficial (los primeros 20 cm) de ambos núcleos dominaban los filos de *Proteobacteria* (19.7 y 32.5%) y *Chloroflexi* (11.9 y 14.1%) seguido por *Bathyarcheota* (5.86%) en el núcleo frío y *Planctomycetes* (9.65%) en el núcleo caliente, y que ambos núcleos eran similares entre ellos en composición, pero variaban las abundancias. En las secciones más profundas dominaban los filos de *Hadesarcheota* en el núcleo frío y *Proteobacteria* en el núcleo caliente.

A partir de los 40 cm, en el núcleo frío (E2N4) se observó que las arqueas y las bacterias se repartían casi a la mitad los filos en casi todas las profundidades, con excepción de tres (80, 200 y 220 cm). Por el lado de las arqueas dominaban los filos *Hadesarcheota* (19.72-35.57%) y *Bathyarcheota* (10.9-13.15%) y por las bacterias los filos *Chloroflexi* (2.8-13.25%) y *Aerophobetes* (10.62-24.27%). La primera profundidad que presentó un patrón diferente fue a los 80 cm, donde se observó que el filo *Chloroflexi* englobaba el 90.3% del total de los OTUs detectados, seguido por *Proteobacteria* con un 9.5% de los OTUs. A los 200 cm son las arqueas las que dominaban prácticamente el 85% de la sección. Por último, a los 220 cm también se observó otro patrón diferente en las abundancias donde *Firmicutes* (27.7%) pasó a ser el filo más abundante, seguido por *Latescibacteria* (23.6%), que solo aparece a esta profundidad, *Actinobacteria* (16.4%) y *Hadesarchea* (10.3%). Y a partir de los 120 cm la abundancia del filo *Chloroflexi* disminuyó hasta prácticamente desaparecer (0.9%), y es desplazado por el filo *Actinobacteria* (8%). En las secciones más profundas de este núcleo se conserva la abundancia de *Hadesarcheota* y aparece el filo *Aigarcheota* (7.38%), el cual pertenece al clado TACK.

En el núcleo caliente de la estación 2 (E2N3) presentó un patrón menos homogéneo y continuo que el núcleo frío, dominado los filos *Actinobacteria* (57-57.7%), *Bacteroidetes* (25-26%) y *Firmicutes* (9.5 - 10.4%) desde los 40 cm hasta los 60 cm. En las siguientes secciones se observó que *Proteobacteria* (41-26%) y *Actinobacteria* (17.2-23.5%) dominaban desde los 80 cm hasta los 100 cm de profundidad. A partir de los 120 cm, fue *Thaumarcheota* es el filo que estaba dominando esta sección (26.5%), junto con *Proteobacteria*, *Chloroflexi* y *Acidobacteria* (24%, 18.8% y 11.5% respectivamente). *Deinococcus-Thermus* dominaba en las profundidades que van de los 120 hasta los 140 cm (24.3-39.2%), mientras que a los 160 cm de profundidad (99 °C) se observó que los filos *Nitrospirae* (27.7%), *Planctomycetes* (23.28%) y *Firmicutes* (22.4%) eran los más abundantes seguidos por *Bacteroidetes* (17.8%), *Proteobacteria* (6.8%) y *Actinobacteria* (1.5%). En las secciones de 180 y 200 cm los filos *Actinobacteria*, *Bacteroidetes*, *Firmicutes* y *Proteobacteria* (6.9-47.6%, 11-40.6%, 1.7-7.53%, 2-76.9% respectivamente) vuelven a ser los más abundantes al igual que a los 40 y 60 cm. Finalmente a los 220 cm de se encontró que *Chloroflexi* y *Omnitrophica* comprendían la mayoría de los OTUs en esta sección (39 y 33.2% respectivamente) solo seguidos por *Proteobacteria* y *Actinobacteria* (16.5 y 10.3% respectivamente). Llama la atención la presencia del filo *Omnitrophica* que solo se detectó en este núcleo y a esta profundidad y con la mayor abundancia de esta sección.



**Figura 12** Abundancias relativas de los organismos procariontes, a nivel filo, a lo largo del perfil de profundidad de la estación 2. A izquierda se observa la abundancia relativa del núcleo frío. En el centro se muestra el perfil de temperatura de ambos núcleos, con línea azul el núcleo frío y con línea roja el núcleo caliente. A la derecha, se puede observar la abundancia relativa el núcleo caliente.

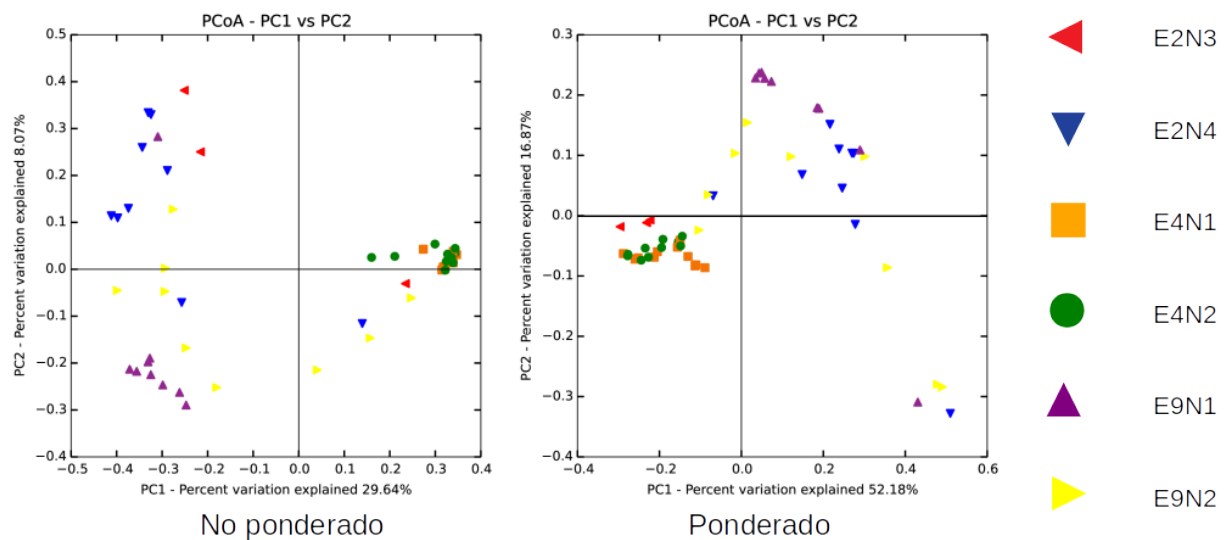
### 3.6. Comparación entre estaciones mediante análisis de coordenadas principales

Se realizaron análisis de coordenadas principales (PCoA), con la métrica UniFrac, para observar si existían relaciones entre las estaciones o las muestras, y si estas relaciones se debían a algún factor específico. Para ello se observaron tres factores principales: el efecto de la estación, el efecto de la profundidad del sedimento y el efecto de la temperatura.

#### 3.6.1. Análisis por estación

El análisis de las muestras mediante PcoAs mostró que las comunidades microbianas variaron entre estaciones (Figura 13). Sin embargo, sí se observaron algunos patrones. El primero es que las muestras de la E4 se agrupaban claramente entre ellas, independientemente de la profundidad, observándose

únicamente una pequeña separación en las muestras más profundas. Junto a estas, también agrupaban las muestras más superficiales (0-20 cm) de ambos núcleos de la estación 2 y el núcleo caliente de la estación 9. Igualmente se observó una agrupación de la mayoría de las muestras del núcleo frío de la estación 9 (E9N1), exceptuado la muestra que se encuentra a 200 m que se situaba junto a muestras que se estaban por debajo de los 100 cm de profundidad. Esta observación es más clara en el análisis ponderado, lo que indicaría que estas muestras contienen elementos taxonómicos similares, aunque no sean exactamente los mismos grupos.. El grupo que contenía las muestras del núcleo frío de la estación 9 agrupaba también con aquellas muestras que se encuentran a menos de 120 cm de profundidad del núcleo caliente de la misma estación. La estación E9N2 es fue la que presentó mayor dispersión y no se obserbaron patrones claros.

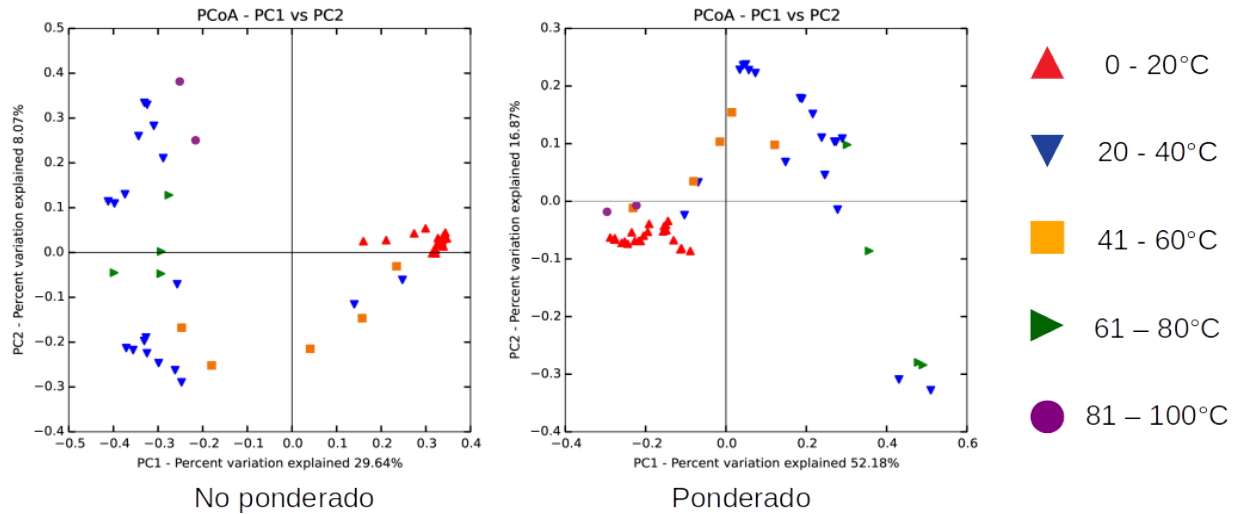


**Figura 13** Análisis de coordenadas principales. A la izquierda se puede ver el análisis no ponderado. A la derecha se puede ver el análisis ponderado. Los diferentes colores indican la estación de la que proviene cada núcleo.

