



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

**ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MORELIA
BIOLOGÍA EXPERIMENTAL**

**ESTUDIO METAGENÓMICO DE ARQUEAS TERMOACIDÓFILAS DE
SEDIMENTOS DEL CAMPO GEOTÉRMICO DE LOS AZUFRES**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN CIENCIAS BIOLÓGICAS

PRESENTA:

ROBERTO MARÍN PAREDES

TUTOR PRINCIPAL: DR. LUIS EDUARDO SERVÍN GARCIDUEÑAS
ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MORELIA. UNAM

COMITÉ TUTOR: DRA. YUNUEN TAPIA TORRES
ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MORELIA. UNAM
DRA. MARIA ESPERANZA MARTÍNEZ ROMERO
CENTRO DE CIENCIAS GENÓMICAS. UNAM

MORELIA, MICHOACÁN 2020
AGOSTO



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

**ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MORELIA
BIOLOGÍA EXPERIMENTAL**

**ESTUDIO METAGENÓMICO DE ARQUEAS TERMOACIDÓFILAS DE SEDIMENTOS DEL
CAMPO GEOTÉRMICO DE LOS AZUFRES**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN CIENCIAS BIOLÓGICAS

PRESENTA:

ROBERTO MARÍN PAREDES

**TUTOR PRINCIPAL: DR. LUIS EDUARDO SERVÍN GARCIDUEÑAS
ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MORELIA. UNAM**

**COMITÉ TUTOR: DRA. YUNUEN TAPIA TORRES
ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MORELIA. UNAM
DRA. MARIA ESPERANZA MARTÍNEZ ROMERO
CENTRO DE CIENCIAS GENÓMICAS. UNAM**

MORELIA, MICHOACÁN 2020

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS

ENTIDAD ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MORELIA

OFICIO CPCB/426/2020

ASUNTO: Oficio de Jurado

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

Me permito informar a usted que en la reunión ordinaria del Subcomité de Biología Experimental y Biomedicina del Posgrado en Ciencias Biológicas, celebrada el día **07 de octubre de 2019** se aprobó el siguiente jurado para el examen de grado de **MAESTRO EN CIENCIAS BIOLÓGICAS** en el campo de conocimiento de **Biología Experimental** del estudiante **MARÍN PAREDES ROBERTO** con número de cuenta **518002859** con la tesis titulada "**Estudio metagenómico de arqueas termoacidófilas de sedimentos del campo geotérmico de Los Azufres**", realizada bajo la dirección del **DR. LUIS EDUARDO SERVÍN GARCIDUEÑAS**, quedando integrado de la siguiente manera:

Presidente: DRA. LUISA ISaura FALCÓN ÁLVAREZ
Vocal: DRA. MAYRA ELENA GAVITO PARDO
Secretario: DRA. MARIA ESPERANZA MARTÍNEZ ROMERO
Suplente: DRA. ROCIO JETZABEL ALCÁNTARA HERNÁNDEZ
Suplente: DR. ANTONIO GONZÁLEZ RODRÍGUEZ

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

ATENTAMENTE
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Cd. Universitaria, Cd. Mx., a 06 de agosto de 2020

COORDINADOR DEL PROGRAMA



DR. ADOLFO GERARDO NAVARRO SIGÜENZA



Agradecimientos institucionales.

Al **Posgrado de Ciencias Biológicas UNAM** por el apoyo y las facilidades para poder realizar mi maestría.

Al **Consejo Nacional de Ciencia y Tecnología (CONACYT)** por la beca que me otorgo con la cual pude realizar mis estudios de maestría (No. Apoyo 477876).

Al **Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la UNAM** por el apoyo económico para poder realizar mi investigación <<No. de proyecto PAPIIT IA210617>>.

Al **Laboratorio Nacional de Análisis y Síntesis Ecológica (LANASE)** por permitirme usar su servidor de cómputo para llevar acabo mis análisis.

A mi tutor el Dr. Luis Eduardo Servín Garcidueñas por haberme aceptado como su alumno de maestría, por el tiempo que dedico a enseñarme todo lo relevante sobre el tema y su paciencia.

A la Dra. Yunuen Tapia Torres y a la Dra. María Esperanza Martínez Romero por formar parte de mi comité tutor y por sus consejos.

Agradezco a los miembros del jurado por su tiempo en la revisión de mi tesis. Gracias por su apoyo.

Agradecimientos personales.

Al Dr. Luis Eduardo Servín Garcidueñas por haberme aceptado como su alumno de maestría. Por su apoyo para poder realizar el cambio de tutor y ser parte de su grupo de trabajo. Por el tiempo que se tomaba en enseñarme todo lo relacionado a los análisis bioinformáticos.

Resumen	1
Abstract.....	2
Introducción.....	3
Campo geotérmico de Los Azufres	5
Antecedentes	8
Arqueas	9
Arqueas ARMAN (Parvarchaeota y Micrarchaeota).....	10
Arqueas ABC plasmatales (<i>Thermoplasmatales</i>).....	11
Arqueas <i>Sulfolobales</i>	12
Justificación	13
Hipótesis.....	13
Objetivos.	14
Objetivo general.	14
Objetivos particulares.....	14
Objetivos adicionales	14
Metodología.....	15
Recolección de muestras	15
Sedimentos fotosintéticos	18
Extracción y secuenciación de ADN de sedimentos fotosintéticos	18
Genomas individuales de arqueas <i>Thermoplasmatales</i>	18
Genoma parcial de <i>Candidatus Cuniculiplasma</i> sp. AZ01.....	19
Análisis de genómica comparativa entre genomas de <i>Cuniculiplasma</i>	20
Anotación funcional del genoma parcial de <i>Ca. Cuniculiplasma</i> sp. AZ01	21
Sedimentos no fotosintéticos	22
Extracción y secuenciación masiva de ADN de sedimentos no fotosintéticos.....	22
Procesamiento de lecturas de secuenciación masiva de ADN.....	22
Posición filogenómica de las poblaciones de arqueas	23
Anotación funcional del metagenoma	24
Secuencias de virus de arqueas <i>Sulfolobales</i>	24
Genoma parcial de <i>Candidatus Aramenus sulfurataquae</i> AZ2	24
Análisis de genómica comparativa entre genomas de <i>Ca. Aramenus sulfurataquae</i>	25
Anotación funcional del genoma parcial de <i>Ca. Aramenus sulfurataquae</i>	26
Secuencias de virus de arqueas <i>Sulfolobales</i>	26
Resultados	27
Sedimentos fotosintéticos	27
Secuenciación masiva de ADN de sedimentos fotosintéticos	27
Genomas individuales de arqueas <i>Thermoplasmatales</i>	27
Genoma parcial de <i>Ca. Cuniculiplasma</i> sp. AZ01	28
Análisis de genómica comparativa entre genomas de <i>Cuniculiplasma</i>	29
Anotación funcional del genoma parcial de <i>Ca. Cuniculiplasma</i> sp. AZ01	31
Sedimentos no fotosintéticos	36

Secuenciación masiva de ADN de sedimentos no fotosintéticos	36
Posición filogenómica de las poblaciones de arqueas	36
Anotación funcional del metagenoma	38
Secuencias de virus de arqueas <i>Sulfolobales</i>	41
Genoma parcial de <i>Ca. Aramenus sulfurataquae</i> AZ2	43
Análisis de genómica comparativa de <i>Ca. Aramenus sulfurataquae</i>	45
Anotación funcional del genoma parcial de <i>Ca. Aramenus sulfurataquae</i> AZ2	47
Secuencias de virus de arqueas <i>Sulfolobales</i>	50
Discusión de los resultados de sedimentos fotosintéticos	51
Discusión de los resultados de sedimentos no fotosintéticos.....	53
Conclusiones	56
Perspectivas.....	58
Bibliografía	59
Anexos	64
Scripts utilizados de las herramientas bioinformáticas	64
Sedimentos fotosintéticos.....	64
Sedimentos no fotosintéticos.....	64
Muestras adicionales de sedimentos del campo geotérmico de Los Azufres.	66

Resumen

El campo geotérmico de Los Azufres alberga una gran cantidad de manifestaciones geotérmicas de origen natural, principalmente manantiales termales y fumarolas que pueden alcanzar temperaturas mayores a los 90 °C y valores de pH menores a 4. En las fumarolas presentes en el campo geotérmico de Los Azufres se acumulan sedimentos que podrían albergar comunidades microbianas poco estudiadas. En el presente proyecto se secuenció un metagenoma de sedimentos no fotosintéticos de fumarolas de una zona rocosa dentro del campo geotérmico de Los Azufres. En el metagenoma de sedimentos no fotosintéticos se encontraron poblaciones de arqueas del orden *Sulfolobales* (*Sulfolobus* sp. A20 (CP017006.1), *Sulfolobus* sp. JCM 16833 (LC382015.1), *Acidianus brierleyi* DSM 1651 (CP029289.1) y *Candidatus Aramenus sulfurataquae* AZ1 (KC749962.1)) y virus de los géneros *Rudivirus* y *Fuselovirus*. En la anotación del metagenoma de sedimentos no fotosintéticos se encontraron genes involucrados en la fijación de carbono, metabolismo del azufre y nitrógeno. Como comunidad las arqueas *Sulfolobales* podrían estar desempeñando funciones ecológicas importantes en la fijación del carbono y metabolismos del azufre y nitrógeno en el campo geotérmico de Los Azufres. A partir de las secuencias metagenómicas de sedimentos no fotosintéticos se logró recuperar un genoma parcial de la arquea *Candidatus Aramenus sulfurataquae* AZ1 al que se le nombró como *Ca. Aramenus sulfurataquae* AZ02. En la anotación funcional del genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 se encontraron genes involucrados en la fijación del carbono y metabolismo del azufre. *Ca. Aramenus sulfurataquae* podría tener una relevancia ecológica importante en la fijación del carbono y metabolismo del azufre. Adicionalmente a los trabajos realizados en el presente proyecto se recuperó el genoma parcial de una arquea del género *Cuniculiplasma* en un metagenoma de sedimentos fotosintéticos. A este genoma parcial se nombró como *Candidatus Cuniculiplasma* sp. AZ01.

Abstract

The Los Azufres geothermal field contains a large number of geothermal springs of natural origin, mainly thermal springs and fumaroles that can reach temperatures above 90 °C and pH values below 4. In the fumaroles found at the Los Azufres geothermal field, sediments accumulate, and they could harbor poorly studied microbial communities. In the current project, a metagenome was sequenced from non-photosynthetic sediments of fumaroles from a rocky area within the Los Azufres geothermal field. Archaeal populations of the order *Sulfolobales* (*Sulfolobus* sp. A20, *Sulfolobus* sp. JCM 16833, *Acidianus brierleyi* DSM 1651 and *Candidatus Aramenus sulfurataquae* AZ1) were found in the metagenome of non-photosynthetic sediments, as well as viruses of the genus *Rudivirus* and *Fusellovirus*. Genes involved in carbon fixation, metabolism of sulfur and nitrogen were found in the annotation of the metagenome of non-photosynthetic sediments. This microbial community could be playing important ecological roles in carbon fixation and in sulfur and nitrogen metabolisms in the Los Azufres geothermal field. From the metagenomic sequences of non-photosynthetic sediments, it was possible to recover a partial genome of the archaea *Candidatus Aramenus sulfurataquae* AZ1, which was named *Ca. Aramenus sulfurataquae* AZ02. Genes involved in carbon fixation and sulfur metabolism were found in the functional annotation of the *Ca. Aramenus sulfurataquae* AZ2 genome. *Ca. Aramenus sulfurataquae* could play an important ecological relevance in carbon fixation and sulfur metabolism. In addition to the work carried out in the present project, the partial genome of an archaea of the genus *Cuniculiplasma* was recovered from a photosynthetic sediment metagenome. This partial genome was named *Candidatus Cuniculiplasma* sp. AZ01.

Introducción

Se ha estimado que en algunas muestras ambientales el 99% de los microorganismos no se pueden cultivar en condiciones de laboratorio debido a que no se tiene el conocimiento sobre requerimientos nutricionales; ya que los científicos han sido incapaces de crear diferentes ambientes con condiciones bióticas y abióticas para el cultivo de la mayoría de los microorganismos (Chaffron *et al.*, 2010; Pham & Kim, 2012; Killham, & Prosser, 2015). En la actualidad se emplean distintas metodologías de secuenciación masiva para analizar la diversidad de comunidades microbianas ambientales, tales como el análisis comparativo de secuencias de genes ribosomales 16S ARNr y más recientemente mediante análisis metagenómicos.

Kunin y colaboradores (2008) definen a la metagenómica como “*la secuenciación masiva de ADN obtenido directamente de una muestra ambiental o serie de muestras*”. La secuenciación masiva de ADN se realiza mediante las tecnologías de secuenciación de alto desempeño. Dentro de las tecnologías de secuenciación de próxima generación se encuentra Illumina. El sistema de secuenciación Illumina se basa en la reacción en cadena de la polimerasa (PCR) de puente para generar grupos de fragmentos de ADN idénticos. Primeramente, fragmentos de ADN que llevan marcadores a los extremos de las secuencias se unen a la superficie de la placa de secuenciación por complementariedad con las bases de los marcadores. Cuando el fragmento de ADN se une a la superficie de la placa se forma un puente de ADN. Posteriormente el puente de ADN es amplificado mediante PCR para así generar grupos de fragmentos de ADN idénticos (Clark *et al.*, 2019; Khodakov *et al.*, 2016). A partir de estos grupos de fragmentos de ADN se realiza la secuenciación por síntesis de ADN usando nucleótidos marcados con fluoróforos. Una de las ventajas de PCR de puente es que se pueden secuenciar los fragmentos de ADN de manera pareada. Illumina es un líder actual tanto por tener bajas las tasas de error de secuenciación como por el costo de secuencia por lectura (Khodakov

et al., 2016). Illumina cuenta con varias plataformas de secuenciación que se adaptan a las necesidades del consumidor. Entre estas plataformas se encuentran MiSeq, HiSeq, NextSeq, etc.

El desarrollo de la metagenómica y las tecnologías de secuenciación de próxima generación han permitido el analizar microorganismos que no se han logrado cultivar (Kunin et al., 2008; Wooley, Godzik, & Friedberg, 2010). Un ejemplo reciente de la importancia de la metagenómica es el trabajo realizado por Zaremba-Niedzwiedzka y colaboradores (2017), quienes realizaron análisis metagenómicos de muestras de sedimentos acuáticos de siete sitios geográficamente distantes y que presentaban diferencias en sus parámetros biológicos y químicos. Los sitios elegidos fueron Castillo de Loki; Parque Nacional de Yellowstone, USA; un acuífero cerca del río Colorado, USA; Radiata Pool, Nueva Zelanda; isla Taketomi, Japón; la bahía Aarhus, Dinamarca y el estuario del río del Roble Blanco, USA. A partir de estos análisis definieron un nuevo superfilo denominado arqueas Asgard. Dentro de estas arqueas se encuentran los filos Lokiarchaeota, Thorarchaeota, Heimdallarchaeota y Odinararchaeota. Las arqueas Asgard codifican un gran número de proteínas características de eucariontes (funciones de citoesqueleto, formación y tráfico vesicular). Además, las arqueas Asgard se afilian con eucariotas en análisis filogenéticos, por lo que Zaremba-Niedzwiedzka y colaboradores (2017) denominan a estas arqueas como el posible ancestro común arqueo-eucariótico.

Lo anterior nos habla de la importancia de la metagenómica en los descubrimientos recientes en los orígenes de la vida. Además, algunos de los sitios donde Zaremba-Niedzwiedzka y colaboradores (2017) realizaron análisis metagenómicos son sitios volcánicos submarinos o terrestres, indicando que en sitios volcánicos es posible encontrar arqueas nuevas y que brinden una mayor información sobre el origen de la vida.

Campo geotérmico de Los Azufres

En algunos sitios volcánicos se pueden encontrar manifestaciones geotérmicas, como manantiales termales, géiseres, fumarolas y volcanes de fango. Las manifestaciones geotérmicas presentan condiciones extremas de acidez y temperatura elevada (Huber *et al.*, 2000).

El Cinturón Volcánico Transmexicano es una cordillera rocosa en México que une a la Sierra Madre Occidental y la Sierra Madre Oriental y se extiende desde el Estado de Veracruz hasta el Estado de Nayarit (Demant, 1978). Dentro del Cinturón Volcánico Transmexicano (Figura 1) se encuentra el campo geotérmico de Los Azufres en el Estado de Michoacán, México (19.7833° N 100.65° W) (Figura 2). El campo geotérmico de Los Azufres alberga una gran cantidad de manifestaciones termales, principalmente manantiales y fumarolas que pueden alcanzar temperaturas mayores a los 90 °C y valores de pH menores a 4 (Tello-Lopez & Suárez-Arriaga 2000) (Figura 2 y Figura 3). Además, se encuentran gases no condensables en todo el campo geotérmico de Los Azufres. El dióxido de carbono (CO₂) es el principal constituyente y se encuentra entre el 70% y 90% del peso total de los gases no condensables. El sulfuro de hidrógeno (H₂S) es el segundo más importante y varía entre el 0.2% y el 13% del peso total de los gases no condensables (Arriaga-Suárez *et al.*, 2000).



Figura 1: Cinturón Volcánico Transmexicano y localización del campo geotérmico de Los Azufres (Modificado de Santoyo-Gutiérrez & Torres-Alvarado, 2010).



Figura 2: Laguna verde del campo geotérmico de Los Azufres localizada al lado de la unidad 17 de la CFE.



Figura 3: Fumarolas del campo geotérmico de Los Azufres en la zona del manantial de Marítaro.



Figura 4: Manantial termal del campo geotérmico de Los Azufres de la zona del manantial de Marítaro.