### 3.6.2. Análisis por temperatura

Al observar las agrupaciones por temperatura se pudo ver que, aunque las muestras más frías se mantienen agrupadas en ambos análisis (ponderado y no ponderado), también se observó que agrupaban con éstas muestras otras se encontraban a casi 100 °C, como se muestra en la Figura 14. Esto nos indicaría que la temperatura es el único factor determinante en estas muestras superficiales, o

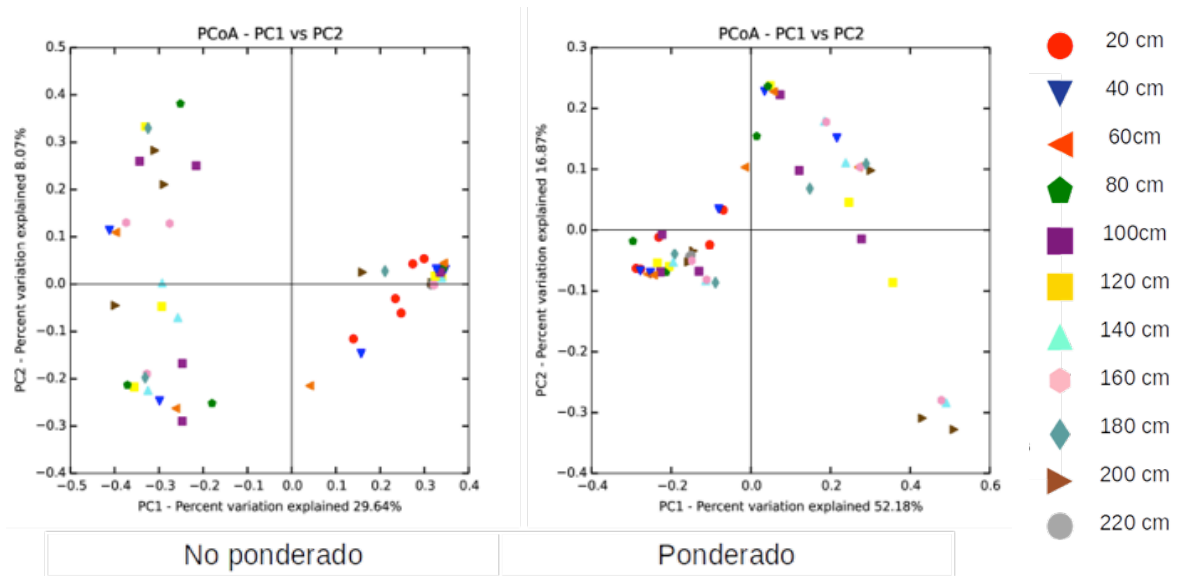
en esas muestras calientes. Cuando se observa el agrupamiento por profundidad en la columna sedimentaria nos encontramos que en la estación 4 y en la estación 9 no parece haber influencia de esta variable, debido a que cada estación forma su propio grupo.



**Figura 14** Análisis de coordenadas principales. A la izquierda se puede ver el análisis no ponderado, A la derecha se puede ver el análisis ponderado. Las muestras se agruparon por intervalos de 20 °C de temperatura.

### 3.6.3. Análisis por profundidad

Finalmente se intentó ver si era la profundidad la que tenía un efecto claro en la distribución de las comunidades. Al observar las agrupaciones por profundidad en la columna sedimentaria se distinguió una clara agrupación en las secciones superficiales independiente de la estación de procedencia. Sin embargo, en las muestras que se encuentran por debajo de los 20 cm ya no muestran patrones de agrupación claros entre las diferentes secciones, como se muestra en la Figura 15. **¡Error! No se encuentra el origen de la referencia. .**



**Figura 15** Análisis de coordenadas principales indicando las profundidades en la columna sedimentaria. A la izquierda se puede ver el análisis no ponderado y a la derecha el análisis ponderado.

### 3.7. Análisis de las familias más abundantes

También se realizó un análisis a nivel de familia, para tener una mejor idea de los principales protagonistas de la actividad microbiana a diferentes profundidades. En esta sección se describen las variaciones en las familias más abundantes por estación. En la Figura 16 se muestran las variaciones de estas familias en la estación 4, la Figura 17 en la estación 9 y por último en la Figura 18 los perfiles los núcleos de la estación 2.

#### 3.7.1. Principales familias de la Estación 4

En la primera sección del núcleo E4N1 (0-20cm), se encontró que las familias *Anaerolineaceae* (8.5%) y *JTB255*(7.5%) fueron las más dominantes, mientras que las otras familias se encontraron en una menor proporción (< 5%). Por debajo de los 20 cm, se observó un cambio en las familias dominantes, donde se observa un aumento de la familia *Anaerolinaceae* (hasta un 21.5%) y una notable disminución de la familia *JTB255*(<1%). Mientras la que familia *Anaerolinaceae* dominó desde los 20 hasta los 220 cm de profundidad en este núcleo, en otras secciones destacaron otras familias: a los 60 cm destacó la presencia de la familia *Planctomycetaceae* (5.13%); desde los 140 a los 180 cm de profundidad se presentó una familia no cultivada de *Lokiarchaeota* (10.4 y 8.6% respectivamente); la familia *ANME-2a-*

2b junto con la familia *Desulfobacteraceae* presentaron un aumento en la abundancia desde los 80 hasta los 120 cm de profundidad (9.6 y 4% respectivamente). En la sección que comprende desde los 180 hasta 200 cm se observó un aumento en la abundancia relativa de una familia del orden de los *Thermoplasmatales* (*Marine Benthic Group D/ DHVEG-1*) (4.8% 6.2%). En la sección más profunda (220 cm) predominaron los OTUs de la familia *Anaerolinaceae* (9.7%) así como la familia *Spirochaetaceae* (9.7%), *Marine Benthic Group D/ DHVEG-1* (4.9%) y una familia no cultivada del filo *Lokiarchaeota* (4.5%). Por debajo de estos se encontraron las familias de *Bathyarchaeota* (3.3%) y una familia no cultivada del filo *Nitrospinae* clase *MD2898-B26* (2.8%).

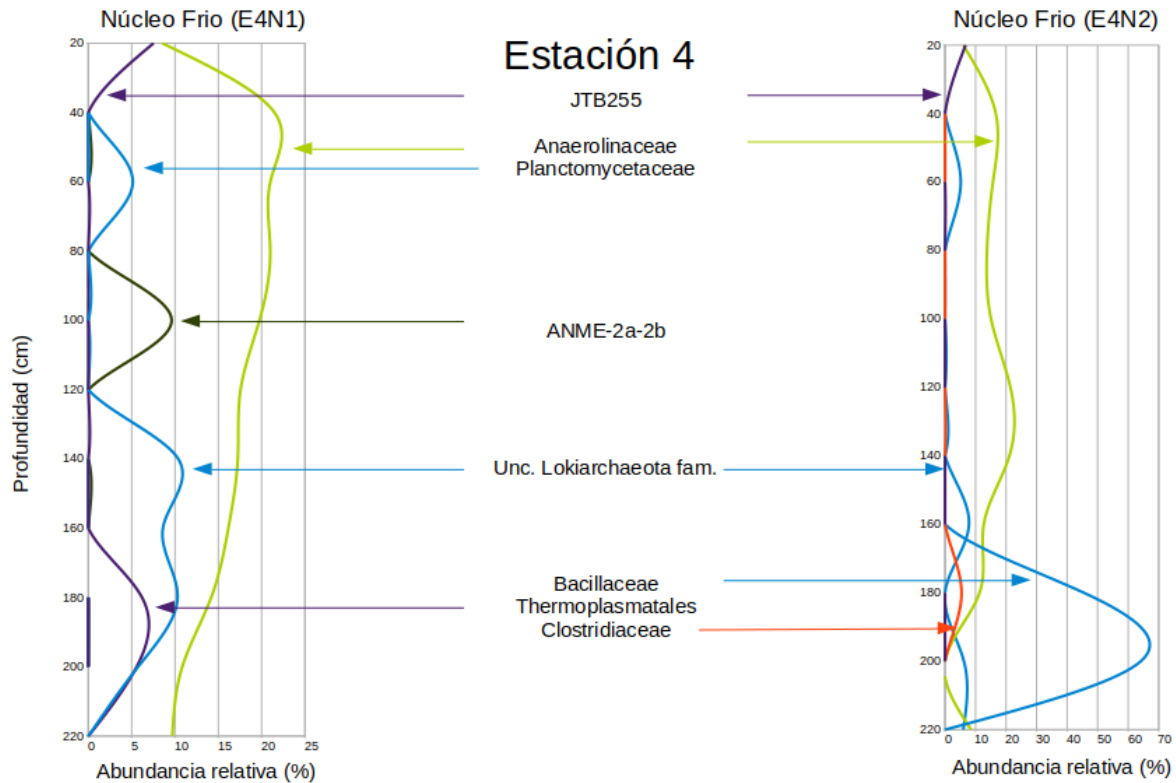
En el núcleo E4N2 también se observó que las familias *Anaerolinaceae* (6.7%) y *JTB255* (5.9%) son las más abundantes en la primera sección, el resto de las familias presentes en esta sección se encontraron debajo del 5% e incluyen a las familias *Desulfobulbaceae*, *Syntrophobacteraceae*, una familia desconocida del filo *Gammaproteobacteria*, la familia *Desulfobulbaceae*, y una familia del filo *Lokiarchaeota*.