Antecedentes

Se han realizado estudios microbiológicos en Los Azufres mediante microscopía electrónica o mediante pruebas bioquímicas para determinar el tipo de microorganismos presentes sobre metales corroídos y dentro de tuberías empleadas para la generación de energía geotérmica. En estos estudios se encontró la presencia de arqueas y bacterias (Navarrette-Bedolla *et al.* 1999; Valdez-Salas *et al.* 2000; Cuevas-Villanueva *et al.* 2006; Castorena *et al.* 2006). En la investigación de 2014 de Brito y colaboradores, se detectó que las comunidades bacterianas de muestras de lodos ácidos de Los Azufres presentaban una diversidad limitada analizando secuencias de genes ribosomales 16S ARNr. En este trabajo se amplificaron genes ribosomales 16S ARNr de bacterias y las muestras fueron colectadas de un sitio con impacto humano que es utilizado como “spa natural” para turistas.

Recientemente se han realizado estudios de diversidad microbiana utilizando enfoques metagenómicos en manifestaciones termales de Los Azufres (Servín-Garcidueñas, 2015). Derivado de estos estudios se logró secuenciar un metagenoma de una solfatara ácida (terreno geológico donde se desprende vapor de agua que contiene ácido sulfhídrico y/o hidrógeno sulfurado). A partir de este metagenoma se logró recuperar el genoma consenso de una población de arqueas que permanecían sin caracterizar y que está relacionada filogenéticamente con arqueas del orden *Sulfolobales*. Esta arquea fue nombrada como *Candidatus Aramenus sulfurataquae* AZ1 (Servín-Garcidueñas & Martínez-Romero 2014). También se reportaron los genomas consenso de dos poblaciones de virus no descritos pertenecientes al género *Rudivirus* y *Fusellovirus* (Servín-Garcidueñas *et al.*, 2013b, 2013a). Por último, a partir del metagenoma de la solfatara ácida también se lograron detectar arqueas pertenecientes al orden *Thermoproteales*, a las que nombraron como *Thermoproteus* sp. AZ2 y *Vulcaniseta* sp. AZ3.

Servín-Garcidueñas (2015) colectaron muestras de sedimentos fotosintéticos depositados sobre rocas alrededor de fumarolas y lograron obtener ADN ambiental que fue utilizado para realizar

la secuenciación de un metagenoma. A partir del metagenoma de sedimentos fotosintéticos se identificaron genes ribosomales 16S y 18S ARNr de bacterias, arqueas y eucariontes. Se detectó la presencia de bacterias pertenecientes a los filos Proteobacteria, Actinobacteria, Nitrospira y Firmicutes. También se detectaron secuencias de eucariontes pertenecientes a los filos Euglenozoa y Ascomycota y una microalga no descrita del género *Cyanidiaceae*. En este metagenoma se encontró la presencia de arqueas pertenecientes al orden *Thermoplasmatales* conocidas como ABC plasmias de los linajes G, E, A, C, I, además de una arquea del filo Micrarchaeota dentro del linaje “*Archaeal Richmond Mine Acidophilic Nanoorganisms*” (ARMAN) (Servín-Garcidueñas, 2015). Se ha propuesto que las arqueas ABC plasma presentan interacciones simbióticas, parasitarias o antagónicas con las arqueas del filo Micrarchaeota y/o Parvarchaeota (Golyshina *et al.* 2017; Chen *et al.* 2018; Golyshina *et al.*, 2019).

Arqueas

Las arqueas son microorganismos procariontes y forman uno de los tres dominios de la vida (Woese *et al.* 1990; DasSarma *et al.*, 2009; Moissl-Eichinger *et al.*, 2018, Blohs *et al.*, 2019), también tienen características únicas que las hacen diferentes de los dominios Bacteria y Eucarya. Una de las características más importantes de las arqueas es la estructura de sus membranas lipídicas (Van de Vossenberg *et al.*, 1998, Blohs *et al.*, 2019). La estructura de sus membranas lipídicas presenta una columna vertebral de lípidos de glicerol-1-fosfato unida a isoprenoides mediante enlaces éter (Van de Vossenberg *et al.* 1998; Koga & Morii 2007; DasSarma *et al.* 2009; Cavicchioli 2011) que las hace más estables que las bicapas lipídicas de las bacterias y más resistentes a la oxidación y a las altas temperaturas que los enlaces éster (Van de Vossenberg *et al.*, 1998). Esto les permite resistir a condiciones extremas de temperatura, acidez, salinidad, etc. La mayoría de las arqueas son quimioautótrofos y son capaces de crecer a partir de compuestos inorgánicos simples, algunas otras

arqueas son heterótrofos y son capaces de crecer a partir de compuestos orgánicos complejos. También existen pocas arqueas fototróficas que usan la energía de la luz para crecer (Blohs *et al.*, 2019; DasSarma *et al.*, 2009). Las arqueas son consideradas como especialistas en la utilización de la energía, porque se encuentran presentes en ambientes con escasez de nutrientes (Blohs *et al.*, 2019). Las arqueas se encuentran distribuidas ampliamente por el mundo tanto en ambientes extremos como no extremos. Por ejemplo, en ambientes de extrema temperatura, acidez, salinidad, en los intestinos de rumiantes, etc., (Cavicchioli, 2011; DasSarma *et al.*, 2009). Las arqueas tiene roles ecológicos en los ambientes en los que viven. Un ejemplo son las arqueas del orden *Sulfolobales* las que contribuyen al ciclo del carbono como productores primarios y consumidores de material orgánico en ambientes de alta temperatura en campos de solfataras acidas alrededor del mundo (Dworkin *et al.* 2006; Boone 2013).

Arqueas ARMAN (Parvarchaeota y Micrarchaeota)

Las arqueas ARMAN pertenecen a los filos Parvarchaeota y Micrarchaeota dentro del superfilo DPANN (Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota y Nanohaloarchaeota) (Rinke *et al.* 2013; Chen *et al.* 2018). Fueron descritas por primera vez en la mina Iron Mountain en Richmond, California (Baker *et al.*, 2006). Las arqueas ARMAN presentan algunos de los genomas más reducidos dentro del dominio de las arqueas con un tamaño de genoma que va desde 0.64 Mb a 1.08 Mb (Chen *et al.*, 2018; Dombrowski *et al.*, 2019). Los genomas de algunas de estas arqueas ARMAN han perdido genes metabólicos esenciales incluidos algunos relacionados con la síntesis de proteínas y nucleótidos (Chen *et al.*, 2018). Se ha propuesto que las arqueas ARMAN dependen de interacciones con otras especies de arqueas, en especial con las arqueas del orden *Thermoplasmatales* (Baker *et al.* 2010; Castelle & Banfield 2018; Golyshina *et al.* 2017, 2019; Dombrowski *et al.* 2019). Recientemente Golyshina y colaboradores (Golyshina *et al.* 2017, 2019) describieron una estrecha relación entre las arqueas del género *Cuniculiplasma* (*Thermoplasmatales*) y las arqueas ARMAN.

Arqueas ABC plasmas (*Thermoplasmatales*)

Las arqueas del orden *Thermoplasmatales* pertenecen al filo *Euryarchaeota* (Boone, 2013). Dentro del orden *Thermoplasmatales* se distinguen cuatro familias: *Thermoplasmaceae* (Darland *et al.*, 1970), *Picrophilaceae* (Schleper *et al.*, 1995), *Ferroplasmaceae* (Golyshina *et al.*, 2000) y *Cuniculiplasmataceae* (Golyshina, Lünsdorf, *et al.*, 2016). El orden *Thermoplasmatales* se caracteriza porque son arqueas acidófilas que crecen en un rango de pH entre 0 y 4 (Boone, 2013). La mayoría de las arqueas *Thermoplasmatales* son termófilas (47-67°C) (Boone, 2013), con excepción de *Ferroplasma acidiphilum* (Golyshina *et al.*, 2000) y *Cuniculiplasma divulgatum* (Golyshina, Lünsdorf, *et al.*, 2016) que tienen temperaturas de crecimiento óptimo de 35/37°C y 37/40°C respectivamente. Las arqueas *Thermoplasmatales* son autótrofos (algunas especies del género *Ferroplasma*) o heterótrofos (*Ferroplasma*, *Cuniculiplasma*, *Thermoplasma* y *Picrophilus*) (Golyshina, Kublanov, *et al.*, 2016; Kämpfer, 2006), en su mayoría son organismos anaerobios facultativos con excepción del género *Picrophilus* que son aerobios obligados (Kämpfer, 2006). Algunas arqueas del género *Thermoplasmatales* son capaces de crecer en presencia de azufre molecular mediante la reducción de azufre generando la formación de sulfuro de hidrógeno (H₂S) (Dworkin *et al.* 2006; Boone 2013;). La nutrición de las arqueas del género *Thermoplasma* en sus respectivos hábitats naturales se basa en alimentarse de células en descomposición de otros organismos como, por ejemplo: *Acidianus brierleyi*, *Bacillus acidocaldarius*, *Cyanidium caldarium* y *Dactylaria gallopava* (Kämpfer, 2006). Del mismo modo, Golyshina *et al.*, (2016) sugieren que el estilo de vida del género *Cuniculiplasma* en su hábitat natural consiste en ser “carroñeros”. Además, Golyshina *et al.*, (2019) mencionan que coincidentemente donde encontraron la presencia de arqueas del género *Cuniculiplasma* también encontraron la presencia de algas acidófilas. Los metabolitos producidos por algas acidófilas podrían ser sustanciales para mantener la vida de las arqueas del género *Cuniculiplasma* en estos ambientes. Por lo que, las arqueas del género *Cuniculiplasma* podrían aprovechar los nutrientes producidos por las algas acidófilas en sus respectivos ambientes, además de

las interacciones con arqueas de los filos Micrarchaeota y Parvarchaeota. Por último, a la fecha se han reportado tres genomas de arqueas del género *Cuniculiplasma* (Golyshina, Lünsdorf, et al., 2016; Krause et al., 2017).

Arqueas *Sulfolobales*

Las arqueas del orden *Sulfolobales*, se encuentran dentro del supergrupo TACK que se encuentra en el filo *Crenarchaeota* y son capaces de crecer en presencia de azufre. Las arqueas del orden *Sulfolobales* se distinguen de otros órdenes de arqueas por sus características morfológicas, fisiológicas y moleculares (Kämpfer, 2006). Las arqueas *Sulfolobales* tienen forma de cocos regulares o irregulares y son principalmente termófilos extremos hasta hipertermófilos. El pH óptimo de crecimiento de las arqueas varía entre 2 y 4, mientras, que la temperatura óptima de crecimiento varía entre 60° y 90°C. Estas arqueas son aerobias, anaerobias facultativas y anaerobias. Las arqueas que crecen en condiciones anaerobias son capaces de reducir u oxidar el azufre elemental (S) a sulfuro de hidrógeno (H₂S) o ácido sulfúrico (H₂SO₄) (Kletzin et al. 2004; Dworkin et al. 2006; Boone 2013). Las arqueas del orden *Sulfolobales* habitan en su mayoría en zonas volcánicas continentales o termales con condiciones de extrema acidez y temperatura tales como aguas termales y campos de solfataras. Las arqueas del orden *Sulfolobales* desempeñan funciones ecológicas relevantes en hábitats ácidos y de alta temperatura en todo el mundo. Las *Sulfolobales* contribuyen al ciclo del carbono como productores primarios y consumidores de material orgánico (Dworkin et al. 2006; Boone 2013). El orden *Sulfolobales* incluye una familia (*Sulfolobaceae*) con siete géneros: *Sulfolobus* (Brock, 1978), *Acidianus* (Seegerer et al., 1986), *Metallosphaera* (Huber et al., 1989), *Stygiolobus* (Seegerer et al., 1991), *Sulfurisphaera* (Kurosawa et al., 1998), *Sulfurococcus* (Golovacheva et al., 1987) y *Saccharolobus* (Sakai & Kurosawa, 2018). Algunas de las cepas tipo de los géneros *Sulfolobus* y *Acidianus* han sido ampliamente estudiadas y son consideradas como organismos modelo de arqueas del orden *Sulfolobales* (Kletzin et al., 2004; Leigh et al., 2011).

Justificación

A la fecha se han realizado pocos estudios de muestras de sedimentos termales y no se han realizado análisis de diversidad microbiana en sedimentos no fotosintéticos en el campo geotérmico de Los Azufres. Las comunidades de arqueas no han sido totalmente descritas dentro del campo geotérmico de Los Azufres y a la fecha solo existen dos genomas de arqueas reportados dentro del campo geotérmico de Los Azufres. Estas dos arqueas corresponden a la arquea del orden *Sulfolobales Candidatus Aramenus sulfurataquae* AZ1, y la arquea del filo Micrarchaeota, *Candidatus Micrarchaeum* sp. AZ1. Solamente el genoma de la arquea del filo Micrarchaeota fue obtenido a partir de un metagenoma de sedimentos fotosintéticos. Por lo que, el presente proyecto explorará la diversidad de arqueas presentes en muestras de sedimentos no fotosintéticos. El tener acceso a sus genomas nos permitirá explorar sus potenciales genéticos y explorar sus contribuciones dentro de la comunidad microbiana.

En estudios previos se identificó la presencia de una arquea relacionada con el género *Cuniculiplasma* mediante genes ribosomales 16S ARNr. Solamente existen tres genomas reportados de arqueas del género *Cuniculiplasma*. Por lo que, sería importante recuperar el genoma de la arquea del género de *Cuniculiplasma* presente en Los Azufres. El obtener un genoma de alguna arquea del género *Cuniculiplasma* presente en Los Azufres permitirá realizar comparaciones genómicas contra los otros genomas reportados previamente. Además, nos permitirá acceder al potencial genético de esta arquea y así determinar cuáles son sus contribuciones a la comunidad microbiana en el campo geotérmico de Los Azufres.

Hipótesis

La presencia y abundancia de poblaciones de arqueas en sedimentos termales no fotosintéticos se favorece a temperaturas superiores a 80°C y valores de pH ácidos, mientras que a estas condiciones se desfavorece la presencia de organismos eucariontes y bacterias.

Objetivos.

Objetivo general.

Realizar estudios genómicos y metagenómicos de arqueas de sedimentos no fotosintéticos del campo geotérmico de Los Azufres, Michoacán.

Objetivos particulares.

1. Realizar el ensamble y la anotación de las secuencias metagenómicas obtenidas.
2. Analizar la composición taxonómica y filogenómica de las comunidades microbianas presentes en los sedimentos no fotosintéticos utilizando marcadores moleculares de arqueas.
3. Recuperar genomas individuales de arqueas presentes en el metagenoma de sedimentos no fotosintéticos.
4. Analizar y comparar ensambles de genomas obtenidos a partir de las secuencias metagenómicas de los sedimentos contra genomas de referencia utilizando técnicas bioinformáticas y de genómica comparativa.

Objetivos adicionales

1. Recuperar un genoma de una arquea del género *Cuniculiplasma* a partir de estudios metagenómicos de sedimentos fotosintéticos.
2. Analizar y comparar el ensamble del genoma de la arquea del género *Cuniculiplasma* obtenido a partir de las secuencias metagenómicas de sedimentos fotosintéticos contra genomas de referencia utilizando técnicas bioinformáticas y de genómica comparativa.

Metodología

Recolección de muestras

Sedimentos no fotosintéticos:

Durante Mayo de 2016 Servin-Garcidueñas colectó muestras de sedimentos no fotosintéticos en el campo geotérmico de Los Azufres en un área rocosa llena de grietas de las que emergen fumarolas calientes, justo enfrente a la laguna fangosa conocida como Los Azufres Natural Spa (Figura 5 y Figura 6) (19.7833° N 100.65° W). Los sedimentos no fotosintéticos se colectaron frotando cajas Petri estériles sobre una roca que se encontraba alrededor de una fumarola. La muestra de sedimentos no fotosintéticos tenía una temperatura de 88.5°C y un pH de 2.8, respectivamente.

El sitio de colecta se escogió por la presencia de fumarolas y sedimentos termales, por el fácil acceso y por el limitado impacto de actividades humanas como el turismo y la industria geotérmica. Los sitios fueron localizados en viajes de exploración.

Sedimentos fotosintéticos:

Los sedimentos fotosintéticos fueron colectados por Servín-Garcidueñas (2015) en 2013. Los sedimentos fotosintéticos fueron colectados en el mismo sitio de colecta anterior (Figura 5 y Figura 7). Esta muestra de sedimentos tenía una temperatura de 67°C y un pH de 3.

La diferencia entre los sedimentos no fotosintéticos y fotosintéticos es que en los sedimentos fotosintéticos se observa un color verde debido a actividad fotosintética de microalgas o cianobacterias (Figura 7). Mientras que en los sedimentos no fotosintéticos no se observa actividad de fotosíntesis y la muestra se asemeja más a cristales de azufre amarillos (Figura 6)



Figura 5: Campo geotérmico de Los Azufres, Michoacán, México. A) Localización de Los Azufres Spa Natural (imagen obtenida de Google Earth). B) Área rocosa llena de fumarolas de donde se colectaron las muestras de sedimentos. C) Sitio turístico de Los Azufres Spa Natural frente a la zona de colecta de muestra de sedimentos. D) Sitio donde se colectaron de sedimentos fotosintéticos en 2013. E) Sitio donde se colectaron sedimentos no fotosintéticos.



Figura 6: Sedimentos no fotosintéticos del área rocosa enfrente de Los Azufres Spa Natural.



Figura 7: Sedimentos fotosintéticos del área rocosa enfrente de Los Azufres Spa Natural.

Sedimentos fotosintéticos

Extracción y secuenciación de ADN de sedimentos fotosintéticos

La extracción de ADN de sedimentos fotosintéticos fue realizada por Servín-Garcidueñas (2015) en 2013 mediante el kit UltraClean Mega (Prep) Soil DNA kit (MoBio Laboratories, Inc., Carlsbad, CA). La secuenciación masiva de ADN fue realizada en 2014 mediante la plataforma Illumina Miseq 2x300 pb (pares de bases) en Macrogen, Corea del Sur. Las lecturas pareadas al ser de un tamaño de 300 pb pueden ser traslapadas para generar lecturas con un tamaño mayor. Servín-Garcidueñas (2015) unió las lecturas pareadas mediante el programa FLASH v.1.2.11 (Magoč & Salzberg 2011). Posteriormente, se ensambló el metagenoma mediante el programa Newbler v.2.3 (454 Life Science). Una vez ensamblado el metagenoma, Servín-Garcidueñas identificó genes ribosomales 16S ARNr mediante los servidores RNAmmer (Lagesen *et al.*, 2007) y WebMGA (Wu *et al.*, 2011).

Genomas individuales de arqueas *Thermoplasmatales*.

A partir del metagenoma ensamblado por Servin-Garcidueñas (2015) se evaluaron y compararon genomas individuales de arqueas *Thermoplasmatales* contra el metagenoma. La evaluación se realizó mediante el programa MetaQUAST v.5.0.0 (Mikheenko *et al.*, 2016). MetaQUAST evalúa qué tan completos se encuentran los genomas individuales contra genomas y/o metagenomas de referencia. En este caso se usó para evaluar y seleccionar genomas individuales de arqueas que se podrían recuperar fácilmente. Los genomas de arqueas *Thermoplasmatales* que se usaron fueron: *Cuniculiplasma divulgatum* PM4 (NZ_LT719092.1), *Cuniculiplasma divulgatum* S5 (LT671858.1), *Cuniculiplasma* sp. C_DKE (MWKS01000001.1), *Ferroplasma acidermus* fer1 (NC_021592.1), *Thermoplasmatales* archaeon A-plasma (ATMF01000001.1), *Thermoplasmatales* archaeon E-plasma (ATME01000001.1), *Thermoplasmatales* archaeon I-plasma (ATMD01000001.1) y *Thermoplasmatales* archaeon G-plasma (ATDV01000001.1).

Genoma parcial de *Candidatus Cuniculiplasma* sp. AZ01.

Servín-Garcidueñas (2015) reportó un gen ribosomal 16S ARNr que presenta una identidad del 100% contra la arquea *Cuniculiplasma divulgatum* PM4 (LT719092.1). Por lo que se tomó la decisión de recuperar el genoma parcial de dicha arquea presente en la muestra de sedimentos fotosintéticos. Este genoma fue nombrado como *Candidatus Cuniculiplasma* sp. AZ01

Como primer paso se usaron las lecturas sobrelapadas del metagenoma de sedimentos fotosintéticos de Servín-Garcidueñas (2015). Estas lecturas sobrelapadas fueron mapeadas contra el genoma de *Cuniculiplasma divulgatum* PM4 (LT719092.1) mediante el programa bowtie2 v.2.3.4.3 (Langmead & Salzberg 2012). El mapeo de las lecturas mediante bowtie2 se realizó de manera sensible y precisa (modo --“very-sensitive”) para obtener solamente las lecturas con mayor similitud al genoma de referencia. Una vez obtenidas las lecturas se procedió a realizar múltiples ensambles mediante el programa SPAdes v.3.12.0 (Bankevich *et al.*, 2012). Los parámetros estadísticos para seleccionar el mejor ensamble fueron que presentara un menor número de contigs, tamaño de contigs más largo, mayor longitud total del ensamble y un valor mayor de N50. Los contigs ensamblados con SPAdes fueron curados para eliminar contigs que no presentan similitud o no pertenecen al genoma de *Cuniculiplasma divulgatum* PM4.

Los contigs se curaron mediante alineamientos en BLAST v.2.10.0 (Altschul *et al.*, 1990) de los contigs ensamblados del genoma de *Cuniculiplasma* obtenido a partir de los sedimentos fotosintéticos contra el genoma de *Cuniculiplasma divulgatum* PM4. Los contigs que no presentaron identidad contra dicho genoma fueron eliminados. Para corroborar que no existieran secuencias de otras arqueas se buscaron marcadores moleculares mediante el servidor AMPHORA2 (<https://pitgroup.org/amphoranet/>) (Kerepesi *et al.*, 2014). AMPHORA2 busca un total de 104 marcadores moleculares de arqueas. Posteriormente, mediante alineamientos en BLAST v.2.10.0 (Altschul *et al.*, 1990) se corroboró que los marcadores moleculares pertenecieran al genoma de

Cuniculiplasma divulgatum PM4. Esto se realizó para comprobar que el genoma parcial de *Cuniculiplasma* obtenido a partir de los sedimentos fotosintéticos no presentara secuencias de otras arqueas.

Análisis de genómica comparativa entre genomas de *Cuniculiplasma*.

El genoma parcial de la arquea ABC plasma *Ca. Cuniculiplasma* sp. AZ01 se comparó contra el genoma de *Cuniculiplasma divulgatum* PM4. La comparación entre ambos genomas se realizó calculando los valores de identidad promedio de nucleótidos (ANI) (Konstantinidis & Tiedje, 2005a), identidad promedio de aminoácidos (AAI) (Konstantinidis & Tiedje, 2005b) y el valor de hibridación ADN-ADN (*in silico* DDH) (Goris *et al.* 2007; Meier-Kolthoff *et al.* 2013). Los valores de ANI y AAI se calcularon mediante el servidor <http://enve-omics.ce.gatech.edu/>. El valor de DDH se calculó mediante el servidor <http://ggdc.dsmz.de/ggdc.php#>.

Los valores de ANI y AAI son medidas de similitud genómica a nivel de nucleótidos/aminoácidos entre regiones codificantes (CDS) de dos genomas. El valor de ANI y AAI se determina mediante algoritmos de BLAST. La hibridación ADN-ADN es una técnica de laboratorio usada para delimitar taxonómicamente especies de arqueas y bacterias. Esta técnica es complicada de realizar y es propensa a cometer errores (Meier-Kolthoff *et al.*, 2013). Con el paso del tiempo y el incremento de genomas disponibles en bases de datos se han desarrollado métodos computacionales para calcular el valor de DDH. Estos métodos se basan en realizar alineamientos mediante BLAST (Meier-Kolthoff *et al.*, 2013).

A nivel funcional se compararon ambos genomas para buscar genes únicos que se encontraran presentes en el genoma parcial de la arquea ABC plasma *Ca. Cuniculiplasma* sp. AZ01. mediante el

servidor “*Rapid Annotation using Subsystem Technology*” (RAST) v.2.0 (<http://rast.theseed.org/>) (Aziz *et al.*, 2008).

Los contigs que sobrelapan del genoma parcial de *Ca. Cuniculiplasma* sp. AZ01 se alargaron mediante el programa OSlay v.1.0 (Richter *et al.*, 2007). Además, OSlay ordena los contigs respecto a un genoma de referencia (*Cuniculiplasma divulgatum* PM4). Una vez alargados y ordenados los contigs del genoma parcial de *Ca. Cuniculiplasma* sp. AZ01 se realizó un alineamiento global contra el genoma de referencia (*Cuniculiplasma divulgatum* PM4) mediante el programa MAUVE v.2.4.0 (Darling *et al.*, 2004). Los alineamientos con MAUVE nos permiten observar el orden genómico, la similitud entre ambos genomas y que tan completo se encuentra el genoma respecto al genoma de referencia.

Anotación funcional del genoma parcial de *Ca. Cuniculiplasma* sp. AZ01

La anotación de genomas se divide en dos pasos. El primer paso es realizar una anotación estructural para localizar las regiones codificantes dentro del genoma. Como segundo paso se realiza una anotación funcional la cual consiste en asignar una función bioquímica a los productos de genes. Existen diferentes herramientas y estrategias para realizar la anotación de genomas (Aubourg & Rouzé, 2001).

Nosotros realizamos la anotación funcional del genoma parcial de la arquea ABC plasma *Ca. Cuniculiplasma* sp. AZ01. mediante el servidor “*Integrated Microbial Genomes and Microbiomes*” (IMG/M) (Chen *et al.*, 2019). El servidor IMG/MG realiza la anotación mediante la tubería de anotación IMG v.5.0.0 (Chen *et al.*, 2019). IMG realiza la anotación de genomas/ mediante seis pasos: 1.- Anotación estructural del genoma. 2.- Anotación funcional del genoma. 3.- Anotación topológica de genes que codifican proteínas (péptido señal y regiones transmembrana). 4.- Anotación taxonómica

de genes que codifican proteínas (distribución filogenética). 5.- Disponibilidad de generación de datos por IMG. 6.- Posibilidad de reanotación. IMG utiliza una amplia variedad de programas o servidores para la realización de los seis pasos anteriores. En el siguiente enlace se encuentran los programas que usa IMG para cada paso (<https://img.jgi.doe.gov/docs/pipelineV5/>) .

Sedimentos no fotosintéticos

Extracción y secuenciación masiva de ADN de sedimentos no fotosintéticos

La extracción de ADN se realizó el mismo día de muestreo. Se usaron 5 g de sedimentos para realizar una sola extracción de ADN usando el *kit UltraClean Mega (Prep) Soil DNA kit* (MoBio Laboratories, Inc., Carlsbad, CA) siguiendo las instrucciones del fabricante. La muestra de ADN se guardó en refrigeración hasta la realización de la secuenciación. La secuenciación masiva del ADN se realizó durante junio de 2016 mediante el secuenciador Illumina Nextseq 550 2x75pb. La secuenciación se realizó en el Laboratorio de Nacional de Análisis y Síntesis Ecológica, en la Escuela Nacional de Estudios Superiores Unidad Morelia de la UNAM.

Procesamiento de lecturas de secuenciación masiva de ADN

Se realizaron múltiples ensambles del metagenoma mediante el programa SPAdes v.3.8.2 (Bankevich *et al.*, 2012). SPAdes realiza un proceso de verificación de calidad de las lecturas. Una vez realizados los ensambles se procedió a realizar una evaluación estadística de los ensambles mediante el programa QUAST v5.0.0 (“*Quality Assessment Tool*”) (Gurevich *et al.*, 2013). Los parámetros estadísticos para seleccionar el mejor ensamble fueron que presentara un menor número de contigs, tamaño de contigs más largo, mayor longitud total del ensamble y un valor mayor de N50.

Posición filogenómica de las poblaciones de arqueas

Mediante el servidor RAST v.2.0 (<http://rast.theseed.org/>) (Aziz *et al.*, 2008) se buscaron los marcos de lectura abierta (ORFs) contenidos en el metagenoma. Una vez encontrados los marcos de lectura abierta se buscaron genes de proteínas de arqueas que sirven como marcadores moleculares mediante el servidor AMPHORA2 (<https://pitgroup.org/amphoranet/>) (Kerepesi *et al.*, 2014). AMPHORA2 busca un total de 104 marcadores moleculares de arqueas. Se identificó a la especie a la que pertenecen los marcadores moleculares mediante búsquedas de BLAST v.2.10.0 (Altschul *et al.*, 1990). Se identificaron cuatro conjuntos de marcadores moleculares pertenecientes a arqueas del orden *Sulfolobales*: *Sulfolobus* sp. A20 (CP017006.1), *Sulfolobus* sp. JCM 16833 (LC382015.1), *Acidianus brierleyi* DSM 1651 (CP029289.1) y *Candidatus Aramenus sulfurataquae* (KC749962.1). Los marcadores moleculares se agruparon en conjuntos de acuerdo a la especie a la que pertenece el marcador. Cada uno de los conjuntos de marcadores moleculares se concatenaron en una sola secuencia. Para la realización del árbol filogenómico se seleccionaron organismos modelo y/o cepas tipo de arqueas del orden *Sulfolobales*. Los proteomas de las arqueas del orden *Sulfolobales* fueron obtenidos del GenBank. Las secuencias concatenadas de los marcadores moleculares y de los proteomas de arqueas del orden *Sulfolobales* se alinearon usando muscle v.3.8.31 (Edgar, 2004; Madeira *et al.*, 2019). El alineamiento se procesó con Gblocks v.0.91.b (Castresana 2000; Talavera & Castresana 2007) para seleccionar bloques conservados y eliminar las posiciones pobremente alineadas. La construcción del árbol filogenómico se realizó mediante un análisis de máxima verosimilitud usando PhyML v.3.0 (Guindon *et al.*, 2010a, 2010b).

Anotación funcional del metagenoma

Los perfiles funcionales del metagenoma de sedimentos no fotosintéticos se anotaron mediante el servidor “*Integrated Microbial Genomes and Microbiomes*” (IMG/M). Para la realización de la anotación de metagenomas IMG se usó la tubería de anotación de metagenomas MAP (Huntemann *et al.*, 2016). Esta tubería de anotación consta solamente de los pasos de anotación estructural y funcional del metagenoma.

Secuencias de virus de arqueas *Sulfolobales*

Mediante la anotación del metagenoma de sedimentos no fotosintéticos se encontraron secuencias pertenecientes a virus, por lo que se decidió realizar una búsqueda de secuencias virales contenidas en el metagenoma. La búsqueda de secuencias virales se realizó mediante el servidor “*Metagenomic Rapid Annotation using Subsystem Technology*” (MG-RAST) v.4.0.3 (Keegan *et al.*, 2016). Una vez obtenidas las secuencias virales se corroboró que si pertenecieran a virus mediante búsquedas de BLAST v. 2.10.0 (Altschul *et al.*, 1990). Las búsquedas de BLAST se realizaron contra las secuencias virales depositadas en la base de datos “*Nucleotide collection*” (nr/nt).

Genoma parcial de *Candidatus Aramenus sulfurataquae* AZ2

En base a los resultados del árbol filogenómico se decidió recuperar el genoma de la arquea *Ca. Aramenus sulfurataquae* AZ1 debido a que presenta alta similitud con las secuencias encontradas de esta arquea en el metagenoma de sedimentos no fotosintéticos. El genoma recuperado de esta arquea a partir de la muestra de sedimentos no fotosintéticos se nombró *Candidatus Aramenus sulfurataquae* AZ2. Las tres arqueas restantes presentaron un mayor número de variabilidad genética lo cual imposibilita el recuperar sus genomas a partir de mapeos contra genomas de referencia.

Las lecturas metagenómicas fueron mapeadas contra el genoma de *Ca. Aramenus sulfurataquae* AZ1 (ASRH00000000.1) mediante el programa bowtie2 v.2.3.4.3 (Langmead & Salzberg 2012). El mapeo de las lecturas mediante bowtie2 se realizó de manera sensible y precisa (modo “--very-sensitive”) para obtener solamente las lecturas con mayor similitud al genoma de referencia. Una vez obtenidas las lecturas se procedió a ensamblarlas mediante el programa SPAdes v.3.12.0 (Bankevich *et al.*, 2012). Los contigs ensamblados con SPAdes fueron curados para eliminar contigs que no presentan similitud o no pertenecen al genoma de *Ca. Aramenus sulfurataquae*. Los contigs se curaron realizando la misma metodología que en la sección sobre la obtención del genoma de *Candidatus Cuniculiplasma* sp. AZ01.

Análisis de genómica comparativa entre genomas de *Ca. Aramenus sulfurataquae*

El genoma de la arquea del orden *Sulfolobales* obtenido a partir de un metagenoma de sedimentos no fotosintéticos se comparó a nivel de secuencia de nucleótidos y aminoácidos contra el genoma de *Ca. Aramenus sulfurataquae* AZ1. La comparación entre ambos genomas a nivel de secuencia de nucleótidos y de aminoácidos se realizó calculando los valores de ANI, AAI (Rodríguez-R & Konstantinidis 2016) y el valor de *in silico* DDH (Goris *et al.*, 2007; Meier-Kolthoff *et al.*, 2013). Los valores de ANI y AAI se calcularon mediante el servidor <http://enve-omics.ce.gatech.edu/>. El valor de DDH se calculó mediante el servidor <http://ggdc.dsmz.de/ggdc.php#>.

Los contigs del genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 se alargaron y ordenaron mediante el programa OSlay v.1.0 (Richter *et al.*, 2007). Una vez alargados y ordenados los contigs del genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 se realizó un alineamiento global contra el genoma de referencia (*Ca. Aramenus sulfurataquae* AZ1) mediante el programa MAUVE v. 2.4.0 (Darling *et al.*, 2004).

Anotación funcional del genoma parcial de *Ca. Aramenus sulfurataquae* AZ2

La anotación funcional del genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 se realizó mediante el servidor “*Integrated Microbial Genomes & Microbiomes*” (IMG/M) (Chen *et al.*, 2019). El servidor IMG/MG realiza la anotación mediante la tubería de anotación IMG v. 5.0.0. (Chen *et al.*, 2019).

Secuencias de virus de arqueas *Sulfolobales*

Se realizó una búsqueda de secuencias virales contenidas en el genoma parcial de la arquea *Ca. Aramenus sulfurataquae* AZ2 mediante el servidor MG-RAST v.4.0.3 (Keegan *et al.*, 2016). Una vez obtenidas las secuencias virales se corroboró que sí pertenecieran a genomas de virus mediante búsquedas de BLAST v.2.10.0 (Altschul *et al.*, 1990).

Resultados

Sedimentos fotosintéticos

Secuenciación masiva de ADN de sedimentos fotosintéticos

A partir de la secuenciación de ADN de sedimentos fotosintéticos se obtuvieron 46,728,050 de lecturas pareadas de un tamaño de ~300 pb. Una vez superpuestas las lecturas, se obtuvieron un total de 19,422,519 lecturas traslapadas de un tamaño de entre 35-592 pb.