A medida que aumenta la profundidad se observó una disminución de la familia *JTB255*, así como un aumento en la abundancia relativa de la familia *Anaerolinaceae* hasta los 140 cm de profundidad. En la sección que va desde 40 hasta los 120 cm de profundidad incrementó la abundancia relativa de la familia *Latescibacteria* pero en proporciones bajas (2.8-2%). Fue más destacable el aumento en la abundancia relativa de las familias *Lokiarchaeota* y *Sandaraniaceae* desde los 60 hasta los 180 cm (2.9-7.8% y 2.1-4.5% respectivamente).

Desde los 100 hasta los 140 cm aumentaron las familias de *Planctomycetaceae*, *Marine Benthic Group D / DHVEG-1*, y una familia del filo *Nitrospinae* clase *MD2898-B26*, observándose un pico de abundancia relativa de una familia no cultivada del filo *Lokiarchaeota* a los 160cm (7.8%).

Por último de 180 a los 200 cm se observó un cambio en la comunidad microbiana donde predominaron los OTUs de la familia *Bacillaceae* (45.5-64.1%). Sin embargo, se observó otro cambio en la sección más profunda donde predominaron los OTUs de las familias *Anaerolinaceae* (8.5%) y una familia no cultivada del filo *Lokiarchaeota* (5.9%), las demás familias mostraron abundancias por debajo del 5%





**Figura 16** Gráfica de líneas que muestra el perfil de profundidad de las familias más abundantes de los núcleos de la estación 4.

### 3.7.2. Principales familias de la Estación 9

La estación 9 al igual que la estación 2, presentó diferencias de temperatura en los núcleos analizados. El núcleo frío de esta estación (E9N1) presentó pocas diferencias a lo largo de la columna sedimentaria, como se observa en la Figura 17. La sección más superficial de este núcleo se encontraba dominada por las familias no cultivadas de los filos *Aminicenantes*(17.4%), *Atribacteria*(10.9%), *Bathyarchaeota*(8.1%), así como dos familias del filo *AC1* (6.5 y 5.6%) y la familia *ANME-1a* (5.4%).

En las secciones que encuentran por debajo de los 20 cm se observaron algunos cambios en las abundancias de las familias, siendo las secciones más profundas (a partir de los 140 cm) las que presentaron una mayor variación. Aunque la familia del filo *AC1* ya se observó en la sección superficial, mostró una mayor proporción a partir de los 20cm y hasta los 180 cm, donde se mantuvo entre el 10 y 15%. También, se observó un cambio en la abundancia de la familia del filo *Aminicenantes*, que fue disminuyendo a medida que aumentaba la profundidad, hasta llegar a un 2% a los 180 cm.

A partir de los 120 cm se observó un aumento en la abundancia de una familia del filo *Bathyarcheota* desde 1.3% en la sección de 120 cm, hasta más de un 30% a 200 cm, para posteriormente disminuir por debajo del 9% en la sección más profunda. La familia *Hadesarcheota* también se encontró distribuida a lo largo de toda la columna sedimentaria, variando desde los 20 cm con un 6.1% hasta aproximadamente un 17% a los 180 cm de profundidad.

En las secciones más profundas, se observó a los 200 cm un aumento en la abundancia de una familia del filo *Aigarcheota* género *Caldiarcheum* (7.2%), y la presencia de familias de los fillos *Bathyarchaeota* (30%), *Hadesarchaeota* (9.5%) y de la clase *ANME-1* (4.7%). A los 220 cm destacaron la familia *Corynebacteriaceae* (24.3%) (posible contaminación) y varias familias no cultivadas de los fillos *Hadesarchaea* (15.1%), *Bacillaceae* (11.6%), *Bathyarchaeota* (8.6%), la familia *Flavobacteriaceae* (7.5%), así como una familia del filo *AK8* (6%).

El análisis del núcleo caliente (E9N2) mostró, que en la sección más superficial, los grupos dominantes fueron una familia no cultivada del filo *Atribacteria* con un 11%, por debajo de ésta se encontraban la familia *Anaerolinaceae* (10%) y la familia *Marine Benthic Group D/ DHVEG-1* (6%). Los demás grupos mostraron abundancias por debajo del 5% e incluían a familias no cultivadas de los fillos *Lokiarchaeota*, *Aminicenantes* y *Thaumarchaeota Marine Group I*.

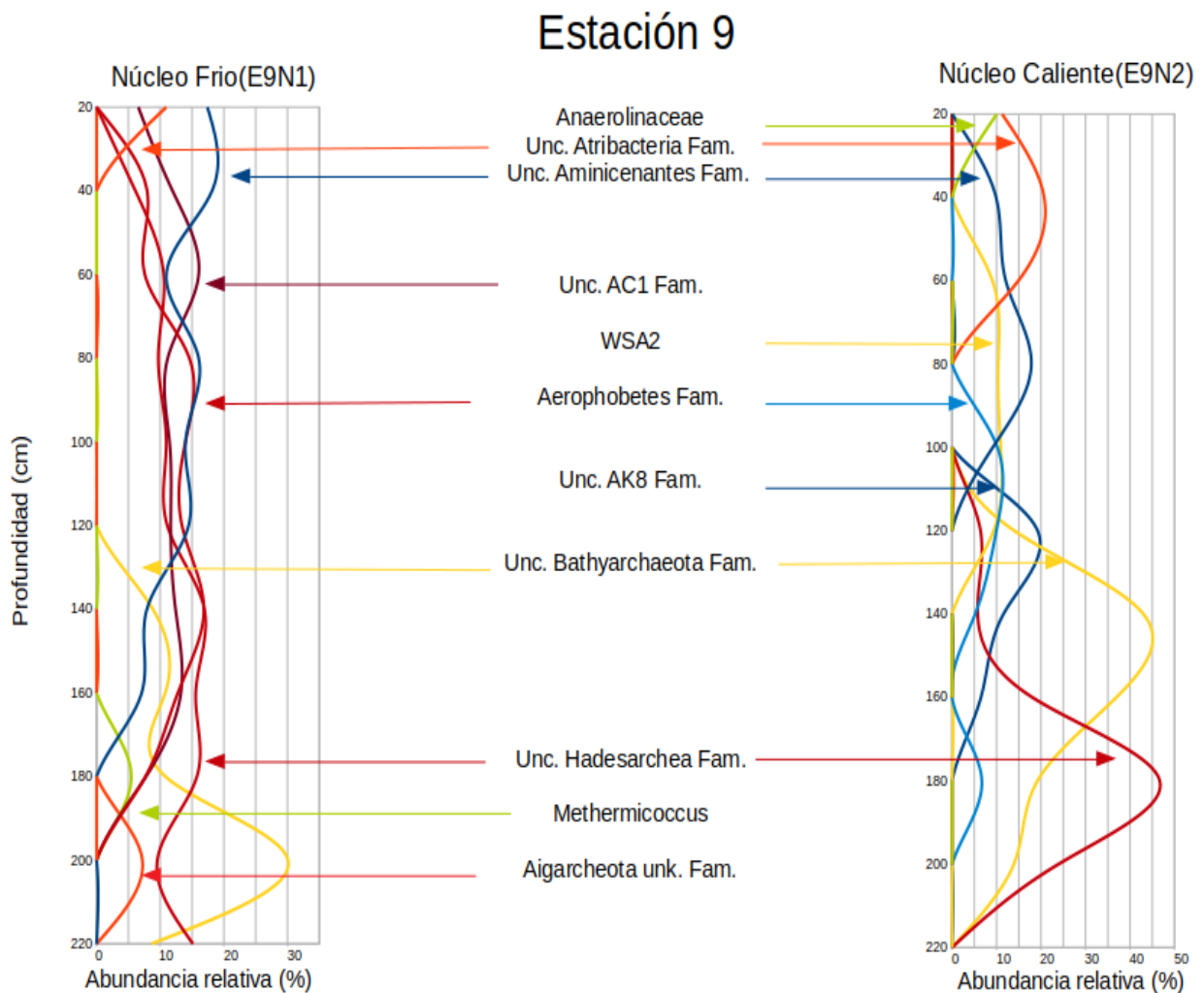
En la sección que va desde los 20 a los 40 cm de profundidad se observó un aumento en una familia no cultivada del filo *Atribacteria* (20%), así como del filo *Aminicenantes* (9%). En las secciones posteriores, la familia *Anaerolinaceae* disminuye a medida que aumenta la profundidad. A los 60 cm, la familia *Atribacteria* disminuyó (14%), al contrario de lo que ocurrió con una familia de la clase *19c-33* del filo *WSA2*, que aumentó y se mantuvo constante hasta los 120 cm.

Ya en las zonas más profundas, de 100 a 120 cm, se observó un aumento en la abundancia de una familia no cultivada del filo *Aerophobetes* y de una familia del filo *AK8*, el cual fue el más abundante a 120 cm de profundidad. A esta profundidad también se observó una familia del filo *Hadesarcheota* con un 6.4% de abundancia. Esta misma familia aumenta notablemente a medida que aumenta la profundidad hasta llegar a poco más del 40% del total en la sección de 180 cm.

A los 120 cm también se observó que familia del filo *Acetothermia* estaba presente con un 13% y que se mantenía por encima del 5% hasta los 160 cm, donde alcanzó un 8.4%. En la sección de los 140 cm de

profundidad, se observó un aumento hasta el 10% en la familia *pMC2A24* del orden *Thermoplasmatales*. y la familia del filo *Bathyarcheota* englobó el 44% de los OTUs observados a esta profundidad.

A partir de los 160 cm empieza a disminuir el número de grupos observados, donde solo predominan familias no cultivadas de los filos *Hadesarcheota* (46%), *Bathyarcheota* (19%) y *Aerophobetes* (6.6%), los cuales disminuyen casi a la mitad en la sección que se encuentra a 200 cm. En la sección más profunda se observa que las familias dominantes son *Corynecateriaceae* (55.8%), *Flavobacteriaceae* (9.3%), *Clostridiaceae* (9.2%), *Porphyromonadaceae* (7.4%) y *Prevotellaceae* (6.7%).



**Figura 17.** Gráfica de líneas que muestra el perfil de profundidad de las familias más abundantes de los núcleos de la estación 9. A la izquierda se muestra el perfil del núcleo frío y a la derecha el del núcleo caliente.

### 3.7.3. Principales familias de la Estación 2

La estación 2 fue la que presentó mayor variación entre núcleos. En la Figura 18 se muestran las variaciones de las principales familias presentes a lo largo del perfil sedimentario de ambos núcleos. En la sección superficial del núcleo frío de esta estación se observó que dominaron las familias no cultivadas de los filos *Aminicenantes* (8.6%) y *Bathyarchaeota* (7.9%). En menor proporción estaban la familia *Anaerolineaceae* (7.2%), una familia del filo AC1 (3.6%) y la familia *Desulfobacteraceae* (3.3%).

A partir de los 20 cm de profundidad se presentó un cambio en la comunidad microbiana. Desde los 20 a los 60 cm dominó una familia no cultivada del filo *Hadesarchaeota* (19%), una familia desconocida del filo *Bathyarchaeota* (7.7%), y se observó un aumento en la abundancia relativa de tres familias del filo *Aerophobetes*, este incremento fue de aproximadamente el 8% para cada una de ellas.

En la sección de los 60 a los 80 cm, llama la atención el incremento de una familia de la clase *Ardenticatenia* (90.3%), que domina esta zona casi en su totalidad. A partir de los 80 cm vuelve a haber un incremento de una familia de *Hadesarchaea* (34%) y a partir de los 100 cm y hasta los 180 cm las familias más abundantes fueron aquellas de los filos *Hadesarchaea* (21-35.9%), seguido por *Bathyarchaeota* (2.4-16.9%) y *Aerophobetes* (3.4–13%).

En la sección de 180 cm predominaron una familia de la clase *20a-9* perteneciente al filo *WSA2* (13.5%), una familia no cultivada del filo *Hadesarchaea* (21.9%), la familia *Pseudomonadaceae* (13.1%), y la familia *Fusobacteriaceae* (3.6%). A partir de los 180 cm y hasta los 200 cm de profundidad se observó un incremento en la abundancia de una familia del filo *Bathyarchaeota* (pasando de un 5.5% a 180 cm hasta 46.7% a los 200cm de profundidad). Y por último, en la última sección de este núcleo, las familias *Bacilliaceae* (24.7%), una familia no cultivada del filo *Latescibacteria* (24%) la familia *Corynebacteriaceae* (12%), una familia no cultivada del filo *Hadesarchaea* (10.2%) y la familia *Comamonadaceae* (7.6%) fueron las más abundantes.

En el núcleo caliente (E2N3) lado derecho de la Figura 13, en la sección superficial dominaron las familias *Anaerolinaceae* (10.58%), *Desulfobacteraceae* (5.9%), *Desulfarculaceae* (3.5%), *JTB255* (3.3%) y las familias no cultivadas del filo *Aminicenantes* y *Latescibacteria* en un 5.3 y 2.9% respectivamente. El resto de las familias, un 47%, se encontraron por debajo del 1%. En las siguientes dos secciones (40 y 60 cm) se observó una alta abundancia de las familias *Corynebacteriaceae* (56.6-57%), y *Flavobacteriaceae* (11.9-17.7%), así como de las familias *Prevotellaceae* (4.1-8.3%) y *Porphyromonadaceae* (3.5-5.8%).

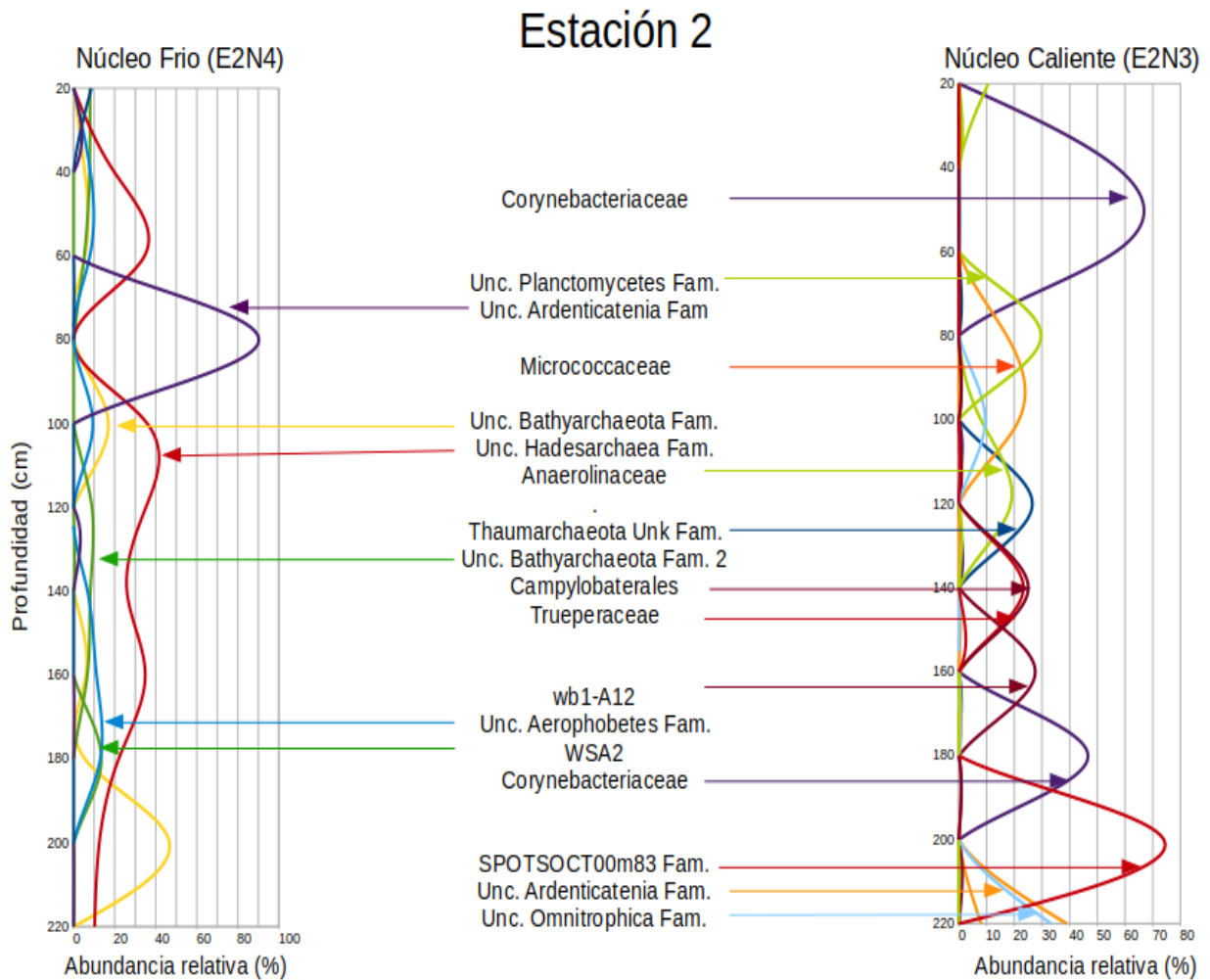
Llamó la presencia de la familia *Corynebacteriaceae* que se cree puede ser una contaminación, ya que es un saprófito humano que se encuentra en la piel.

En la sección que se encuentra a 100cm se presentó un aumento en la proporción de la familia *Micrococcaceae* (21.7%), así como de una familia no cultivada del filo *Latescibacteria* (6.1%). También se presentó una familia del filo *Omnitrophica* (9.4%) y se volvió a observar la familia *Anaerolineaceae* (7.9%) que se encontraba en la sección superficial.

A los 120 cm de profundidad destacó el aumento en la abundancia de la familia *Anaerolineaceae* (18.7%) y la gran abundancia de una familia desconocida del filo *Thaumarchaeota* (26%), así como de una familia no cultivada de la clase *HTA4* (13%) y una familia del subgrupo 17 del filo *Acidobacteria* (11.4%).

En la sección que se encuentra a 140 cm de profundidad se encontraba dominada por las familias *Helicobacteraceae* (25%) y *Trueperaceae* (23.2%), en menor familias no cultivadas de los filios *Planctomycetes* (13.3%) y *Thaumarchaeota* (11.2%) así como las familias *Caulobacteraceae* (9.2%) y *Cenarchaeaceae* (8.6%).

En las secciones más profundas, destacó a los 160 cm el aumento en una familia del filo *Planctomycetes* (23.6%), en mayor proporción se encontraba la familia *wb1-A12* (27.5%), así como las familias *Bacteroidaceae* (17.2%), *Flavobacteriaceae* (11.9%) y *Comamonadaceae* (5.9%). A los 180 cm dominaban las familias *Corynebacteriaceae* (46.6%), *Prevotellaceae* (28%), *Flavobacteriaceae* (10.8%) y *Veillonellaceae* (3%). En la siguiente sección, se observó de una familia no cultivada de la clase *SPOTSOCT00m83* (74%), así como, en menor presencia, las familias *Flavobacteriaceae* (11.6%), *Micrococcaceae* (4.1%), *Desulfobacteraceae* (2.3%) y *Anaerolineaceae* (1.9%). Y en la última sección se encontró que las familias más dominantes fueron las de los filios *Ardenticatenia* y *Omnitrophica* (38.9 y 33.2%), así como de las familias *Xanthomonadaceae* (13.4%) y *Micrococcaceae* (8%) en menor proporción..