Genomas individuales de arqueas *Thermoplasmatales*

En la Figura 8 se muestra la representación gráfica de las fracciones de los genomas individuales de arqueas *Thermoplasmatales* presentes en el metagenoma. A partir de esta gráfica y con los análisis filogenéticos realizados por Chen *et al.*, (2018) se tomó la decisión de recuperar el genoma individual de la arquea del género *Cuniculiplasma*. Esta decisión se tomó porque se encuentra presente el 82.825% del genoma de *C. divulgatum* PM4 en el metagenoma de sedimentos fotosintéticos. Además, porque es una arquea del orden *Thermoplasmatales* que presenta una mayor distancia filogenética con las otras arqueas del orden *Thermoplasmatales* (Chen *et al.* 2018) y así evitar recuperar secuencias muy conservadas de otras arqueas.

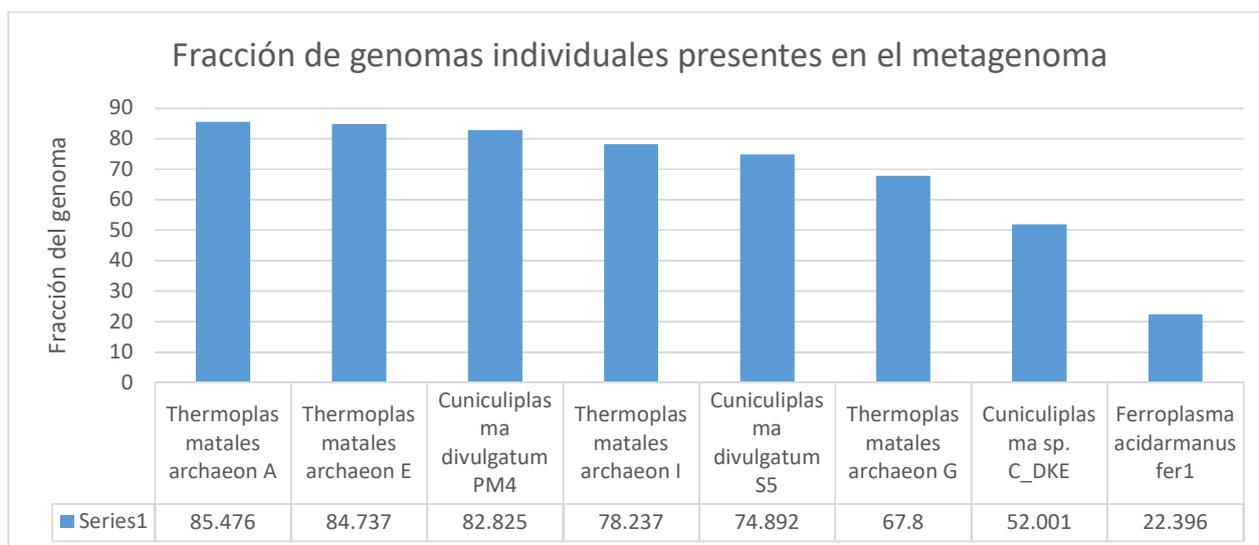


Figura 8: Representación gráfica de las fracciones de genomas individuales de arqueas *Thermoplasmatales* presentes en el metagenoma.

Genoma parcial de *Ca. Cuniculiplasma* sp AZ01

Con los mapeos mediante bowtie2 se obtuvieron 2,481,448 lecturas traslapadas de un tamaño entre 35-592 pb y que representan el 12.776% del total de las lecturas traslapadas. El genoma parcial de *Ca. Cuniculiplasma* sp. AZ01 ensamblado a partir de las lecturas de sedimentos fotosintéticos tiene 277 contigs que contienen 1,858,996 pb. Presenta un valor de N50 de 18,461 y una cobertura del genoma de 346.05X. El contenido de GC del genoma es de 37.3%. Este genoma parcial de *Ca. Cuniculiplasma* sp. AZ01 representa el 91.332% del tamaño total del genoma usado como referencia de *Cuniculiplasma divulgatum* PM4 (NZ_LT719092). El valor de 91.332% viene de obtener el genoma curado de *Ca. Cuniculiplasma* sp. AZ01.

Análisis de genómica comparativa entre genomas de *Cuniculiplasma*

En la Tabla 1 se muestran los valores de ANI, AAI y DDH que se obtuvieron al comparar el genoma de *Ca. Cuniculiplasma* sp. AZ01 contra los genomas de *Cuniculiplasma divulgatum* PM4 y S5. Valores de ANI mayores a 95%, AAI mayores a 90% y valores de DDH mayores a 70% nos permiten circunscribir taxonómicamente especies procariotas (Goris *et al.*, 2007; Meier-Kolthoff *et al.*, 2013; Rodriguez & Konstantinidis, 2014; Ciufu *et al.*, 2018; Jain, Rodriguez *et al.*, 2018; Medlar *et al.*, 2018).

Tabla 1: Métricos de identidad de secuencia entre los genomas disponibles de *Cuniculiplasma*. Las comparaciones se realizaron contra el genoma de *Ca. Cuniculiplasma* sp. AZ01.

	<i>Ca. Cuniculiplasma</i> sp. AZ01	<i>Cuniculiplasma divulgatum</i> PM4	<i>Cuniculiplasma divulgatum</i> S5
Número de contigs	277	1 cromosoma	1 cromosoma
Tamaño del genoma	1.85 MB	1.87 MB	1.93 MB
Contenido de G+C	37.3%	37.2%	37.3%
ANI%	100	99.12	98.66
AAI%	100	99.31	97.88
<i>In silico</i> DDH%	100	91.50	86.50

La comparación a nivel de genes entre ambos genomas de *Cuniculiplasma* nos reveló que no existen genes únicos presentes en el genoma parcial de *Ca. Cuniculiplasma* sp. AZ01. El que no existieran genes únicos en el genoma parcial de *Ca. Cuniculiplasma* sp. AZ01 podría ser debido a que el genoma aún se encuentra en un estado parcial. A pesar de lo anterior, se encontró que comparten 1,765 secuencias codificantes de proteínas completas o parciales con una identidad mayor al 70% (Figura 9).

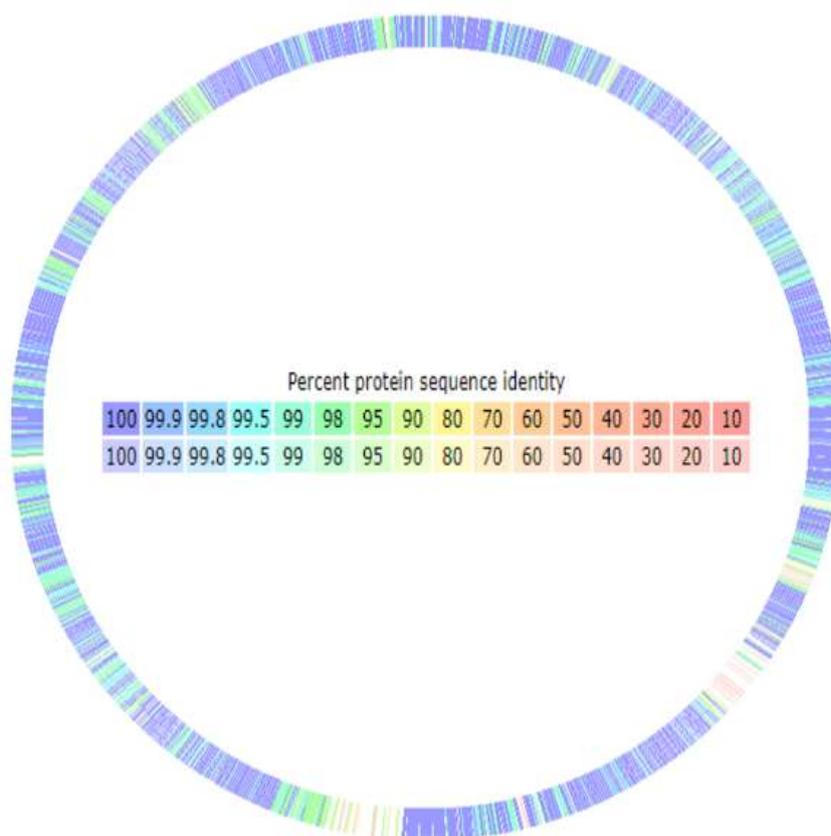


Figura 9: Comparación a nivel de secuencias de proteínas entre el genoma de referencia de *Cuniculiplasma divulgatum* PM4 y el genoma parcial de *Ca. Cuniculiplasma* sp. AZ01. Imagen generada mediante el servidor RAST (Aziz *et al.*, 2008).

En la Figura 10 se aprecia el alineamiento global realizado mediante MAUVE. En este alineamiento se observa que el genoma parcial de *Ca. Cuniculiplasma* sp. AZ01 se encuentra casi completo y presenta alta similitud contra el genoma de *Cuniculiplasma divulgatum* PM4. Esto se corrobora con los valores de ANI y DDH obtenidos previamente.

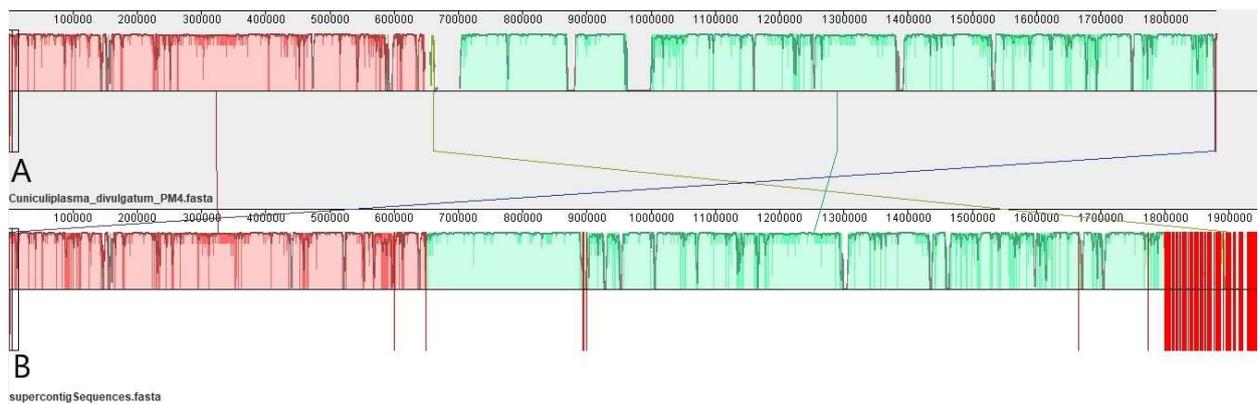


Figura 10: Alineamiento global realizado mediante MAUVE entre el genoma de *Cuniculiplasma divulgatum* PM4 y el genoma parcial de *Ca. Cuniculiplasma* sp. A) Genoma de *Cuniculiplasma divulgatum* PM4. B) Genoma parcial de *Ca. Cuniculiplasma* sp. AZ01

Anotación funcional del genoma parcial de *Ca. Cuniculiplasma* sp. AZ01

Mediante la anotación funcional de la arquea *Ca. Cuniculiplasma* sp. AZ01 se encontraron un total de 2,056 genes que codifican proteínas. De las cuales solo 1,472 genes tienen una función predicha contra los 584 genes que no tienen función conocida. En la Tabla 2 se muestran 1547 genes agrupados en el “Cluster of Orthologous Groups” (COGs).

Tabla 2: Genes asociados a categorías funcionales COGs (“Cluster of orthologous groups”) del genoma de *Ca. Cuniculiplasma* sp. AZ01 (Chen *et al.*, 2019).

Categoría COG	Número de genes	Porcentaje
Biosíntesis de metabolitos secundarios, transporte y catabolismo	33	1.83%
Control de ciclo celular, división celular, partición cromosómica	10	0.55%
Citoesqueleto	2	0.11%
Dinámica y estructura de cromatina	1	0.06%
Función desconocida	107	5.92%
Mecanismos de defensa	62	3.43%
Mecanismos de transducción de señales	27	1.49%
Mobiloma: profago, transposones	20	1.11%
Modificación post-transduccional, recambio proteico, chaperonas	73	4.04%
Motilidad celular	10	0.55%
Genes no agrupados en COGs	559	26.54%
Pared celular/membrana/sobre biogénesis	66	3.65%
Producción y conversión de energía	127	7.02%
Replicación, recombinación y reparación	85	4.7%
Solo predicción de función general	211	11.67%
Traducción, biogénesis y estructura de ribosomas	182	10.07%
Tráfico intracelular, secreción y transporte vesicular	10	0.55%
Transcripción	106	5.86%
Transporte y metabolismo de aminoácidos	205	11.34%
Transporte y metabolismo de carbohidratos	142	7.85%
Transporte y metabolismo de coenzimas	94	5.2%
Transporte y metabolismo de iones inorgánicos	101	5.59%
Transporte y metabolismo de lípidos	72	3.98%
Transporte y metabolismo de nucleótidos	62	3.43%

El potencial genético de la arquea *Ca. Cuniculiplasma* sp. AZ01 incluye todas las enzimas involucradas en glucólisis. En el caso de la enzima fosfoglucoasa isomerasa solamente se logró identificar una secuencia parcial mediante búsquedas de Blast usando el gen de la fosfoglucoasa isomerasa del genoma de *Cuniculiplasma divulgatum* PM4.

Los genes involucrados en el ciclo del ácido cítrico fueron identificados en el genoma parcial de *Ca. Cuniculiplasma* sp. AZ01 a excepción de las enzimas 2-oxoglutarato deshidrogenasa y fumarasa. La ausencia de estas enzimas también se reporta para el genoma de *Cuniculiplasma divulgatum* (Golyshina, Kublanov, *et al.*, 2016). A pesar de la ausencia de la enzima 2-oxoglutarato deshidrogenasa, *Cuniculiplasma* es capaz de llevar a cabo la descarboxilación del 2-oxoglutarato a succinato. La reacción de descarboxilación en *Cuniculiplasma* es llevada a cabo mediante la enzima 2-oxoácido: aceptor oxidoreductasa subunidad alfa y beta. Sin embargo, no existe alguna enzima que pueda llevar a cabo la reacción de adición de agua (H₂O) del fumarato a malato en los genomas de *Ca. Cuniculiplasma* sp. AZ01 y *Cuniculiplasma divulgatum* (Golyshina, Kublanov, *et al.*, 2016)

En los procesos de replicación y transcripción del ADN y traducción del ARN se encontraron todos los genes involucrados en dichos procesos. En la síntesis de pirimidinas se encuentran presentes todos los genes involucrados. Sin embargo, para la síntesis de purinas encontramos fragmentos del gen IMP ciclohidrolasa, esta enzima lleva a cabo la reacción de transformación del 5-formaminoimidazol-4-carboxamida ribotido (FAICAR) a Inosina monofosfato (IMP). Todas las demás enzimas involucradas en la síntesis de purinas se encuentran presentes.

En la síntesis de aminoácidos, *Ca. Cuniculiplasma* sp. AZ01 carece de los genes para la síntesis de histidina, leucina, isoleucina, prolina y valina. Esto se confirma con el genoma de *Cuniculiplasma divulgatum* PM4 que también carece de los genes para la síntesis de dichos aminoácidos (Golyshina, Kublanov, *et al.*, 2016).

Para el metabolismo del azufre se encontraron los genes que codifican para las enzimas sulfuro: quinona oxidoreductasa y una sulfotransferasa. La enzima sulfuro: quinona oxidoreductasa cataliza la oxidación del sulfuro de hidrógeno (H_2S). La enzima sulfotransferasa transfiere grupos sulfito entre moléculas. Mediante búsquedas de BLAST se identificó el gen que codifica para la enzima cisteína sintasa. La cisteína sintasa lleva a cabo la síntesis de cisteína mediante sulfuro de hidrógeno (H_2S) y O-acetil-L-serina. Respecto al metabolismo del nitrógeno, la única enzima encontrada es la glutamina sintetasa que realiza la síntesis de glutamina mediante amoníaco y glutamato.

También se encontraron genes que codifican para diferentes tipos de transportadores. Existe una proteína permeasa transportador ABC de la familia FeCT que actúa como transportador de absorción de quelato de hierro. En relación a lo anterior también se encontró el transportador Fe^{3+} -sideróforo tipo-ABC que libera sideróforos al medio extracelular para reducir el Fe^{3+} a Fe^{2+} y pueda ser utilizado por microorganismos. Existe un transportador ABC de proteína de membrana de la familia CUT1 que transporta carbohidratos al interior de la célula mediante hidrólisis de ATP (Figura 11). También se encontró el sistema de transporte permeasa de tipo-ABC de dipéptidos/oligopéptidos/níquel [aminoácidos] y el sistema de transporte de aniones tipo-ABC.