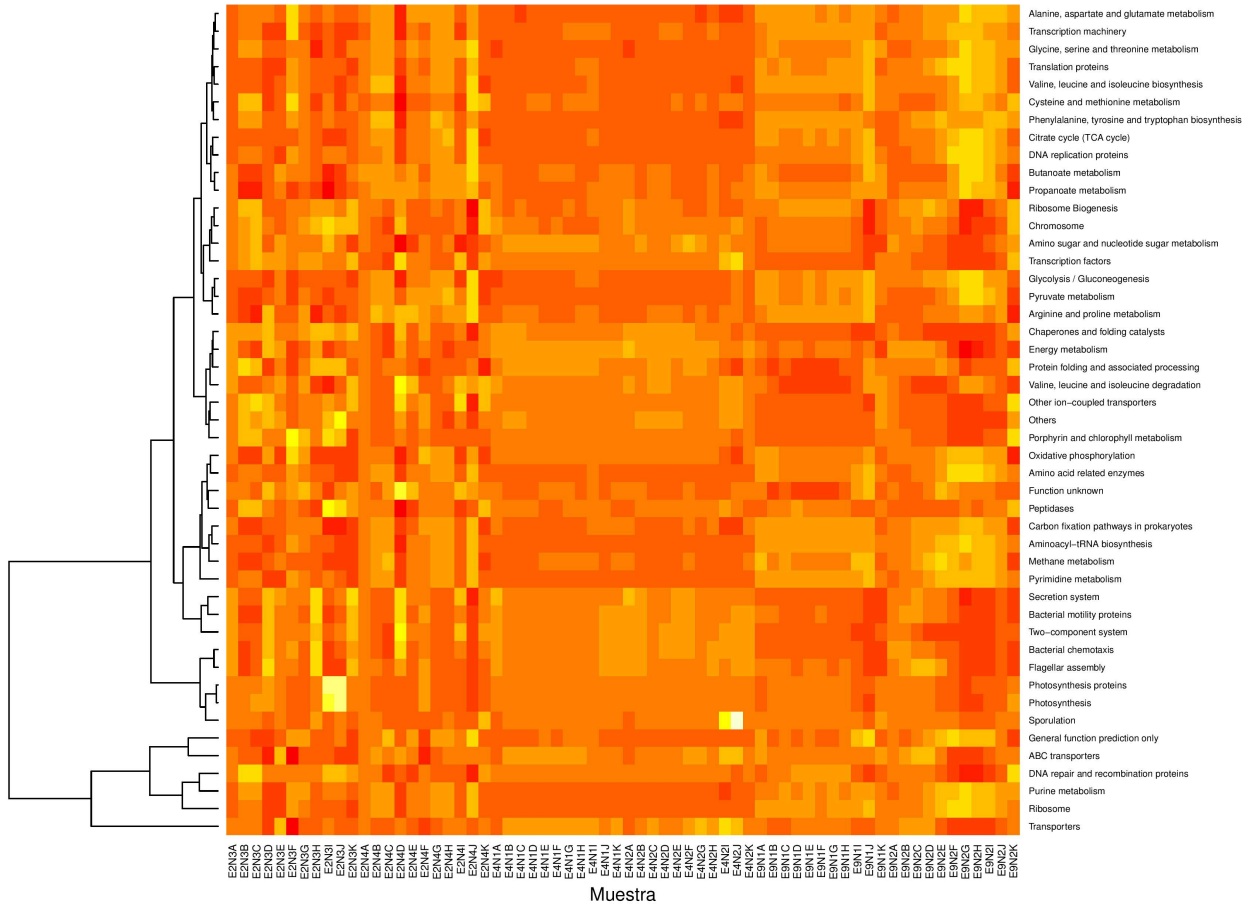


**Figura 18** Gráfica de líneas que muestra el perfil de profundidad de las familias más abundantes de los núcleos de la estación 2. A la izquierda se muestra el perfil del núcleo frío y a la derecha el del núcleo caliente.

### 3.8. Predicción funcional utilizando herramientas bioinformáticas

Se realizó una predicción de las funciones que se podrían estar llevando cabo por los microorganismos que se encuentran en los diferentes estratos. Para ello se utilizaron los programas PICRUSt y el Tax4Fun

Predicción funcional utilizando PICRUSt y Tax4Fun. Los resultados fueron muy parecidos en ambos programas como se puede ver en los *heatmaps* de la Figura 19 que muestra los resultados obtenidos con el PICRUSt y la Figura 20, que se construyó usando el programa Tax4Fun. La estación 4 se ve que ambos núcleos apenas presentan diferencias tanto entre núcleos como por profundidad. En cambio, el núcleo 2 es más diverso en término de sus funciones al igual que el 9. Sin embargo, el núcleo 9 presenta resultados más variables usando el programa Tax4Fun que cuando se usó el PICRUSt.



**Figura 19** Heatmap de las funciones predichas por el programa PICRUSt

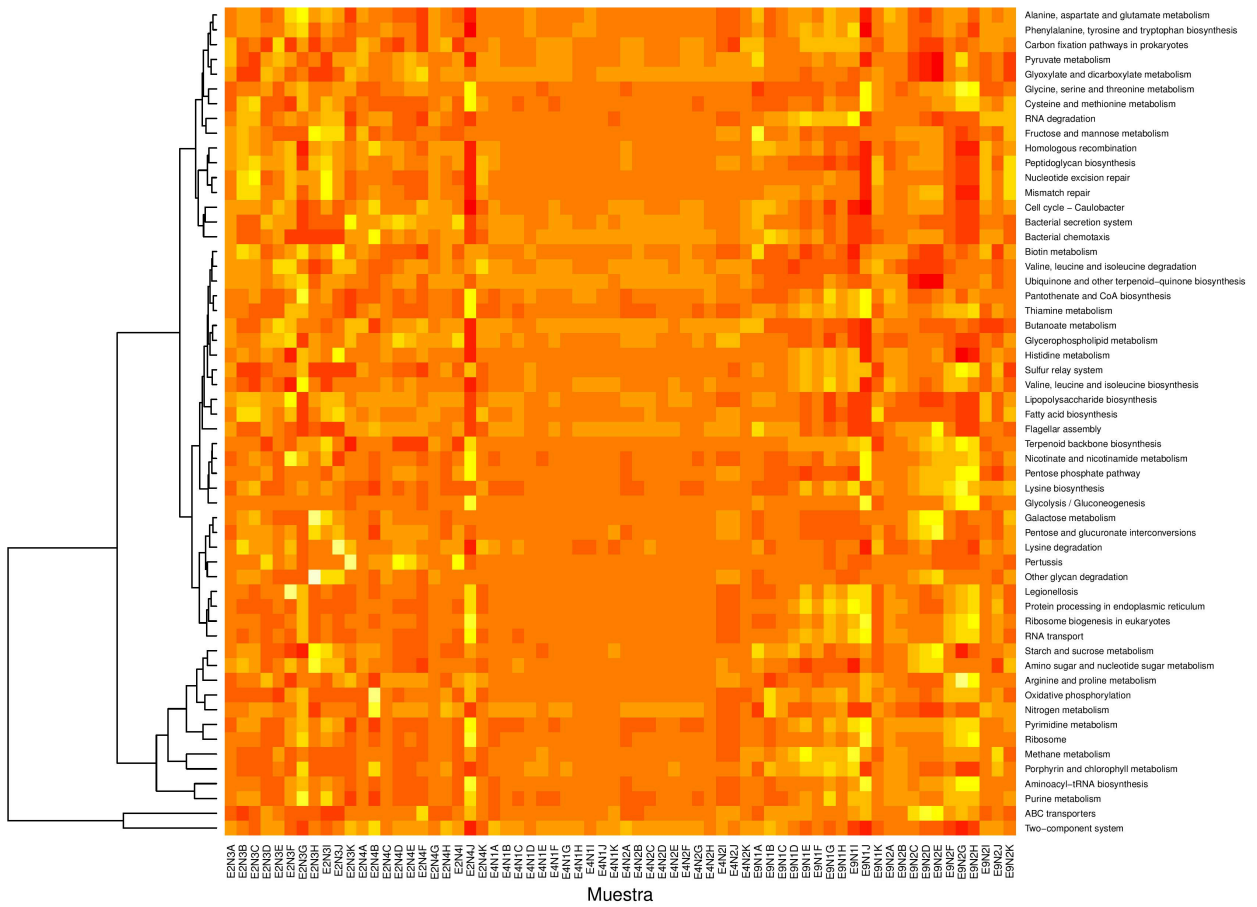


Figura 20. Heatmap de las funciones más abundantes predichas por Tax4Fun



## Capítulo 4. Discusión

---

El presente trabajo constituye una contribución al estudio de comunidades de organismos procarioes en zonas marinas de actividad geotermal, describiendo por primera vez la microbiota procariota de sedimentos procedentes de la Falla Wagner, mediante el análisis de sus metagenomas. Hasta la fecha solo se había caracterizando la microbiota asociada a sedimentos en el golfo de California en la cuenca de Guaymas (Dhillon *et al.*, 2003 y 2005; Meyer *et al.*, 2013). Sin embargo, en la falla Wagner no existen trabajos publicados que describan la diversidad microbiana que habita en esta zona. Con este fin, se seleccionaron 6 núcleos de 220 metros de longitud, procedentes de tres estaciones de muestreo: E4, E2, E9. Dos de las estaciones, las procedentes de la zona de la falla, registraron diferentes temperaturas que oscilaron entre los 25 y 42 °C para el núcleo E2N4, entre 50 y 99 °C para el E2N3, entre 24 y 37 °C para el E9N1 y entre 35 y 83 °C para el E9N2. Y para los núcleos de la estación E4, procedente del centro de la cuenca y usada como referencia externa, las temperaturas oscilaron entre los 14 y 16 °C para ambos núcleos. Los núcleos se dividieron en secciones de 20 cm para estudiar los posibles cambios en la comunidad a lo largo de la columna sedimentaria.

La extracción de ADN mostró rendimientos muy variables. Se observaron rendimientos que van desde los 0.1 µg/g hasta los 3 µg/g, lo que son rendimientos bastante bajos, aunque esto no es extraño y ya se han reportado rendimientos similares por algunos autores utilizando diferentes métodos de extracción de ácidos nucleicos en diferentes sedimentos (Alain *et al.*, 2011). Sin embargo son muy inferiores a lo reportado por otros estudios donde se utilizan métodos tradicionales (Smalla, *et al.* 1993) pero utilizando muestras de suelos y no de sedimentos marinos. Los bajos rendimientos tuvieron efecto en las amplificaciones, sobre todo en aquellas que estaban en los últimos 100 cm de la E2N3. Las dificultades en la extracción posiblemente se debieron a la baja densidad de microorganismos capaces de vivir a más de 90 °C, que era la temperatura media de esa zona del núcleo.