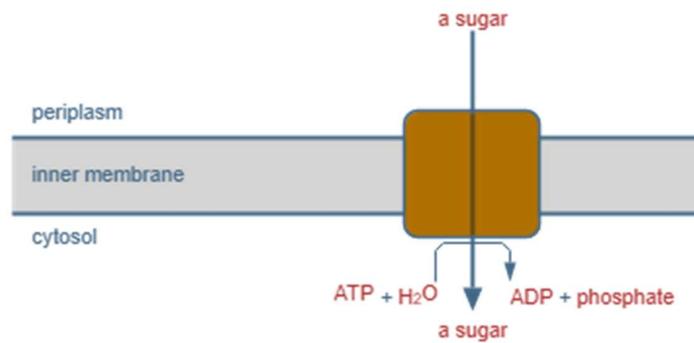


Figura 11: Transporte de carbohidratos mediante transportadores ABC. Modificado de <https://biocyc.org/gene?orgid=RINT718255&id=G13CX-3243-MONOMER#> (Karp *et al.*, 2018)

Dentro del genoma de *Ca. Cuniculiplasma sp. AZ01* se encontraron también los transportadores de superfamilias del máximo facilitador (MFS). DHA3 (MFS) funciona como un antiporte de aminoácidos o drogas: H^+ , algunos pueden transportar nucleótidos (Saier *et al.* 1999; Reddy *et al.* 2012). DHA1 (MFS) transportador antiporte drogas: H^+ (Reddy *et al.*, 2012). El transportador UMF12 (MFS) está posiblemente involucrado en el transporte de aminoácidos en arqueas (Reddy *et al.*, 2012). La familia 9 de MFS que son transportadores simporter de fosfato: H^+ (familia PHS). La familia 11 de MFS que son transportadores antiporte de oxalato: formiato (familia OFA). La familia 14 de MFS son transportadores simporte anion:cation (familia ACS) y pueden transportar aniones inorgánicos u orgánicos. Los transportadores de aniones inorgánicos cotransportan Na^+ , mientras que, los transportadores de aniones orgánicos cotransportan H^+ (Paulsen & Saier 1998). Por último, se encontró un transportador de aminoácidos, bombas de eflujo para sulfitos y arsénico y el transportador simporte núcleo base:cación (familia NCS1).

Sedimentos no fotosintéticos

Secuenciación masiva de ADN de sedimentos no fotosintéticos

A partir de la secuenciación de ADN se obtuvieron 60, 553, 942 lecturas pareadas de un tamaño de 75 pb. Las lecturas del metagenoma fueron depositadas en la base de datos de NCBI en el archivo de “*Sequence Read Archive*” bajo el número de acceso SRR7716004. Del total de lecturas se ensamblaron 56,000,00 lecturas pareadas. El metagenoma ensamblado tiene 1,726 contigs con una longitud mínima de 1,000 pb que contienen 11,295,008 pb. El valor de N50 es de 17,474 pb y presenta un contenido de GC de 39.2%.

Posición filogenómica de las poblaciones de arqueas

Mediante el servidor AMPHORA2 se encontraron cuatro conjuntos completos de 98 proteínas marcadoras de arqueas que contienen 25,962 caracteres, lo que indica que existen cuatro poblaciones de arqueas del orden *Sulfolobales* en el metagenoma secuenciado. La posición filogenómica de las cuatro poblaciones de arqueas dentro del orden *Sulfolobales* se observa en la Figura 12. Además, se encontraron cuatro genes ribosomales 16S ARNr de arqueas *Sulfolobales* (Tabla 3). Estas poblaciones de arqueas se nombraron como AZsed1 a AZsed4.

Tabla 3: Secuencias de genes ribosomales 16S ARNr de arqueas identificadas en el metagenoma de sedimentos no fotosintéticos.

Población	Tamaño	Especie	Identidad
AZsed1	1,498 pb	<i>Candidatus Aramenus sulfurataquae</i> (KC749962.1)	100%
AZsed2	1,497 pb	<i>Acidianus brierleyi</i> DSM 1651 (CP029289.1)	100%
AZsed3	1,387 pb	<i>Sulfolobus</i> sp. JCM 16833 (LC382015.1)	100%
AZsed4	1,494 pb	<i>Sulfolobus</i> sp. A20 (CP017006.1)	100%

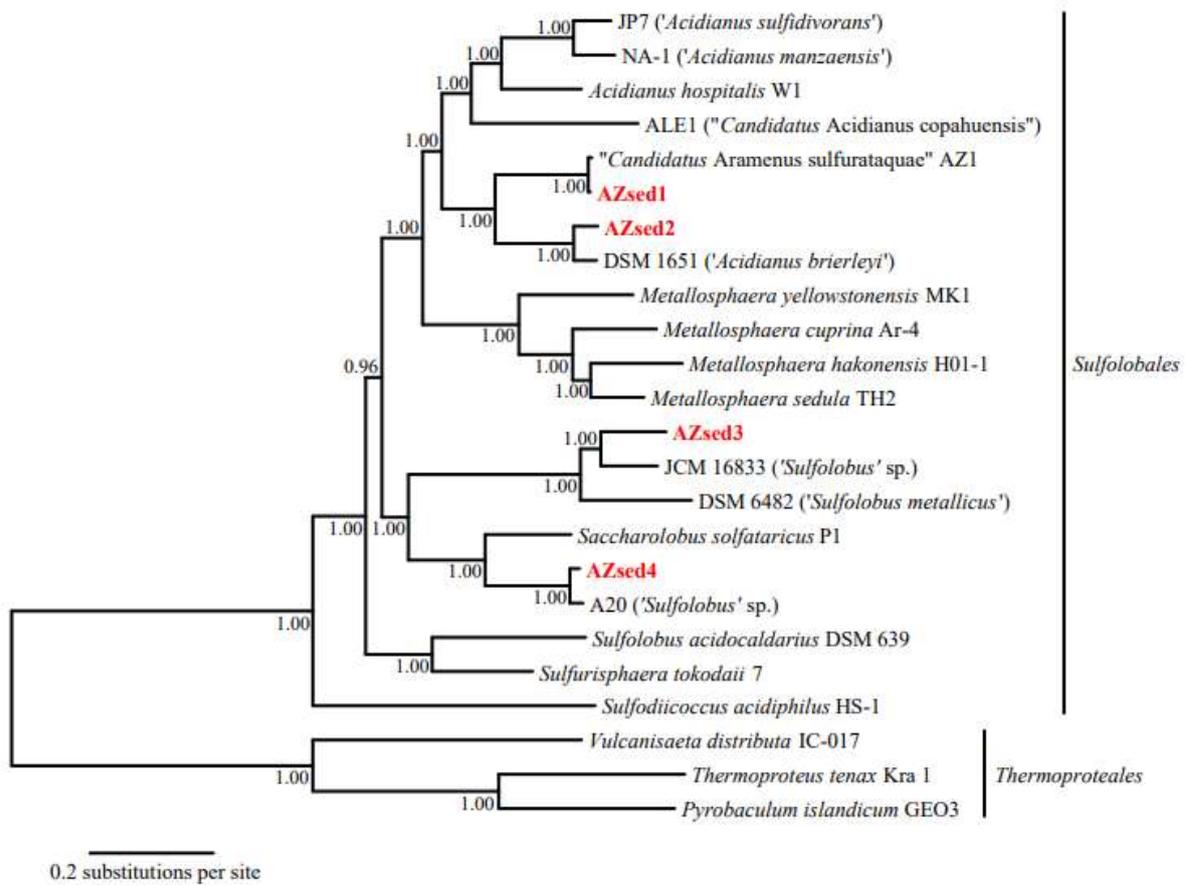


Figura 12: Análisis filogenómico que muestra las relaciones evolutivas de las arqueas *Sulfolobales* identificadas en este estudio. La filogenia se basa en una alineación de secuencia concatenada de 98 proteínas marcadoras de arqueas que contienen 25,962 caracteres

Anotación funcional del metagenoma

Mediante el servidor IMG/MER se lograron predecir 13,413 secuencias de genes que codifican proteínas.

Además, se encontró un total de 134 posibles secuencias de CRISPR.

Tabla 4: Genes asociados a categorías funcionales COGs (“*Cluster of orthologous groups*”)(Chen *et al.*, 2019).

Categoría COG	Número de genes	Porcentaje
Biosíntesis de metabolitos secundarios, transporte y catabolismo	191	2.19%
Control de ciclo celular, división celular, partición cromosómica	71	0.81%
Citoesqueleto	1	0.01%
Dinámica y estructura de cromatina	13	0.15%
Estructuras extracelulares	5	0.06%
Función desconocida	404	4.64%
Mecanismos de defensa	228	2.62%
Mecanismos de transducción de señales	176	2.02%
Mobiloma: profago, transposones	158	1.81%
Modificación post-transduccional, recambio proteico, chaperonas	364	4.18%
Modificación y procesamiento de ARN	4	0.05%
Motilidad celular	66	0.76%
Genes no agrupados en COGs	6327	46.46%
Pared celular/membrana/sobre biogénesis	315	3.62%
Producción y conversión de energía	705	8.09%
Replicación, recombinación y reparación	353	4.05%
Solo predicción de función general	1088	12.49%
Traducción, biogénesis y estructura de ribosomas	840	9.64%
Tráfico intracelular, secreción y transporte vesicular	52	0.60%
Transcripción	518	5.95%
Transporte y metabolismo de aminoácidos	916	10.51%
Transporte y metabolismo de carbohidratos	540	6.20%
Transporte y metabolismo de coenzimas	558	6.40%
Transporte y metabolismo de iones inorgánicos	486	5.35%
Transporte y metabolismo de lípidos	316	3.63%
Transporte y metabolismo de nucleótidos	364	4.18%

En la Tabla 4 se observan las secuencias de proteínas anotadas asociadas a agrupaciones de grupos ortólogos (COGs). La base de datos de proteínas COG se generó comparando proteínas predichas y conocidas en todos los genomas microbianos completamente secuenciados para inferir conjuntos de ortólogos.

En el metagenoma se encuentran presentes la mayoría de las proteínas de las rutas metabólicas de glucólisis, ciclo del ácido cítrico, fosforilación oxidativa, metabolismo de lípidos, biosíntesis de la columna vertebral de terpenoides.

Para la vía de fijación de carbono en procariontes, se identificó la enzima deshidrogenasa de monóxido de carbono. La deshidrogenasa de monóxido de carbono cataliza la siguiente reacción $\text{CO} + \text{H}_2\text{O} + \text{A (cofactor)} \rightarrow \text{CO}_2 + \text{AH}_2$. Esta reacción de oxidación del monóxido de carbono a dióxido de carbono permite a los organismos procariontes usar dióxido de carbono como fuente principal de carbono. Además, se encontraron enzimas que participan en el ciclo 3-hidroxiacetil-CoA/4-hidroxiacetil-CoA. El ciclo 3-hidroxiacetil-CoA/4-hidroxiacetil-CoA es el medio por el cual las arqueas autotróficas pueden fijar dióxido de carbono. Por último, se encontró la enzima acetil-CoA/propionil-CoA carboxilasa que presenta un rol clave en la carboxilación del dióxido de carbono.

Se identificaron varias enzimas involucradas en el metabolismo del azufre. Se encontró la enzima oxigenasa/reductasa de azufre la cual cataliza la oxidación y reducción de azufre ($4 \text{S} + 4 \text{H}_2\text{O} + \text{O}_2 \rightarrow 2 \text{HS} + 2 \text{HSO}_3^- + 2 \text{H}^+$). Además, se encuentra presente la enzima sulfato adeniltransferasa ($\text{ATP} + \text{SO}_4^{2-} \rightarrow \text{fosfato} + \text{adenil sulfato}$) que inicia la reacción de reducción asimilativa de sulfatos (Figura 13). La enzima fosfoadenosina fosfosulfato sintetasa que interviene en la reacción de asimilación no se encontró presente. Se encontraron las enzimas fosfoadenosina-fosfosulfato reductasa y la enzima fosfoadenosina-fosfosulfato, dichas enzimas se encuentran involucradas en la ruta de reducción asimiladora de sulfatos. Por último, se encontró la enzima reductasa asimilativa de

sulfitos (NADP) que cataliza la reducción de sulfito a sulfuro de hidrógeno ($3 \text{ H}_2\text{O} + \text{SH} + 3 \text{ NADP}^+ \rightarrow 4 \text{ H}^+ + 3 \text{ NADPH} + \text{SO}_3^-$).

Además, existe la presencia de una familia de transportadores de fosfato inorgánico. Este transportador también permite el paso de sulfatos inorgánicos.

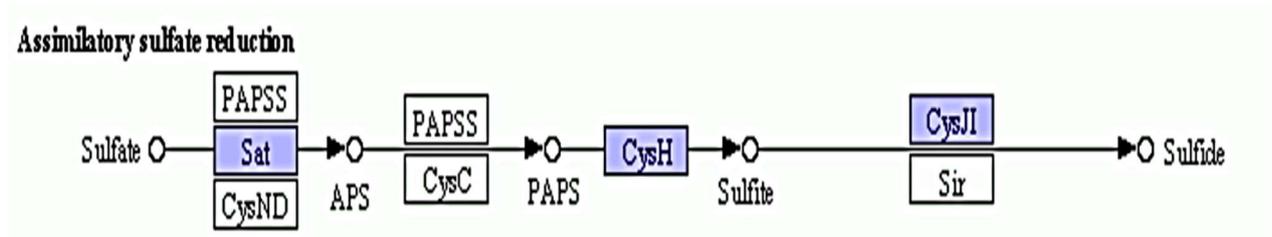


Figura 13: Ruta de la reducción asimilativa de sulfatos (imagen generada en el servidor IMG/MER)(Chen *et al.*, 2019).

En el metabolismo de nitrógeno se encontraron los transportadores: sistema de transporte tipo ABC nitrato/sulfonato/bicarbonato, componente ATP, de la familia NitT/TauT y el transportador MFS, familia NNP, transportador nitrato/nitrito. Se encuentran presentes las enzimas para la ruta de reducción asimilativa de nitratos (Figura 14). La primera enzima es la ferredoxina nitrato reductasa que lleva a cabo la reacción de reducción de nitrato a nitrito ($\text{NO}_3^- + 2 \text{ ferredoxina reducida} + 2 \text{ H}^+ \rightarrow \text{NO}_2^- + \text{H}_2\text{O} + 2 \text{ ferredoxina oxidada}$). Posterior a esta reacción el nitrito es reducido nuevamente a amoníaco mediante la enzima ferredoxina nitrito reductasa ($\text{NO}_2^- + 6 \text{ ferredoxina reducida} + 7 \text{ H}^+ \rightarrow \text{NH}_3 + 2 \text{ H}_2\text{O} + 6 \text{ ferredoxina oxidada}$). Por último, el amoníaco entra a la síntesis de aminoácidos mediante la enzima glutamina sintetasa.

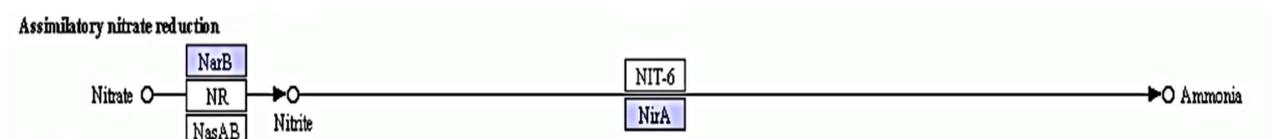


Figura 14: Ruta de la asimilación de nitratos (imagen generada en el servidor IMG/MER)(Chen *et al.*, 2019).

También se encontraron genes que codifican proteínas para los procesos de replicación del ADN, transcripción del ADN y síntesis de proteínas. Para el proceso de replicación del ADN se encontraron todos los elementos que conforman el complejo de replicación de arqueas (ADN polimerasa PolB, helicasa, primasa, proteínas SSB, clamp, clamp loader, ADN ligasa y RNase H). En el proceso de transcripción se encontraron presentes genes que codifican para todas las subunidades de la ARN polimerasa de arqueas y los factores de transcripción TFIIB, TBP, la subunidad alfa del factor de transcripción TFIIE1, la proteína de reparación de escisión de ADN ERCC-2 (XDP). Para la síntesis de proteínas se encontraron presentes los 20 aminoacil-ARNt sintetasa. Los factores de iniciación eIF1, eIF2, eIF2B y eIF5B, factores de elongación EF-Tu y el factor de liberación eRF1.

Secuencias de virus de arqueas *Sulfolobales*

Se encontraron un total de 19 contigs que presentan identidad con los genomas de “*Sulfolobales Mexican Rudivirus 1*” (NC_019413.1), “Los Azufres archaeal virus 1” (secuencia parcial) (KF547926.1), “*Sulfolobales Mexican Fusellovirus 1*” (KC618393.1) y “Los Azufres archaeal virus 2” (secuencia parcial) (KF547927.1). En la Tabla 5 se aprecia un listado de los contigs que contienen secuencias virales y la especie de virus con la que presenta mayor identidad. Para el genoma de SMR 1 (Figura 15) y la secuencia parcial de Los Azufres archaeal virus 1 (Figura 16) se pudieron realizar alineamientos globales mediante el programa MAUVE. En los alineamientos globales se aprecia que se encuentran presentes la mayor parte de las secuencias de los virus SMR 1 y “Los Azufres archaeal virus 1”. Sin embargo, no presentan una alta identidad entre las secuencias. Ambos alineamientos globales presentan contigs metagenómicos que no alinearon con los genomas de SMR1 y “Los Azufres archaeal virus 1”. Los contigs que no alinearon se debe a que son secuencias repetidas y de baja identidad de los genomas de SMR1 y “Los Azufres archaeal virus 1” que MAUVE no alinea.

Tabla 5: Listado de contigs del metagenoma que contienen secuencias virales.

Contig ID	Longitud del contig	Identidad	Especie	Nucleótidos alineados
102	21,934	74%	"Sulfolobales Mexican <i>Rudivirus</i> 1" (NC_019413.1)	3,644
		75%		1,106
		84%		413
		84%		413
311	7,980	73%		1,374
436 ¹	5,153	71%		338
514 ²	4,260	87%		248
644	3,228	74%		2,185
677 ³	3,033	80%		506
854 ⁴	2,318	70%	1,366	
1205 ⁵	1,564	71%	1,367	
1267 ⁶	1,473	77%	547	
Contig	Longitud del contig	Identidad	Especie	Nucleótidos alineados
384	6,157	84%	"Los Azufres archaeal virus 1", secuencia parcial (KF547926.1)	1,236
424	5,318	77%		2,830
464	4,770	80%		2,893
581*	3,729	100%		39
596	3,598	79%		2,355
1216	1,549	87%		465
1273	1,463	81%		620
1411	1,271	76%		712
1458**	1,238	78%		570
Contig	Longitud del contig	Identidad	Especie	Nucleótidos alineados
578	3,749	74%	"Sulfolobales Mexican <i>Fusellovirus</i> 1" (KC618393.1)	407
Contig	Longitud del contig	Identidad	Especie	Nucleótidos alineados
1140	1,657	74%	"Los Azufres archaeal virus 2", secuencia parcial (KF547927.1)	1,657

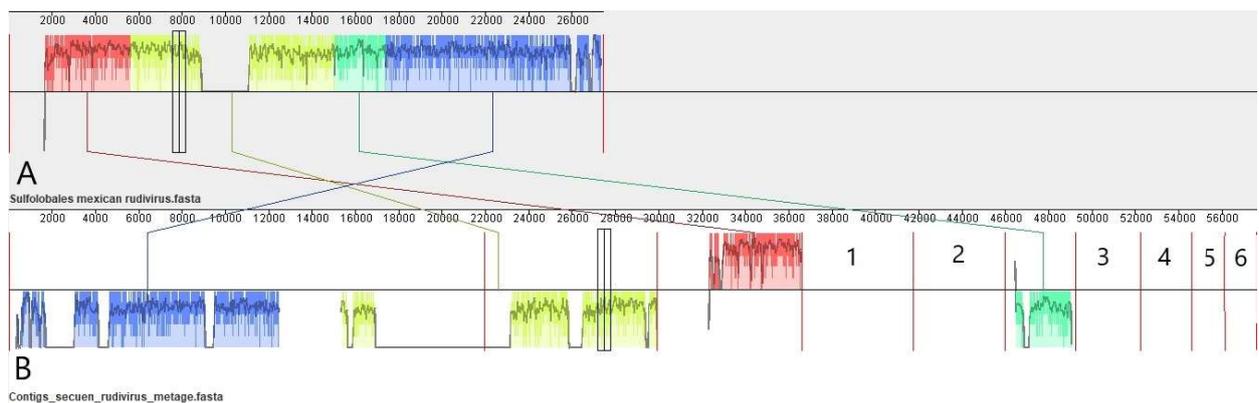


Figura 15: Alineamiento global entre el genoma de "Sulfolobales Mexican *Rudivirus* 1" y los contigs del metagenoma que contienen secuencias virales de *Rudivirus*. A) Genoma de "Sulfolobales Mexican *Rudivirus* 1". B) Secuencias virales dentro del metagenoma que presentan identidad contra el genoma de "Sulfolobales Mexican *Rudivirus* 1".

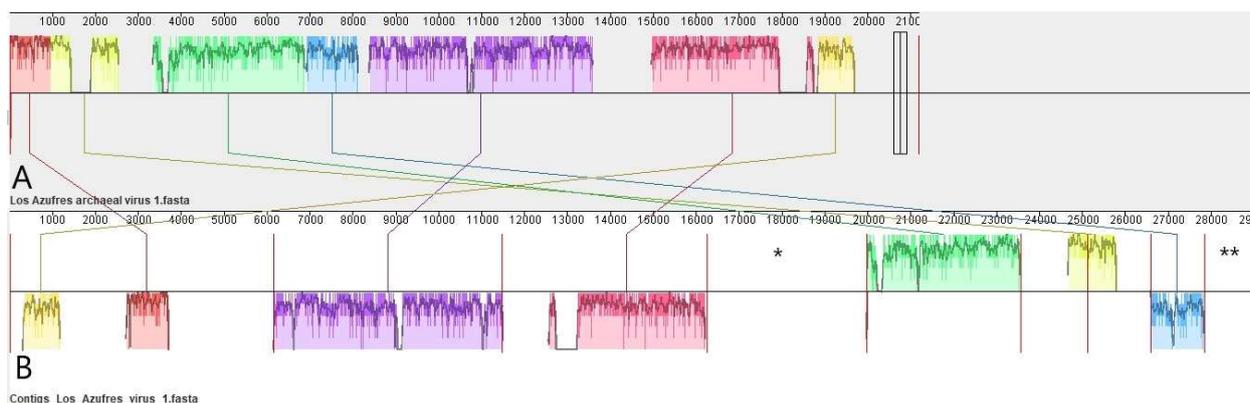


Figura 16: Alineamiento global entre la secuencia parcial de “Los Azufres archaeal virus 1” y los contigs del metagenoma que contienen secuencias de dicho virus. A). Secuencia parcial de “Los Azufres archaeal virus 1”. B). Secuencias virales dentro del metagenoma que presentan identidad contra la secuencia parcial de “Los Azufres archaeal virus 1.

Genoma parcial de *Ca. Aramenus sulfurataquae* AZ2

Mediante los mapeos de bowtie2 se obtuvieron 2,224,262 lecturas pareadas con un tamaño entre 35-76 pb y representan el 3.673% del total de las lecturas del metagenoma de sedimentos no fotosintéticos.

El genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 ensamblado a partir de las lecturas de sedimentos no fotosintéticos tiene 379 contigs que contienen 1,681,874 pb. Presenta un valor de N50 de 11,611 pb y una cobertura del genoma de 84.31X. El genoma tiene un contenido de GC de 47.8%. El genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 representa el 90.927% del tamaño total del genoma usado como referencia de *Ca. Aramenus sulfurataquae* AZ1 (ASRH000000000.1).

El genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 es el tercer genoma secuenciado y ensamblado dentro de la zona del campo geotérmico de Los Azufres. Además de la distancia entre los tres sitios de colecta (Figura 17), también presentan diferencias en condiciones ambientales. Del sitio uno se colectó una muestra de agua que presentaba una temperatura de 89°C y un pH de 3.7. En esta muestra se encontró únicamente la presencia de la arquea *Ca. Aramenus sulfurataquae* AZ1 y dos virus. Del sitio dos se colectó una muestra de lodo que presentaba una temperatura de 65°C y un pH

de 3.0. En esta muestra se encontró la presencia de arqueas del orden *Thermoproteales* (*Thermoproteus* sp. y *Vulcaniseta* sp.), de dos genomas de virus y la arquea *Ca. Aramenus sulfurataquae* AZ1. Del sitio tres se colectó una muestra de sedimentos no fotosintéticos que contenían cristales de azufre. En esta muestra se encontró la presencia de arqueas de orden *Sulfolobales* y se encontraron secuencias de virus.



Figura 17: Localización de los sitios de colecta de muestras donde se encontró la presencia de la arquea *Ca. Aramenus sulfurataquae*. El sitio de colecta uno (pH 3.7, T 89°C) se encuentra en la zona de manantiales termales de Marítaro. El sitio de colecta dos (pH 3.0, T 65°C) se encuentra en una solfatara ácida enfrente de la central eléctrica Los Azufres CFE. El sitio de colecta tres (pH 2.8, T 88.5°C) se encuentra en una zona rocosa enfrente de Los Azufres Spa natural.

Análisis de genómica comparativa de *Ca. Aramenus sulfurataquae*

En la Tabla 6 se muestran los valores de ANI, AAI y DDH que se obtuvieron al comparar los genomas de *Ca. Aramenus sulfurataquae*. Valores de ANI mayores a 95%, AAI mayores a 90% y valores de DDH mayores a 70% nos permiten circunscribir taxonómicamente especies procariotas (Goris *et al.*, 2007; Meier-Kolthoff *et al.*, 2013; Rodriguez & Konstantinidis, 2014; Ciufu *et al.*, 2018; Jain, Rodriguez *et al.*, 2018; Medlar *et al.*, 2018).

Tabla 6: Métricos de identidad de secuencia entre los genomas disponibles de *Ca. Aramenus*.

	<i>Ca. Aramenus sulfurataquae</i> AZ2	<i>Ca. Aramenus sulfurataquae</i> AZ1
Numero de contigs	379	46
Tamaño del genoma	1.68 Mb	1.79 Mb
Contenido de G+C	47.8%	47%
ANI%	100	97.91
AAI%	100	97.74
DDH%	100	81.70

La comparación a nivel de genes entre ambos genomas de *Ca. Aramenus sulfurataquae* nos reveló que no existen genes únicos presentes en el genoma parcial de *Ca. Aramenus sulfurataquae* AZ2. El que no existieran genes únicos en el genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 podría ser debido a que el genoma aún se encuentra en un estado parcial. A pesar de lo anterior, se encontró que comparten 1869 secuencias de proteínas completas o parciales con una identidad mayor al 70% (Figura 18).

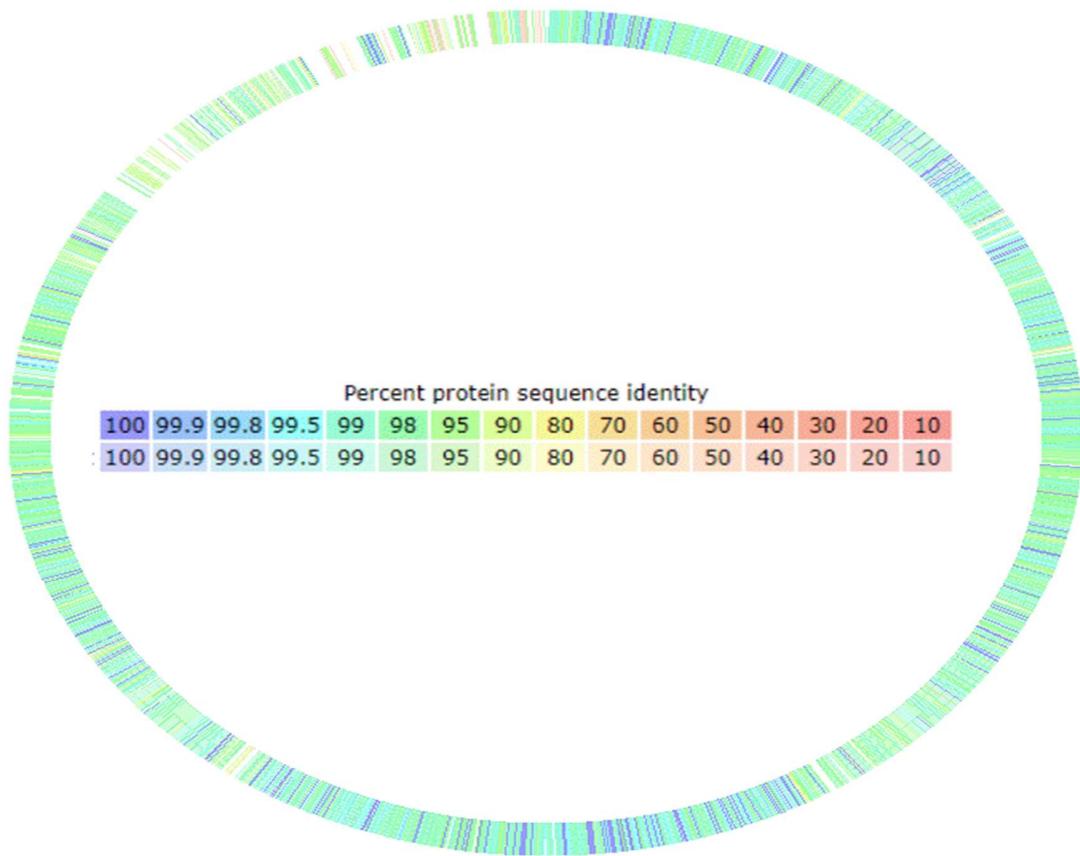


Figura 18: Comparación a nivel de secuencias de proteínas entre los genomas de *Ca. Aramenus sulfurataquae* (imagen generada mediante el servidor RAST(Aziz *et al.*, 2008).

En la Figura 19 se aprecia el alineamiento global realizado mediante MAUVE entre los genomas de *Ca. Aramenus sulfurataquae*. En este alineamiento se observa que el genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 se encuentra casi completo y presenta alta similitud contra el genoma de *Ca. Aramenus sulfurataquae* AZ1. Esto se corrobora con los valores de ANI y DDH obtenidos previamente.

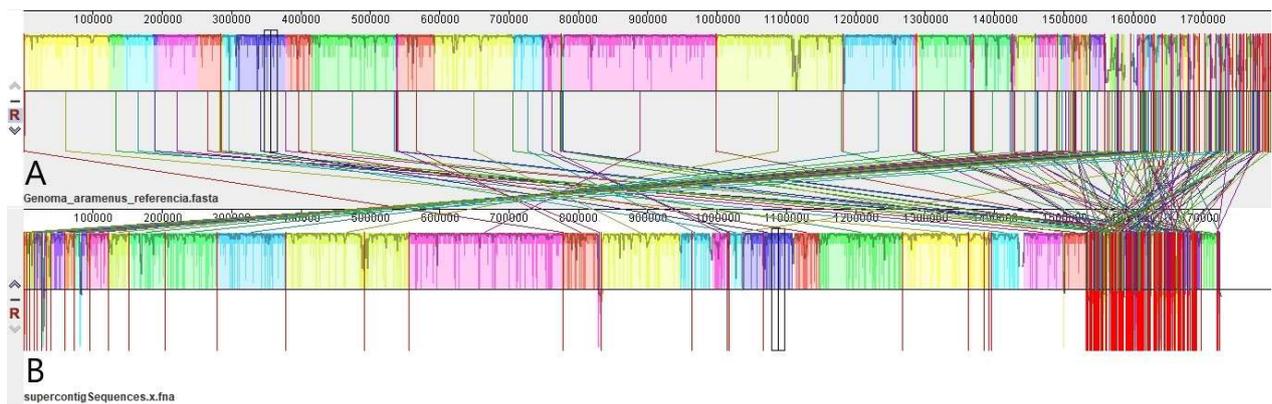


Figura 19: Alineamiento global realizado mediante MAUVE entre los genomas de *Ca. Aramenus sulfurataquae*. A) Genoma de *Ca. Aramenus sulfurataquae* AZ1. B) Genoma parcial de *Ca. Aramenus sulfurataquae* AZ2.

Anotación funcional del genoma parcial de *Ca. Aramenus sulfurataquae* AZ2

En la anotación funcional de la arquea *Ca. Aramenus sulfurataquae* AZ2 se encontraron un total de 2,151 genes que codifican proteínas. De los cuales solo 1,553 genes tienen una función predicha contra los 598 genes que no tienen función conocida. En la Tabla 7 se muestran 1643 genes agrupados en el “Cluster of Orthologous Groups” (COGs).

Tabla 7: Genes asociados a categorías funcionales COGs (“*Cluster of orthologous groups*”) del genoma de *Ca. Aramenus sulfurataquae AZ2* (Chen *et al.*, 2019).

Categoría COG	Número de genes	Porcentaje
Biosíntesis de metabolitos secundarios, transporte y catabolismo	28	1.46%
Control de ciclo celular, división celular, partición cromosómica	17	0.89%
Dinámica y estructura de cromatina	3	0.16%
Estructuras extracelulares	2	0.1%
Función desconocida	113	5.91%
Mecanismos de defensa	46	2.4%
Mecanismos de transducción de señales	47	2.46%
Mobiloma: profago, transposones	1	0.05%
Modificación post-transduccional, recambio proteico, chaperonas	87	4.55%
Modificación y procesamiento de ARN	1	0.03%
Motilidad celular	16	0.84%
Genes no agrupados en COGs	560	25.42%
Pared celular/membrana/sobre biogénesis	56	2.93%
Producción y conversión de energía	152	7.95%
Replicación, recombinación y reparación	72	3.76%
Solo predicción de función general	229	11.97%
Traducción, biogénesis y estructura de ribosomas	223	12.18%
Tráfico intracelular, secreción y transporte vesicular	14	0.73%
Transcripción	100	5.23%
Transporte y metabolismo de aminoácidos	208	10.87%
Transporte y metabolismo de carbohidratos	107	5.59%
Transporte y metabolismo de coenzimas	136	7.11%
Transporte y metabolismo de iones inorgánicos	106	5.54%
Transporte y metabolismo de lípidos	61	3.19%
Transporte y metabolismo de nucleótidos	78	4.08%

Se encontraron todos los genes para llevar a cabo la fijación del carbono mediante el ciclo 3-hidroxipropionato/4-hidroxi butirato. Aunque la secuencia del gen que codifica la enzima fosfoenolpiruvato carboxilasa se encuentra en fragmentos de secuencia. Esta enzima lleva a cabo la siguiente reacción: fosfoenolpiruvato + GDP + CO₂ --> Oxalacetato + GTP, el oxalacetato producido entra al ciclo de Krebs.

Todas las enzimas para realizar el proceso de glucólisis se encuentran presentes en el genoma. Para el ciclo de Krebs no se encontró la enzima malato deshidrogenasa la cual lleva la reacción de oxidación de malato a oxalacetato. Se encuentran presentes todos los elementos para realizar el proceso de fosforilación oxidativa.

En la síntesis de nucleótidos se encontraron todas las enzimas involucradas en la síntesis de pirimidinas. Sin embargo, en la síntesis de purinas, no se encontró la presencia del gen IMP ciclohidrolasa que lleva a cabo la síntesis de IMP a partir del ribonucleótido 5-formaminoimidazol-4-carboxamida (FAICAR). Respecto a la biosíntesis de aminoácidos el genoma parcial de *Ca. Aramenus sulfurataquae* AZ2 no cuenta con varios genes involucrados en la biosíntesis de lisina, isoleucina, leucina, histidina y valina. El genoma de *Ca. Aramenus sulfurataquae* tiene la capacidad de llevar a cabo los procesos de replicación, transcripción y traducción.

En el metabolismo del azufre se encuentran presentes los genes de cisteína sintetasa que realiza la síntesis de cisteína a partir de sulfuro de hidrógeno y la enzima azufre oxigenasa/reductasa que lleva a cabo la siguiente reacción: $4 S + H_2O + O_2 \rightarrow 2 H_2S + 2 SO_3^{2-}$. En el metabolismo del nitrógeno se encontraron los genes de glutamina sintetasa y glutamato deshidrogenasa. Estos genes están involucrados en la síntesis de glutamina y glutamato a partir de amoníaco.