Otro de los problemas a los que se pudo deber la dificultad en la obtención de material genético procedente de las muestras de sedimentos estudiados en este trabajo, es la posible coextracción de ácidos húmicos impidiendo una buena amplificación en algunas de las muestras. La existencia de ácidos húmicos en los sedimentos marinos ya ha sido bastante reportada en otros trabajos como factor causante de esta inhibición (Tsai y Olson, 1992; Smalla *et al.*, 1993). Aunque se utilizaron diferentes métodos para sobrellevar la presencia de inhibidores de la PCR, no fue posible obtener fragmentos de

calidad aceptable en algunas muestras, sobre todo en aquellas muestras que presentaban las temperaturas más altas.

Sin embargo, y a pesar de las dificultades técnicas, se obtuvo un buen número de secuencias (2 261 113), que permitieron llevar a cabo el trabajo de manera exitosa, descartándose únicamente 560 512 secuencias, que no cumplían con los criterios de calidad necesarios para incluirlas en los análisis.

#### 4.1 Relaciones observadas entre las comunidades

Los análisis de Coordenadas Principales (PCoAs) mostraron que las secciones superficiales de los núcleos fueron muy similares entre ellas. Lo mismo ocurrió con las muestras de toda la columna sedimentaria de ambos núcleos de la estación 4. Las muestras más superficiales se encontraron dominadas en su mayoría por OTUs de las familias *JTB255-MBG* y *Anaerolineaceae*. Esta semejanza entre las muestras superficiales entre todas las estaciones, podría deberse a que estamos estudiando un área relativamente reducida en el golfo de California, donde se ha observado la presencia de corrientes submarinas (Carriquiry *et al.*, 2001). Estas corrientes podrían estar provocando que la capa superficial del sedimento fuese más dinámica y ayude a mantener una mayor diversidad microbiana (Galand *et al.*, 2016) al tiempo que hace que esta sección sea más homogénea. Las pocas diferencias encontradas en algunas de las muestras superficiales pueden deberse a factores abióticos como la temperatura superficial del sedimento o a la diferencia en la batimetría de la región.

De igual manera, las similitudes encontradas en las muestras de la estación 4, las cuales proceden del centro de la cuenca, podrían explicarse por la alta tasa de sedimentación de la zona, lo cual no estaría permitiendo que se dé una estratificación del sedimento que propicie una zonación marcada en las comunidades microbianas. Sin embargo, una de las cosas que más llama la atención en esta estación, es que el núcleo 2 (E4N2) presenta dos secciones con una comunidad muy diferente al resto del perfil. Este cambio se presenta entre los 160 y 200 cm de profundidad, donde abundan los OTUs de la familia *Bacillaceae*, compuesta por 3 géneros principales *Anoxybacillus*, *Anaerobacillus*, y *Bacillus*, siendo este último el más abundante (80%). Los miembros de estos géneros se caracterizan por que son capaces de formar endosporas (Mandic-Mulec *et al.*, 2015). En particular, se han estudiado varias especies del género *Bacillus* en diversos ambientes, incluyendo los sedimentos marinos, por su capacidad de formación de endosporas y por su potencial biotecnológico (Ivanova *et al.*, 1999; Miranda *et al.*, 2007;

Ettoumi *et al.*, 2009). Otros miembros de este género también han sido encontrados en ambientes geotermales presentando capacidad de oxidación de Mn(II), sin embargo, por norma general, se encuentran en estado vegetativo cuando se encuentran en temperaturas fuera del su rango de tolerancia (Dick *et al.*, 2006). Por su parte, los miembros de la familia *Anoxybacillus* generalmente son termófilos o termófilos moderados y toleran un rango de temperatura que va desde los 50 hasta los 65 °C, muy pocas cepas son mesófilas, y suelen ser anaerobios estrictos, aunque también existen cepas aerobias, así como anaerobias facultativas (Poli *et al.*, 2006; Pikuta *et al.*, 2009; Goh *et al.*, 2013), así que no está claro cual es el factor que promueve que sean tan abundantes en esta sección. Este es, por tanto, el primer trabajo que describe una alta abundancia de este género a temperaturas tan bajas, sin embargo, es posible que hayan sido acarreados en un fenómeno de erupción submarina en las zonas de alta temperatura hace aproximadamente 40 años, que coincidiría con los sismos registrados en 1969 cuyas magnitudes rondaron entre los 4 y 5.5 en la escala Richter (Thatcher and Brune, 1971; Alvarez *et al.*, 2009).

Las agrupaciones observadas en las muestras que se encuentran a partir de los 40 cm bsm de la E9N1, nos indican que en particular este núcleo es muy distinto en composición microbiana de los otros analizados, lo cual no se explica por la temperatura superficial del sedimento. Posiblemente podría tratarse de una zonación específica en este punto, ya que se comparten OTUs de familias de los filos *Aerophobetes*, *Aminicenantes*, *AC1*, *Hadesarchaeota* y *Bathyarchaeota*. No obstante, cuando se realiza el análisis tomando en cuenta las abundancias relativas solo aquellas muestras más profundas empiezan a separarse. Esto concuerda con lo observado en los análisis taxonómicos donde empieza a aumentar las abundancias de ciertos grupos a medida que aumenta la profundidad en la columna sedimentaria. En el núcleo caliente de la estación 9 (E9N2) se observó que en ambos análisis cada una de las secciones se va alejando de la anterior, lo que nos indicaría que, debido a que este fue uno de los núcleos que presentaba un alto gradiente de temperatura, se está manifestando una sucesión microbiana probablemente dirigida por la temperatura.

Se observó una tercera agrupación de las muestras E9N2P140, E9N2P160, E9N1P200 y E2N4P200, las primeras dos presentan temperaturas de 80 °C, las otras dos presentan temperaturas que rondan entre los 40 °C. Estas se encuentran dominadas en su mayoría por OTUs pertenecientes a familias de los filos *Bathyarchaeota*, *AK8*, *Hadesarchaea* y *Acetothermia*. El segundo grupo se caracteriza, además de los otros grupos, por presentar también una alta abundancia de los miembros del filo *Aigarchaeota*. Estas agrupaciones podrían deberse a que son muestras muy profundas, lo que concuerda con lo que se ha

observado en otros estudios, donde encuentra la presencia de estos grupos, principalmente, a altas profundidades en el subsuelo marino (Yu *et al.*, 2018; Baker *et al.*, 2016)

## 4.2 Análisis de las principales familias

Los resultados del análisis de las comunidades a nivel familia mostraron que hay un grupo de familias que se presentaron de manera abundante en determinadas secciones o estaciones. Una de estas familias es la *JTB255-MBG*, que es una familia ubicua de los sedimentos marinos superficiales, tanto en sedimentos someros como en sedimentos profundos (Bowman *et al.*, 2005; Bienhold *et al.*, 2016; Dyksma *et al.*, 2016; Mußmann *et al.*, 2017). En algunos estudios se ha visto que miembros de esta familia son los principales fijadores de carbono orgánico en sedimentos superficiales (Dyksma, 2016).

Otra de las principales familias encontradas en los núcleos analizados fue la *Anaerolineaceae*. La familia *Anaerolinaceae* está compuesta por bacterias anaerobias obligadas y ha sido ampliamente estudiado su potencial papel en la degradación anaerobia de hidrocarburos de cadena larga y su conversión a metano. Miembros de esta familia se han encontrado en diversos ambientes tanto en sedimentos marinos, ventilas hidrotermales, suelos y emanaciones de hidrocarburos terrestres, pero principalmente se encuentra en zonas caracterizadas por su anaerobiosis y ricas en hidrocarburos (Elshahed *et al.*, 2003; Savage *et al.*, 2010; Nunoura *et al.*, 2013; Sutton *et al.*, 2013; Liang *et al.*, 2015).

Los miembros del filo *Aerophobetes* han sido observados anteriormente en sedimentos profundos, generalmente en áreas relacionadas con emanaciones de metano y ambientes anóxicos (Wang *et al.*, 2016). En nuestras muestras se encontraron en mayor abundancia entre los 40 y 200 cm de profundidad en los núcleos cercanos a la falla Wagner E9N1, y E9N2, lo que podría indicar que los sedimentos los bajos niveles de oxígeno en estos puntos.

La distribución de miembros del filo *Lokiarchaeota* o *Deep Sea Marine Group* a lo largo de la columna sedimentaria de los núcleos de la estación 4 concuerda con lo que se ha estudiado sobre la ubiquidad de miembros de este filo en los sedimentos marinos, y se ha observado que estos pueden estar correlacionados con la presencia de materia orgánica (Jørgensen *et al.*, 2013; Spang *et al.*, 2015).

Otro cambio observado es el aumento abrupto en la abundancia de la familia *ANME-2a-2b* desde los 80 hasta los 120cm de profundidad en el primer núcleo de la estación 4, así como de la familia *Desulfulobacteraceae* (9.6 y 4% respectivamente). Esto corresponde con lo que se ha estudiado sobre la

sintrofia entre metanotrofos y bacterias reductoras de sulfato (Boetius *et al.*, 2000; Heijs 2007), y nos podría indicar la presencia de metano en esta sección.

Las muestras del núcleo caliente de la estación 2 que se encontraban a 40, 60 y 180 cm de profundidad, así como las muestras de las secciones más profundas de los núcleos de la estación 9 presentaron problemas en la amplificación. Algunos de las familias encontradas en estas muestras de sedimento analizadas, como *Corynebacteriaceae* y *Prevotellaceae* han sido encontradas en el microbioma humano, son organismos saprofitos de la piel (Downes *et al.*, 2008; Bernard K., 2012), lo que sugiere que en este caso puede ser una contaminación en la muestra a lo largo de la cadena de procesamiento, por lo que fueron eliminadas de los análisis de componentes principales para evitar que nos afectaran las agrupaciones observadas. En el futuro convendrá repetir las, con especial cuidado para evitar una posible contaminación.