El genoma parcial de *Ca. Aramenus sulfurataquae* contiene transportadores tipo ABC-2. Sin embargo, no se encontró que tipo de sustrato transportan. El genoma también contiene genes que

codifican para la superfamilia de facilitadores mayores (MFS). Dentro del genoma también se encontró un gen que codifica para un transportador de citrato y/o una proteína de resistencia a arsénico ArsB. Un gen que codifica para un transportador tipo ABC dipéptido/oligopéptido/níquel. Existen otros transportadores tales como: transportador ABC de antibióticos, transportador permeasa ABC para péptidos, permeasa de aminoácidos, transportador de alta afinidad de níquel, transportador ABC de azúcares, transportador de molibdeno ModA, transportador ATPasa de cationes, transportador antiporter sodio/protón, transportador ABC de nitrato y un transportador de fosfatos.

Secuencias de virus de arqueas *Sulfolobales*

No se encontraron secuencias de virus integrados en el genoma parcial de *Ca. Aramenus sulfurataquae* recuperado en el presente trabajo.

Discusión de los resultados de sedimentos fotosintéticos

A la fecha se ha analizado la diversidad microbiana mediante la identificación de genes ribosomales presentes en un metagenoma de sedimentos fotosintéticos de Los Azufres (Servin-Garcidueñas 2015). Las comunidades microbianas de los sedimentos fotosintéticos que se han estudiado en Los Azufres están integradas por bacterias, arqueas y eucariotas. Se han detectado bacterias pertenecientes a los filos Proteobacteria, Actinobacteria, Nitrospira y Firmicutes. Respecto a eucariontes, se han detectado secuencias de organismos pertenecientes a los filos Euglenozoa y Ascomycota. También se encontró la presencia abundante de microalgas del género *Cyanidiaceae*. En cuanto a arqueas, se encontraron diversos linajes (linajes G, E, A, C, I) pertenecientes al orden *Thermoplasmatales*. Además de una arquea del filo Micrarchaeota (Servin-Garcidueñas 2015). Hasta ahora se ha podido reportar el genoma de Micrarchaeota dentro de un análisis de genómica comparativa para determinar el metabolismo potencial de este filo de arqueas (Chen *et al.*, 2018). No se han obtenido otros genomas individuales de arqueas presentes en el metagenoma de sedimentos fotosintéticos. Por lo que, el presente trabajo se enfocó en obtener el genoma de una arquea perteneciente al género *Candidatus Cuniculiplasma* cuyas secuencias genómicas son abundantes en el metagenoma de sedimentos fotosintéticos.

El genoma parcial de *Ca. Cuniculiplasma* sp. AZ01 que se obtuvo en el presente es el cuarto genoma reportado para el género *Cuniculiplasma*. Los genomas de *Cuniculiplasma divulgatum* S5 y PM4 fueron obtenidos de aguas ácidas de minas de cobre de Canteras, España y de Mynydd Parys, Anglesey, Reino Unido (Golyshina, Kublanov, *et al.*, 2016; Golyshina, Lünsdorf, *et al.*, 2016). El genoma de *Cuniculiplasma* sp. C_DKE fue obtenido a partir de un enriquecimiento de una muestra de biofilm en las montañas Harz, Alemania (Krause *et al.*, 2017).

El genoma parcial de *Ca. Cuniculiplasma* sp. AZ01 tiene un contenido de GC igual al de los genomas de *Cuniculiplasma divulgatum* PM4 y S5 (Tabla 1). El tamaño del genoma de *Ca.*

Cuniculiplasma sp. AZ01 es menor al de los genomas de *Cuniculiplasma divulgatum* PM4 y S5 probablemente a que el genoma de *Ca. Cuniculiplasma* sp. AZ01 se encuentra en un estado parcial (Tabla 1). De acuerdo con los parámetros de identidad genómica se encontró que el genoma de *Ca. Cuniculiplasma* sp. AZ01 es altamente conservado con respecto a los genomas de *Cuniculiplasma divulgatum* PM4 y S5 (Tabla 1). El genoma de *Ca. Cuniculiplasma* sp. AZ01 no presenta genes únicos en comparación con los genomas de *Cuniculiplasma divulgatum* PM4 y S5.

Se lograron anotar 1,472 genes que codifican proteínas con función conocida y 584 genes que codifican proteínas sin función conocida. Mayormente se encontraron genes involucrados en la traducción, biogénesis y estructura de los ribosomas; transporte y metabolismo de aminoácidos; producción y conversión de energía. *Ca. Cuniculiplasma* sp. AZ01 carece de las enzimas 2-oxoglutarato deshidrogenasa y fumarasa. La ausencia de estas enzimas también se reporta para el genoma de *Cuniculiplasma divulgatum* (Golyshina, Kublanov, *et al.*, 2016). A pesar de la ausencia de la enzima 2-oxoglutarato deshidrogenasa *Cuniculiplasma* es capaz de realizar la descarboxilación del 2-oxoglutarato a succinato. La reacción de descarboxilación en *Cuniculiplasma* es llevada a cabo mediante la enzima 2-oxoácido: aceptor oxidoreductasa subunidad alfa y beta. Sin embargo, no existe alguna enzima capaz de realizar la reacción de adición de agua (H₂O) del fumarato a malato en los genomas de *Ca. Cuniculiplasma* sp. AZ01 y *Cuniculiplasma divulgatum* (Golyshina, Kublanov, *et al.*, 2016). En *Ca. Cuniculiplasma* sp. AZ01 se encontraron los genes involucrados en la síntesis de nucleótidos (purinas y pirimidinas). De modo que podría ser posible que *Ca. Cuniculiplasma* sp. AZ01 este compartiendo nucleótidos con la Micrarchaeota reportada previamente (Chen *et al.*, 2018).

Por último, se encontraron una gran variedad de proteínas capaces de transportar una gran variedad de nutrientes dentro y fuera de la célula. *Ca. Cuniculiplasma* sp. AZ01 podría estar incorporando nutrientes producidos por la microalga del género *Cyanidiaceae* que se encontró presente en el metagenoma de sedimentos fotosintéticos. Golyshina y colaboradores (2019)

mencionan que coincidentemente donde se aisló *Cuniculiplasma* también se encontró la presencia de algas acidófilas, las cuales podrían estar aportando nutrientes a *Cuniculiplasma*.

Discusión de los resultados de sedimentos no fotosintéticos

En este trabajo se secuenció el primer metagenoma de sedimentos no fotosintéticos del campo geotérmico de Los Azufres. La diversidad microbiana encontrada en los sedimentos no fotosintéticos estuvo representada principalmente por especies de arqueas del orden *Sulfolobales*. En las secuencias metagenómicas se detectó la presencia de cuatro poblaciones de arqueas del orden *Sulfolobales*. Las cuatro poblaciones presentes en el metagenoma fueron *Ca. Aramenus sulfurataquae*, *Acidianus brierleyi* y dos arqueas del género *Sulfolobus*. Es común encontrar arqueas del orden *Sulfolobales* en ambientes termales ricos en azufre. La presencia de arqueas del orden *Sulfolobales* se debe probablemente a la abundancia de azufre y a las condiciones ambientales de temperaturas superiores a 90 °C y pH ácidos que presentan este tipo de sedimentos de Los Azufres. Las arqueas *Sulfolobales* se han aislado exclusivamente de campos de solfataras. Se han aislado cepas del género *Sulfolobus* en varios lugares geográficos, incluidos algunos ubicados en los Estados Unidos de América (Brock *et al.* 1972; Brock 1978), China (Xiang *et al.*, 2003; Luo *et al.* 2015) Islandia (Brock 1978; Huber & Stetter 1991), Italia (de Rosa *et al.* 1974; Brock 1978; Zillig *et al.* 1980), Japón (Yeats *et al.*, 1982; Jan *et al.* 1999; Suzuki *et al.* 2002;), Nueva Zelanda (Bohlool, 1975) y las Indias Occidentales (Brock, 1978).

La anotación del metagenoma de sedimentos no fotosintéticos logró predecir 13,413 secuencias de genes codificantes de proteínas. También fue posible la identificación de genes involucrados en la fijación del carbono y en los metabolismos del azufre y del nitrógeno.

Dentro del metagenoma también se encontraron genomas parciales de virus que presentaron identidad contra los genomas de los virus “*Sulfolobales* Mexican Rudivirus 1”, “*Sulfolobales* Mexican

Fusellovirus 1” y “Los azufres archaeal virus 1 y 2”. Estos genomas fueron recuperados previamente de Los Azufres a partir de muestras de manantiales termales de Los Azufres. Las akeas de los géneros *Sulfolobus*, *Acidianus* y *Stygiolobus* son los organismos hospederos de *Rudivirus* (King *et al.*, 2012b). Los hospederos de *Fusellovirus* se encuentran limitados a arqueas tales como: *Sulfolobus shibatae*, *Sulfolobus solfataricus* y *Sulfolobus islandicus* (King *et al.*, 2012a). *Rudivirus* y *Fusellovirus* presentan una estrecha relación con arqueas del orden *Sulfolobales*.

A partir de las lecturas del metagenoma de sedimentos no fotosintéticos fue posible recuperar un tercer genoma de la arquea *Ca. Aramenus sulfurataquae*. Además de la distancia entre los tres sitios de colecta, también se presentan diferencias en condiciones ambientales (Figura 17). La arquea *Ca. Aramenus sulfurataquae* se ha detectado únicamente en la zona del campo geotérmico de Los Azufres. Considerando que el genoma de *Ca. Aramenus sulfurataquae* AZ2 se encuentra en estado parcial, su contenido de GC y tamaño del genoma es similar al del genoma de *Ca. Aramenus sulfurataquae* AZ1 (Tabla 6). De acuerdo a los parámetros de identidad genómica se encontró que el genoma de *Ca. Aramenus sulfurataquae* AZ2 se encuentra conservado respecto al genoma de *Ca. Aramenus sulfurataquae* AZ1 (Tabla 6). No se encontraron genes únicos presentes en el genoma de *Ca. Aramenus sulfurataquae* AZ2. En consecuencia, *Ca. Aramenus sulfurataquae* es capaz de sobrevivir en diferentes tipos de ambientes contrastantes (acuoso y no acuoso) con diferentes condiciones de temperatura (65 a 89 °C) y de pH (2.8 a 3.7). Proponemos que *Ca. Aramenus sulfurataquae* es una especie con un metabolismo versátil y con las adaptaciones necesarias para sobrevivir en ambientes termales con condiciones fisicoquímicas diversas de Los Azufres.

Mediante la anotación funcional del genoma parcial de la arquea *Ca. Aramenus sulfurataquae* AZ2 se encontraron 1533 proteínas con función predica y 598 proteínas que no tienen una función conocida. Mayormente se encontraron genes involucrados en la traducción, biogénesis y estructura de los ribosomas; transporte y metabolismo de aminoácidos; transporte y metabolismo de coenzimas.

Dentro del genoma parcial de la arquea *Ca. Aramenus sulfurataquae* AZ2 se encuentran presentes todos los genes involucrados en la fijación del carbono en procariontes. Como se mencionó anteriormente las arqueas *Sulfolobales* contribuyen al ciclo del carbono como productores primarios y consumidores de material orgánico (Dworkin *et al.* 2006; Boone 2013). Además, se encontraron genes involucrados en el metabolismo del azufre. Estos genes les permiten a las arqueas producir ácido sulfúrico, agua y sulfuro de hidrógeno que es liberado al medio ambiente (Kämpfer, 2006).

Conclusiones

Sedimentos fotosintéticos.

- La diversidad de arqueas presentes en el metagenoma de sedimentos fotosintéticos está integrada por arqueas del orden *Thermoplasmatales* y por una arquea del filo Micrarchaeota. Se detectó la presencia de microalgas del género *Cyanidiaceae* y de hongos acidófilos. Por último, se encontraron bacterias pertenecientes a los filos Proteobacteria, Actinobacteria, Nitrospira y Firmicutes
- Se recuperó el genoma parcial de la arquea *Ca. Cuniculiplasma* sp. AZ01 que representa el cuarto genoma del género *Cuniculiplasma* en ser reportado. Este es el primer genoma de una arquea *Thermoplasmatales* en ser obtenido a partir de datos metagenómicos de Los Azufres. El genoma de *Ca. Cuniculiplasma* sp. AZ01 se encuentra casi completo, con buena cobertura y será útil para realizar análisis de genómica comparativa.
- *Ca. Cuniculiplasma* sp. AZ01 presenta los genes involucrados en la síntesis de nucleótidos, por lo que podría existir una transferencia de nucleótidos hacia Micrarchaeota.
- *Ca. Cuniculiplasma* sp. AZ01 presenta una abundancia de proteínas transportadoras. Estas proteínas podrían transportar una variedad amplia de nutrientes producidos por microalgas.
- El genoma de *Ca. Cuniculiplasma* sp. AZ01 se encuentra conservado respecto a los genomas de *Cuniculiplasma divulgatum* PM4 y S5. La identidad genómica entre los genomas de *Cuniculiplasma* es alta y en general son genomas idénticos. Su potencial genético también se encuentra conservado.

Sedimentos no fotosintéticos.

- En este trabajo se secuenció el primer metagenoma de sedimentos no fotosintéticos del campo geotérmico de Los Azufres. Este ha sido el primer metagenoma secuenciado con la infraestructura del Laboratorio Nacional de Síntesis Ecológica (LANASE), de la Escuela Nacional de Estudios Superiores Unidad Morelia (ENES Morelia), y del Laboratorio de Microbiómica. El secuenciador NextSeq 500 de Illumina con el que cuenta el LANASE es apto para explorar el contenido tanto de genomas individuales como de metagenomas. Las lecturas metagenómicas generadas a partir del secuenciador del LANASE son de un tamaño de 75 pb y al ser lecturas tan pequeñas se dificulta el obtener genomas individuales completos de microorganismos de interés a partir de datos metagenómicos.
- La diversidad microbiana presente en el metagenoma de sedimentos no fotosintéticos está dominada por arqueas del orden *Sulfolobales*.
- La arqueas *Sulfolobales* encontradas en el metagenoma de sedimentos no fotosintéticos desempeñan un rol importante en la fijación del carbono, metabolismo del azufre y metabolismo del nitrógeno.
- Se encontró la presencia de secuencias genómicas novedosas de virus de arqueas *Sulfolobales* de Los Azufres.
- Se obtuvo un tercer genoma de la arquea *Ca. Aramenus sulfurataquae*. *Ca. Aramenus sulfurataquae* podría estar adaptada para convivir con diferentes comunidades de microorganismos y adaptarse a las diferentes condiciones ambientales en las que se encuentra presente.
- La arquea *Ca. Aramenus sulfurataquae* podría estar utilizando el dióxido de carbono como fuente principal de carbono, por lo que podría desempeñar un rol importante en la fijación del carbono de los sedimentos termales de Los Azufres.

Perspectivas

- Se propone cultivar a las diferentes arqueas presentes en Los Azufres con énfasis en la arquea *Ca. Aramenus sulfurataquae* y *Ca. Cuniculiplasma* sp. AZ01. Tener cultivadas a estas arqueas nos permitiría obtener la secuenciación de genomas de cepas individuales que facilitarían los análisis de potenciales genéticos y de genómica comparativa.
- Se sugiere realizar análisis biogeoquímicos para determinar el rol biológico de las arqueas *Sulfolobales* en los ciclos de carbono y azufre.
- Se recomienda llevar a cabo análisis de metatranscriptómica para facilitar la anotación de genomas y para determinar que genes se están expresando en las diferentes arqueas y bajo qué condiciones. A partir de estos datos podremos analizar el potencial impacto que tienen en sus funciones en el medio ambiente o en la comunidad microbiana a la que pertenecen.
- Tratar de obtener imágenes mediante hibridación fluorescente in situ (FISH) de las arqueas. Esto nos permitirá explorar las relaciones que presentan las arqueas *Thermoplasmatales* con las arqueas *Micrarchaeota*.
- Sería importante realizar secuenciaciones masivas de ADN mediante la plataforma de secuenciación PacBio o la tecnología NanoPore para obtener genomas individuales completos y cerrados de arqueas.
- Finalmente se propone continuar con la recuperación de genomas de otras arqueas presentes en los metagenomas analizados en el presente trabajo. Así mismo se recomienda continuar con los análisis de otros metagenomas de sedimentos de Los Azufres que ya fueron secuenciados (Anexo) con el objetivo de determinar la distribución de arqueas dentro del campo geotérmico de Los Azufres.