Finalmente, la co-ocurrencia, en distintas secciones, de miembros de los filos *Hadesarchaea* y *Bathyarchaea* podría darse por relaciones de comensalismo por parte de estos últimos, donde los productos metabólicos de *Hadesarchaea* son utilizados para potenciar el metabolismo de los miembros del filo *Bathyarchaeota*. Estudios recientes subrayan la capacidad de miembros del filo *Hadesarchaea* de fijar carbono inorgánico, lo cual resulta útil en los sedimentos profundos donde escasean las fuentes de carbono orgánico, este proceso se lleva a cabo mediante la oxidación del monóxido de carbono, (Sokolova *et al.*, 2009; Baker *et al.*, 2016), mientras que los miembros del filo *Bathyarchaeota* son capaces de metabolizar el carbono recalcitrante, proveniente de la lignina, para sobrellevar la escasez de carbono orgánico en sedimentos profundos (Yu *et al.*, 2018). Los miembros de este filo generalmente se encuentran en las secciones más profundas del sedimento marino. Es posible que *Bathyarchaeota* este incorporando el CO<sub>2</sub> producido por miembros del filo *Hadesarchaeota* promoviendo así el metabolismo acetogénico (Yu *et al.*, 2018).

## Capítulo 5. Conclusiones

---

Las comunidades microbianas procariotas presentes en los primeros 20 cm de los núcleos analizados presentan similitudes entre todas las estaciones de muestreo. Esta sección en todas las estaciones están dominadas en su mayoría por los filos *Proteobacteria*, *Chloroflexi*, *Acidobacteria* y *Planctomycetes*.

Todas las secciones de ambos núcleos de la E4 presenta la misma estructura, por lo que la profundidad no parece ser un factor determinante. Estas muestras se agrupan tanto en los análisis de presencia-ausencia como en los que se tiene en cuenta la abundancia de algunos taxones.

Las secciones subsuperficiales, a partir de los 40 cm, entre la estación de referencia y las estaciones que se encuentran cercanas a la falla Wagner son diferentes en su composición microbiana procariota.

Los perfiles de profundidad entre núcleos fríos y núcleos calientes presentan claras diferencias tanto en la composición como en la abundancia de la comunidad procariota, a lo largo de la columna sedimentaria.

La profundidad parece tener efecto sobre la estructura de la comunidad aunque el factor más determinante parece ser la temperatura, en los núcleos estudiados.

Los miembros del dominio *Archaea* fueron más dominantes en secciones más profundas de los núcleos de las estaciones cercanas a la falla Wagner, exceptuando el núcleo caliente de la estación 2 donde se presentó la temperatura más alta.

## Literatura citada

---

- Alain K., Callacc N., Ciobanub M.-C., Reynaudb Y., Duthoitb F., Jebbara M. 2011. DNA extractions from deep subseafloor sediments: Novel cryogenic-mill-based procedure and comparison to existing protocols. *Journal of Microbiological Methods* 87(3): 355-362.
- Alvarez L.G., Suárez-Vidal F., Mendoza-Borunda R., González-Escobar M. 2009. Bathymetry and active geological structures in the Upper Gulf of California. *Boletín de la Sociedad Geológica Mexicana* 61(1):129-141.
- Anantharaman, K., Breier, J., Dick, G. 2016. Metagenomic resolution of microbial functions in deep-sea hydrothermal plumes across the Eastern Lau Spreading Center. *The ISME Journal* 10, 225-239. doi:10.1038/ismej.2015.81.
- Aragón-Arreola, M., Morandi, M., Martin-Barajas, A., Delgado-Argote, L., González-Fernández, A. 2005. Structure of the rift basins in the central Gulf of California: Kinematic implications for oblique rifting. *Tectonophysics* 409: 19-38.
- Aßhauer K.P., Wemheuer B., Daniel R., Meinicke P. 2015. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics*. 31(17):2882-4.
- Badhai J., Ghosh T.S., Das S.K. 2015. Taxonomic and functional characteristics of microbial communities and their correlation with physicochemical properties of four geothermal springs in Odisha, India. *Frontiers in Microbiology*. 26(6):1166.
- Baker B.J., Saw J.H., Lind A.E., Lazar C.S., Hinrichs K.-U., Teske A.P. & Ettema T.J.G. 2016. Genomic inference of the metabolism of cosmopolitan subsurface *Archaea*, *Hadesarchaea*. *Nature Microbiology* DOI: 10.1038/NMICROBIOL.2016.2
- Bazylinski, D., Wirsén, C., Jannasch, H. 1989. Microbial utilization of naturally occurring hydrocarbons at the Guaymas Basin hydrothermal vent site. *Applied and Environmental Microbiology* 55(11): 2831-2836.
- Bernard K. 2012. The Genus *Corynebacterium* and other medically relevant coryneform-like bacteria. *Journal of Clinical Microbiology* 50(10):3152-3158.
- Bienhold C., Zinger L., Boetius A., Ramette A. 2016. Diversity and Biogeography of Bathyal and Abyssal Seafloor Bacteria. *PLoS One*. 11(1): e0148016. doi: 10.1371/journal.pone.0148016.
- Boetius A., Ravensschlag K., Schubert C.J., Rickert D., Widdel F., Gieseke A., Amann R., Jørgensen B.B., Witte U., Pfannkuche O. 2000. A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* 407(6804):623-6.
- Boudreau, B. 1997. Diagenetic models and their implementation: Modelling transport and reactions in aquatic sediments. Springer-Verlag, Berlin, Germany.
- Bowman J.P., McCammon S.A., Dann A.L. 2005. Biogeographic and quantitative analyses of abundant uncultivated *gamma-proteobacterial* clades from marine sediment. *Microbial Ecology* 49: 451-460.
- Caporaso J.G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.D., Costello E.K., Fierer N., Peña A.G., Goodrich J.K., Gordon J.I., Huttley G.A., Kelley S.T., Knights D., Koenig J.E., Ley R.E., Lozupone C.A., McDonald D., Muegge B.D., Pirrung M., Reeder J., Sevinsky J.R., Turnbaugh P.J., Walters W.A., Widmann J., Yatsunenko T., Zaneveld J., Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7(5):335-6.

- Carriquiry J.D., Sánchez A., Camacho-Ibar V.F. 2001. Sedimentation in the northern Gulf of California after cessation of the Colorado River discharge. *Sedimentary Geology*. 114(1-2): 37-62.
- Covarrubias, M. 2016. Caracterización de las comunidades microbianas en sedimentos del Golfo de México, mediante análisis metagenómico. Tesis de Maestría en Ciencias. Centro de investigación Científica y de Educación Superior de Ensenada, Baja California. 82 pp.
- Dhillon A., Lever M., Lloyd K.G., Albert D.B., Sogin M.L., Teske A. 2005. Methanogen diversity evidenced by molecular characterization of methyl coenzyme M reductase A (*mcrA*) genes in hydrothermal sediments of the Guaymas Basin. *Applied and Environmental Microbiology*. 71(8):4592-601.
- Dhillon A., Teske A., Dillon J., Stahl D.A., Sogin M.L. 2003. Molecular characterization of sulfate-reducing bacteria in the Guaymas Basin. *Applied and Environmental Microbiology*. 69(5):2765-72.
- D'Hondt, S., Jørgensen, B., Miller, D., Batzke, A., Blake, R., Cragg, B., *et al.* 2004. Distributions of microbial activities in deep sub seafloor sediments. *Science* 306: 2216–2.
- Dick G.J., Lee Y.E., Tebo B.M. 2006. Manganese(II)-oxidizing *Bacillus* spores in Guaymas Basin hydrothermal sediments and plumes. *Applied and Environmental Microbiology*. 72(5):3184-90.
- Downes J., Hooper S.J., Wilson M.J., Wade W.G. 2008. *Prevotella histicola* sp. nov., isolated from the human oral cavity. *International Journal of Systematic and Evolutionary Microbiology*. 58(Pt 8):1788-91.
- Dyksma S., Bischof K., Fuchs B.M., Hoffmann K., Meier D., Meyerdierks A., Pjevac P., Probandt D., Richter M., Stepanauskas R., Mußmann M. Ubiquitous *Gammaproteobacteria* dominate dark carbon fixation in coastal sediments 2016. *ISME Journal* 10(8):1939-53.
- Edgar R.C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19):2460-2461.
- Edgcomb, V., Kysela, D., Teske, A., Gomez, A., y Sogin, M. 2002. Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *Proceedings of the National Academy of Sciences of the United States of America* 99(11): 7658-7662.
- Elshahed M.S., Senko J.M., Najar F.Z., Kenton S.M., Roe B.A., Dewers T.A., Spear J.R., Krumholz L.R. 2003. Bacterial diversity and sulfur cycling in a mesophilic sulfide-rich spring. *Applied Environmental Microbiology*. 69(9):5609-21.
- Ettoumi B., Raddadi N., Borin S., Daffonchio D., Boudabous A., Cherif A. 2009 Diversity and phylogeny of culturable spore-forming *Bacilli* isolated from marine sediments. *Journal of Basic Microbiology* 49, S13 – S23.
- Galand P.E., Lucas S., Fagervold S.K., Peru E., Pruski A.M., Vétion G., Dupuy C., Guizien K. 2016. Disturbance Increases Microbial Community Diversity and Production in Marine Sediments. *Frontiers in Microbiology*. 7: 1950.
- Goh K.M., Kahar Ummirul Mukminin, Chai Y.Y., Chong C.S., Chai K.P., Ranjani V., Illias R.M., Chan K.-G. 2013. Recent discoveries and applications of *Anoxybacillus*. *Applied Microbiology and Biotechnology* 97(4):1475–1488.
- Grijalva, N., 1986. Investigación de la Energía Geotérmica en la Depresión de Wagner en el Golfo de California. Latitud 31° 00' al 31° 15' y Longitud 113° 50' al 114° 10'. Reportes no publicados preparados para la Comisión Federal de Electricidad, Cerro Prieto, BC, bajo Contrato No. CCP-CLS-002/86. Fecha del primer Reporte: 10/1/1986, (22 pp). Fecha del segundo Reporte: 11/22/1986, (38pp). Inédito.



- Heijs S.K., Haese R.R., van der Wielen P.W., Forney L.J., van Elsas J.D. 2007. Use of 16S rRNA gene based clone libraries to assess microbial communities potentially involved in anaerobic methane oxidation in a Mediterranean cold seep. *Microbial Ecology* 53(3):384-98.
- Ivanova E.P., Vysotskii M.V., Svetashev V.I., Nedashkovskaya O.I., Gorshkova N.M., Mikhailov V.V., Yumoto N., Shigeri Y., Taguchi T., Yoshikawa S. 1999. Characterization of *Bacillus* strains of marine origin. *International Microbiology*. 2(4):267-71.
- Jørgensen B., D'Hondt S. 2006. A starving majority deep beneath the seafloor. *Science*, 314: 932–934.
- Jørgensen S.L., Thorseth I.H., Pedersen R.B., Baumberger T., Schleper C. 2013. Quantitative and phylogenetic study of the Deep Sea Archaeal Group in sediments of the Arctic mid-ocean spreading ridge. *Frontiers in Microbiology* 4:299.
- Karl, D. 1995. Ecology of free-living, hydrothermal vent microbial communities. In: Karl, D.M. (Ed.), *The Microbiology of Deep-Sea Hydrothermal Vents*. CRC Press, Boca Raton, FL, pp. 35-124.
- Karl, D., Taylor, G., Novitsky, J., Jannash, H., Wirsén, C., Pace, N., Lane, D., Olsen, G., y Giovannoni, S. 1988. A microbiological study of Guaymas Basin high temperature hydrothermal vents. *Deep Sea Research Part A. Oceanographic Research Papers*, 35(5): 777-791.
- Kozich, J., Wescott, S., Baxter, N., Highlander, S., Schloss, P. 2013. Development of a Dual-Index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, 79(17): 5112-5120.
- Langille M.G.I., Zaneveld J., Caporaso J.G., McDonald D., Knights D., Reyes J.A., Clemente J.C., Burkepille D.E., Vega T.R.L., Knight R., Beiko R.G., Huttenhower C. 2014. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*. *Nature Biotechnology* 31(9): 814–821.
- Liang B., Wang L.-Y., Mbadinga S. M., Liu J.-F., Yang S.-Z., Gu J.-D., and Mu B.-Z. 2015. *Anaerolineaceae* and *Methanosaeta* turned to be the dominant microorganisms in alkanes-dependent methanogenic culture after long-term of incubation. *AMB Express* 5: 37. doi: 10.1186/s13568-015-0117-4.
- Lonsdale, P. 1989. Geology and tectonic history of the Gulf of California. En: Winterer E.L., Hussong D.M., Decker R.W. (Eds). *The Geology of North America, The Eastern Pacific Ocean and Hawaii*. Geological Society of America, Boulder, Colorado. pp. 499-521.
- Lozupone C., Knight R. 2005. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology* 71(12):8228–8235.
- Lozupone C., Lladser M.E., Knights D., Stombaugh J., Knight R. 2011. UniFrac: an effective distance metric for microbial community comparison. *ISME Journal* 5(2):169-72.
- Madigan, M., Martinko, J., Bender, K., Buckley, D., Stahl, D. 2016. *Brock biology of microorganisms*, 14 ed. Pearson, 1032 pp. ISBN: 978-0321897398.
- Mandic-Mulec I., Stefanic P., van Elsas J.D. 2015. Ecology of *Bacillaceae*. *Microbiol Spectrum* 3(1): TBS-0017-2013.
- McCollom, T. 2000. Geochemical constraints on primary productivity in submarine hydrothermal vent plumes. *Deep-sea research I* 47: 85-101. doi: 10.1016/S0967-0637(99)00048-5.
- Meyer S., Wegener G., Lloyd K.G., Teske A., Boetius A., Ramette A. 2013. Microbial habitat connectivity across spatial scales and hydrothermal temperature gradients at Guaymas Basin. *Frontiers in Microbiology*. 4: 207. doi: 10.3389/fmicb.2013.00207.

- Miranda C.A., Martins O.B., Clementino M.M. 2008. Species-level identification of *Bacillus* strains isolates from marine sediments by conventional biochemical, 16S rRNA gene sequencing and inter-tRNA gene sequence lengths analysis. *Antonie van Leeuwenhoek*. 93(3):297-304.
- Mußmann M., Pjevac P., Krüger K. & Dykma S. 2017. Genomic repertoire of the *Woeseiaceae/JTB255*, cosmopolitan and abundant core members of microbial communities in marine sediments. *The ISME Journal* 11(15): 1276–1281.
- Nunoura T., Oida H., Nakaseama M., Kosaka A., Ohkubo S.B., Kikuchi T., Kazama H., Hosoi-Tanabe S., Nakamura K., Kinoshita M., Hirayama H., Inagaki F., Tsunogai U., Ishibashi J., Takai K. 2010. Archaeal diversity and distribution along thermal and geochemical gradients in hydrothermal sediments at the Yonaguni Knoll IV hydrothermal field in the Southern Okinawa trough. *Applied and Environmental Microbiology* 76(4):1198-211.
- Orcutt, B., Sylvan, J., Knab, N. y Edwards, K. 2011. Microbial ecology of the dark ocean above, at, and below the seafloor. *Microbiology and molecular biology reviews* 75(2):361-422.
- Persuad, P., Stock, J., Steckler, M., Martin-Barajas, A., Diebold, J., Gonzalez-Fernandez, A., Mountain, G. 2003. Active deformation and shallow structure of the Wagner, Consag and Delfin Basins, northern Gulf of California, Mexico. *Journal of geophysical research* 108(B7)2355.
- Pikuta E.V. 2009. Genus IV. Anoxybacillus En: Vos P., Garrity G., Jones D., Krieg N.R., Ludwig W., Rainey F.A., Schleifer K.H., Whitman W.B. (Eds.) *Bergey's Manual of systematic bacteriology* Vol. 3: The *Firmicutes*. Springer Science & Business Media. New York. 134-141
- Poli A., Esposito E., Lama L., Orlando P., Nicolaus G., de Appolonia F., Gambacorta A., Nicolaus B. 2006. *Anoxybacillus amylolyticus* sp. nov., a thermophilic amylase producing bacterium isolated from Mount Rittmann (Antarctica). *Systematic and Applied Microbiology* 29(4):300-7.
- Prol-Ledesma, R., Torres-Vera, M., Rodolfo-Metalpa, R., Angeles, C., Lechuga-Deveze, C., Villanueva-Estrada, R., Shumilin, E., Robinson, C. 2013. High heat flow and ocean acidification at a nascent rift in the northern Gulf of California. *Nature Communications* 4:1388 10.1038/ncomms2390.
- Pérez-Villatoro F. Dávila S., Prol-Ledesma R.M., Juárez K. 2017. Archaeal communities at gas-venting shallow basins in the northern Gulf of California (Wagner and Consag Basins). En: *Goldschmidt Conference 2017 Paris, Francia, 13-18 de agosto*.
- Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., Peplies J., Glöckner F.O. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*. doi: 10.1093/nar/gks1219.
- Savage K.N., Krumholz L.R., Gieg L.M., Parisi V.A., Suflita J.M., Allen J., Philp R.P., Elshahed M.S. 2010. Biodegradation of low-molecular-weight alkanes under mesophilic, sulfate-reducing conditions: metabolic intermediates and community patterns. *FEMS Microbiology Ecology* 72(3):485-95.
- Smalla K., Cresswell N., Mendonca-Hagler L.C., Wolters A., van Elsas J.D. 1993. Rapid DNA extraction protocol from soil for polymerase chain reaction-mediated amplification. *Journal of Applied Bacteriology* 74(1):78-85.
- Sokolova T.G., Henstra A.-M., Sipma J., Parshina S.N., Stams A.J.M., Lebedinsky A.V. 2009. Diversity and ecophysiological features of thermophilic carboxydrotrophic anaerobes. *FEMS Microbiology Ecology*. 68(2):131-41.

- Spang A., Saw J.H., Jørgensen S. L., Zaremba-Niedzwiedzka K., Martijn J., Lind A.E., van Eijk R., Schleper C., Guy L., and Ettema T.J. 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521(7551): 173–179.
- Sutton N.B., Maphosa F., Morillo J.A., Abu A.-S.W., Langenhoff A.A., Grotenhuis T., Rijnaarts H.H., Smidt H. 2013. Impact of long-term diesel contamination on soil microbial community structure. *Applied and Environmental Microbiology* 79(2):619-30.
- Teske, A., Hinrichs, K., Edgcomb, V., de vera, G., Kysela, D., Sylva, S., Sogin, M., Jannasch, H. 2002. Microbial Diversity of Hydrothermal Sediments in the Guaymas Basin: Evidence for Anaerobic Methanotrophic Communities. *Applied and Environmental Microbiology*, 68(4): 1994-2007. doi: 10.1128/AEM.68.4.1994–2007.2002.
- Thatcher W., Brune J.N. 1971. Seismic Study of an Oceanic Ridge Earthquake Swarm in the Gulf of California. *Geophysical Journal of the Royal Astronomical Society*. 22:473-489.
- Tsai Y.L., Olson B.H. 1992. Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. *Applied and Environmental Microbiology*. 58(7):2292-5.
- Wagner I.D., Wiegand J. 2008. Diversity of thermophilic anaerobes. *Annals of the New York Academy of Sciences*. 1125:1-43. doi: 10.1196/annals.1419.029.
- Wang Y., Gao Z.-M., Li J.-T., Bougouffa S., Tian R.M., Bajic V.B., Qian P.-Y. 2016. Draft genome of an *Aerophobetes* bacterium reveals a facultative lifestyle in deep-sea anaerobic sediments. *Science Bulletin* 61(15): 1176-1186.
- Yu T., Wu W., Liang W., Lever M. A., Hinrichs K.-U., Wang F. 2018. Growth of sedimentary *Bathyarchaeota* on lignin as an energy source. *Proceedings of the National Academy of Sciences of the United States of America* 115(23): 6022-6027