Bibliografía

- Alfaro-Cuevas-Villanueva, R., Cortes-Martinez, R., García-Díaz, J. J., Galvan-Martinez, R., & Torres-Sanchez, R. (2006). Microbiologically influenced corrosion of steels by thermophilic and mesophilic bacteria. *Materials and Corrosion*, 57(7), 543–548. <https://doi.org/10.1002/maco.200503948>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Arriaga-Suarez, M. c., Tello-Lopez, M. R. ocí., & Samaniego, F. (2000). GEOCHEMICAL EVOLUTION OF THE LOS AZUFRES , MEXICO , GEOTHERMAL RESERVOIR . Part II : NON-CONDENSIBLE GASES. *Proceedings World Geothermal Congress. Kyushu-Tohoku, Japan.*, 2227–2233.
- Aubourg, S., & Rouzé, P. (2001). Genome annotation. *Plant Physiology and Biochemistry*, 39(3–4), 181–193. [https://doi.org/10.1016/S0981-9428\(01\)01242-6](https://doi.org/10.1016/S0981-9428(01)01242-6)
- Aziz, R. K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R. A., ... Zagnitko, O. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*, 9. <https://doi.org/10.1186/1471-2164-9-75>
- Baker, B. J., Comolli, L. R., Dick, G. J., Hauser, L. J., Hyatt, D., Dill, B. D., ... Banfield, J. F. (2010). Enigmatic, ultrasmall, uncultivated Archaea. *Proceedings of the National Academy of Sciences of the United States of America*, 107(19), 8806–8811. <https://doi.org/10.1073/pnas.0914470107>
- Baker, B. J., Tyson, G. W., Webb, R. I., Flanagan, J., Hugenholtz, P., Allen, E. E., & Banfield, J. F. (2006). Lineages of acidophilic archaea revealed by community genomic analysis. *Science*, 314(5807), 1933–1935. <https://doi.org/10.1126/science.1132690>
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, 19(5), 455–477. <https://doi.org/10.1089/cmb.2012.0021>
- Blohs, M., Mahnert, A., Spang, A., Dombrowski, N., Krupovic, M., & Klingl, A. (2019). Archaea – An Introduction. In T. M. B. T.-E. of M. (Fourth E. Schmidt (Ed.), *Reference Module in Life Sciences* (pp. 243–252). Oxford: Academic Press. <https://doi.org/10.1016/b978-0-12-809633-8.20884-4>
- Bohlool, B. Ben. (1975). Occurrence of Sulfolobus acidocaldarius, an extremely thermophilic acidophilic bacterium, in New Zealand hot springs - Isolation and immunofluorescence characterization. *Archives of Microbiology*, 106(3), 171–174. <https://doi.org/10.1007/BF00446520>
- Boone, D. (2013). Bergey’s Manual of Systematic Bacteriology—Second Edition-vol. One. *The Archaea and the Deeply Branching and ...*
- Brito, E. M. S., Villegas-Negrete, N., Sotelo-González, I. A., Caretta, C. A., Goñi-Urriza, M., Gassie, C., ... Guyoneaud, R. (2014). Microbial diversity in Los Azufres geothermal field (Michoacán, Mexico) and isolation of representative sulfate and sulfur reducers. *Extremophiles*, 18(2), 385–398. <https://doi.org/10.1007/s00792-013-0624-7>
- Brock, T. D. (1978). The Genus Sulfolobus. In T. D. Brock (Ed.) (pp. 117–179). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4612-6284-8_6
- Brock, T. D., Brock, K. M., Belly, R. T., & Weiss, R. L. (1972). Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Archiv Für Mikrobiologie*, 84(1), 54–68. <https://doi.org/10.1007/BF00408082>
- Castelle, C. J., & Banfield, J. F. (2018). Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. *Cell*. <https://doi.org/10.1016/j.cell.2018.02.016>
- Castorena, G., Mugica, V., Le Borgne, S., Acuña, M. E., Bustos-Jaimes, I., & Aburto, J. (2006). Carbazole biodegradation in gas oil/water biphasic media by a new isolated bacterium Burkholderia sp. strain IMP5GC. *Journal of Applied Microbiology*, 100(4), 739–745. <https://doi.org/10.1111/j.1365-2672.2005.02799.x>
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17(4), 540–552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- Cavicchioli, R. (2011). Archaea - Timeline of the third domain. *Nature Reviews Microbiology*. <https://doi.org/10.1038/nrmicro2482>
- Chaffron, S., Rehrauer, H., Pernthaler, J., & Von Mering, C. (2010). A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Research*, 20(7), 947–959. <https://doi.org/10.1101/gr.104521.109>
- Chen, I. M. A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., ... Kyrpides, N. C. (2019). IMG/M v.5.0: An integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Research*, 47(D1), D666–D677. <https://doi.org/10.1093/nar/gky901>
- Chen, L. X., Méndez-García, C., Dombrowski, N., Servín-Garcidueñas, L. E., Eloë-Fadrosh, E. A., Fang, B. Z., ... Shu, W. S. (2018). Metabolic versatility of small archaea Micrarchaeota and Parvarchaeota. *ISME Journal*, 12(3), 756–775. <https://doi.org/10.1038/s41396-017-0002-z>

- Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S., ... DiCuccio, M. (2018). Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *International Journal of Systematic and Evolutionary Microbiology*, 68(7), 2386–2392. <https://doi.org/10.1099/ijsem.0.002809>
- Clark, D. P., Pazdernik, N. J., & McGehee, M. R. (2019). Chapter 8 - DNA Sequencing. In D. P. Clark, N. J. Pazdernik, & M. R. B. T.-M. B. (Third E. McGehee (Eds.) (pp. 240–269). Academic Cell. <https://doi.org/https://doi.org/10.1016/B978-0-12-813288-3.00008-2>
- Darland, G., Brock, T. D., Samsonoff, W., & Conti, S. F. (1970). A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile. *Science*, 170(3965), 1416–1418. <https://doi.org/10.1126/science.170.3965.1416>
- Darling, A. C. E., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Research*, 14(7), 1394–1403. <https://doi.org/10.1101/gr.2289704>
- DasSarma, S., Coker, J. A., & DasSarma, P. (2009). Archaea (overview). *Encyclopedia of Microbiology*, 1–23. <https://doi.org/10.1016/B978-012373944-5.00108-5>
- de Rosa, M., Gambacorta, A., Millonig, G., & Bu'Lock, J. D. (1974). Convergent characters of extremely thermophilic acidophilic bacteria. *Experientia*, 30(8), 866–868. <https://doi.org/10.1007/BF01938328>
- Demant, A. (1978). Características del Eje Neovolcánico Transmexicano y sus problemas de interpretación. *Revista Mexicana de Ciencias Geológicas*, 2(2), 172–187.
- Dombrowski, N., Lee, J. H., Williams, T. A., Offre, P., & Spang, A. (2019). Genomic diversity, lifestyles and evolutionary origins of DPANN archaea. *FEMS Microbiology Letters*, 366(2). <https://doi.org/10.1093/femsle/fnz008>
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Golovacheva, R. S., Valiejo-Roman, K. M., & Troitsky, A. V. (1987). *Sulfurococcus mirabilis* gen. nov. sp. nov., a new thermophilic archaeobacterium oxidizing sulfur. *Mikrobiologiya*, 56(1), 100–107.
- Golyshina, O. V., Bargiela, R., & Golyshin, P. N. (2019). Cuniculiplasmataceae, their ecogenomic and metabolic patterns, and interactions with 'ARMAN.' *Extremophiles*. <https://doi.org/10.1007/s00792-018-1071-2>
- Golyshina, O. V., Kublanov, I. V., Tran, H., Korzhenkov, A. A., Lünsdorf, H., Nechitaylo, T. Y., ... Golyshin, P. N. (2016). Biology of archaea from a novel family Cuniculiplasmataceae (Thermoplasmata) ubiquitous in hyperacidic environments. *Scientific Reports*, 6. <https://doi.org/10.1038/srep39034>
- Golyshina, O. V., Lünsdorf, H., Kublanov, I. V., Goldenstein, N. I., Hinrichs, K. U., & Golyshin, P. N. (2016). The novel extremely acidophilic, cell-wall-deficient archaeon Cuniculiplasma divulgatum gen. nov., sp. nov. represents a new family, Cuniculiplasmataceae fam. nov., of the order Thermoplasmatales. *International Journal of Systematic and Evolutionary Microbiology*, 66(1), 332–340. <https://doi.org/10.1099/ijsem.0.000725>
- Golyshina, O. V., Pivovarova, T. A., Karavaiko, G. I., Kondrat'eva, T. F., Moore, E. R. B., Abraham, W. R., ... Golyshin, P. N. (2000). *Ferroplasma acidiphilum* gen. nov., sp. nov., an acidophilic, autotrophic, ferrous-iron-oxidizing, cell-wall-lacking, mesophilic member of the Ferroplasmaceae fam. nov., comprising a distinct lineage of the Archaea. *International Journal of Systematic and Evolutionary Microbiology*, 50(3), 997–1006. <https://doi.org/10.1099/00207713-50-3-997>
- Golyshina, O. V., Toshchakov, S. V., Makarova, K. S., Gavrillov, S. N., Korzhenkov, A. A., La Cono, V., ... Golyshin, P. N. (2017). "ARMAN" archaea depend on association with euryarchaeal host in culture and in situ. *Nature Communications*, 8(1). <https://doi.org/10.1038/s41467-017-00104-7>
- Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., & Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*, 57(1), 81–91. <https://doi.org/10.1099/ijms.0.64483-0>
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010a). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321. <https://doi.org/10.1093/sysbio/syq010>
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010b). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321. <https://doi.org/10.1093/sysbio/syq010>
- Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUASt: Quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
- Huber, G., Spinnler, C., Gambacorta, A., & Stetter, K. O. (1989). *Metallosphaera sedula* gen. nov. sp. nov. Represents a New Genus of Aerobic, Metal-Mobilizing, Thermoacidophilic Archaeobacteria. *Systematic and Applied Microbiology*, 12(1), 38–47. [https://doi.org/10.1016/S0723-2020\(89\)80038-4](https://doi.org/10.1016/S0723-2020(89)80038-4)
- Huber, G., & Stetter, K. O. (1991). *Sulfolobus metallicus*, sp. nov., a Novel Strictly Chemolithoautotrophic Thermophilic Archaeal Species of Metal-Mobilizers. *Systematic and Applied Microbiology*, 14(4), 372–378. [https://doi.org/10.1016/S0723-2020\(11\)80312-7](https://doi.org/10.1016/S0723-2020(11)80312-7)
- Huber, R., Huber, H., & Stetter, K. O. (2000). Towards the ecology of hyperthermophiles: Biotopes, new isolation strategies and novel metabolic properties. *FEMS Microbiology Reviews*. <https://doi.org/10.1016/S0168->

- Huntemann, M., Ivanova, N. N., Mavromatis, K., Tripp, H. J., Paez-Espino, D., Tennessen, K., ... Kyrpides, N. C. (2016). The standard operating procedure of the DOE-JGI Metagenome Annotation Pipeline (MAP v.4). *Standards in Genomic Sciences*, 11(1). <https://doi.org/10.1186/s40793-016-0138-x>
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., & Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-018-07641-9>
- Jan, R. L., Wu, J., Chaw, S. M., Tsai, C. W., & Tsen, S. Der. (1999). A novel species of thermoacidophilic archaeon, *Sulfolobus yangmingensis* sp. nov. *International Journal of Systematic Bacteriology*, 49(4), 1809–1816. <https://doi.org/10.1099/00207713-49-4-1809>
- Kämpfer, P. (2006). *The Family Streptomycetaceae, Part I: Taxonomy BT - The Prokaryotes*. (M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, & E. Stackebrandt, Eds.), *The Prokaryotes*. New York, NY: Springer New York. <https://doi.org/10.1007/0-387-30743-5>
- Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., ... Subhraveti, P. (2018). The BioCyc collection of microbial genomes and metabolic pathways. *Briefings in Bioinformatics*, 20(4), 1085–1093. <https://doi.org/10.1093/bib/bbx085>
- Keegan, K. P., Glass, E. M., & Meyer, F. (2016). MG-RAST, a metagenomics service for analysis of microbial community structure and function. In *Methods in Molecular Biology* (Vol. 1399, pp. 207–233). https://doi.org/10.1007/978-1-4939-3369-3_13
- Kerepesi, C., Bánky, D., & Grolmusz, V. (2014). AmphoraNet: The webserver implementation of the AMPHORA2 metagenomic workflow suite. *Gene*, 533(2), 538–540. <https://doi.org/10.1016/j.gene.2013.10.015>
- Khodakov, D., Wang, C., & Zhang, D. Y. (2016). Diagnostics based on nucleic acid sequence variant profiling: PCR, hybridization, and NGS approaches. *Advanced Drug Delivery Reviews*. <https://doi.org/10.1016/j.addr.2016.04.005>
- King, A. M. Q., Adams, M. J., Carstens, E. B., & Lefkowitz, E. J. (Eds.). (2012a). Fuselloviridae. In *Virus Taxonomy* (pp. 183–186). San Diego: Elsevier. <https://doi.org/10.1016/b978-0-12-384684-6.00016-1>
- King, A. M. Q., Adams, M. J., Carstens, E. B., & Lefkowitz, E. J. (Eds.). (2012b). Rudiviridae. In *Virus Taxonomy* (pp. 311–315). San Diego: Elsevier. <https://doi.org/10.1016/b978-0-12-384684-6.00029-x>
- Kletzin, A., Urich, T., Müller, F., Bandejas, T. M., & Gomes, C. M. (2004). Dissimilatory Oxidation and Reduction of Elemental Sulfur in Thermophilic Archaea. *Journal of Bioenergetics and Biomembranes*. <https://doi.org/10.1023/B:JOB.0000019600.36757.8c>
- Koga, Y., & Morii, H. (2007). Biosynthesis of Ether-Type Polar Lipids in Archaea and Evolutionary Considerations. *Microbiology and Molecular Biology Reviews*, 71(1), 97–120. <https://doi.org/10.1128/mmbr.00033-06>
- Konstantinidis, K. T., & Tiedje, J. M. (2005a). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7), 2567–2572. <https://doi.org/10.1073/pnas.0409727102>
- Konstantinidis, K. T., & Tiedje, J. M. (2005b). Towards a genome-based taxonomy for prokaryotes. *Journal of Bacteriology*, 187(18), 6258–6264. <https://doi.org/10.1128/JB.187.18.6258-6264.2005>
- Krause, S., Bremges, A., Münch, P. C., McHardy, A. C., & Gescher, J. (2017). Characterisation of a stable laboratory co-culture of acidophilic nanoorganisms. *Scientific Reports*, 7(1). <https://doi.org/10.1038/s41598-017-03315-6>
- Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K., & Hugenholtz, P. (2008). A Bioinformatician's Guide to Metagenomics. *Microbiology and Molecular Biology Reviews*, 72(4), 557–578. <https://doi.org/10.1128/mmbr.00009-08>
- Kurosawa, N., Itoh, Y. H., Iwai, T., Sugai, A., Uda, I., Kimura, N., ... Itoh, T. (1998). *Sulfurisphaera ohwakuensis* gen. nov., sp. nov., a novel extremely thermophilic acidophile of the order Sulfolobales. *International Journal of Systematic Bacteriology*, 48(2), 451–456. <https://doi.org/10.1099/00207713-48-2-451>
- Lagesen, K., Hallin, P., Rødland, E. A., Stærfeldt, H. H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, 35(9), 3100–3108. <https://doi.org/10.1093/nar/gkm160>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Leigh, J. A., Albers, S. V., Atomi, H., & Allers, T. (2011). Model organisms for genetics in the domain Archaea: Methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiology Reviews*, 35(4), 577–608. <https://doi.org/10.1111/j.1574-6976.2011.00265.x>
- Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., ... Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, 47(W1), W636–W641. <https://doi.org/10.1093/nar/gkz268>
- Magoc, T., & Salzberg, S. L. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, 27(21), 2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>
- Medlar, A. J., Törönen, P., & Holm, L. (2018). AAI-profiler: Fast proteome-wide exploratory analysis reveals

- taxonomic identity, misclassification and contamination. *Nucleic Acids Research*, 46(W1), W479–W485. <https://doi.org/10.1093/nar/gky359>
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, 14(1), 60. <https://doi.org/10.1186/1471-2105-14-60>
- Mikheenko, A., Saveliev, V., & Gurevich, A. (2016). MetaQUAST: Evaluation of metagenome assemblies. *Bioinformatics*, 32(7), 1088–1090. <https://doi.org/10.1093/bioinformatics/btv697>
- Moissl-Eichinger, C., Pausan, M., Taffner, J., Berg, G., Bang, C., & Schmitz, R. A. (2018). Archaea Are Interactive Components of Complex Microbiomes. *Trends in Microbiology*, 26(1), 70–85. <https://doi.org/10.1016/j.tim.2017.07.004>
- Navarrete-Bedolla, M., Ballesteros-Almanza, M. L., Sánchez-Yañez, J. M., Valdéz-Salas, B., & Hernández-Duque, G. (1999). Biocorrosion in a geothermal power plant. *Materials Performance*, 38(4).
- Paul, E. A., Killham, K., & Prosser, J. I. (2015). Chapter 3 – The Bacteria and Archaea. *Soil Microbiology, Ecology and Biochemistry*. <https://doi.org/10.1016/B978-0-12-415955-6.00003-7>
- Pham, V. H. T., & Kim, J. (2012). Cultivation of unculturable soil bacteria. *Trends in Biotechnology*. <https://doi.org/10.1016/j.tibtech.2012.05.007>
- Pruesse, E., Peplies, J., & Glöckner, F. O. (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*, 28(14), 1823–1829. <https://doi.org/10.1093/bioinformatics/bts252>
- Reddy, V. S., Shlykov, M. A., Castillo, R., Sun, E. I., & Saier, M. H. (2012). The major facilitator superfamily (MFS) revisited. *FEBS Journal*, 279(11), 2022–2035. <https://doi.org/10.1111/j.1742-4658.2012.08588.x>
- Rédei, G. P. (2008). Major Facilitator Superfamily (MFS). *Encyclopedia of Genetics, Genomics, Proteomics and Informatics*, 62(1), 1142–1142. https://doi.org/10.1007/978-1-4020-6754-9_9778
- Richter, D. C., Schuster, S. C., & Huson, D. H. (2007). OSLay: Optimal syntenic layout of unfinished assemblies. In *Bioinformatics* (Vol. 23, pp. 1573–1579). <https://doi.org/10.1093/bioinformatics/btm153>
- Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J. F., ... Woyke, T. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, 499(7459), 431–437. <https://doi.org/10.1038/nature12352>
- Rodríguez-R., L. M., & Konstantinidis, K. T. (2014). Bypassing cultivation to identify bacterial species. *Microbe*. <https://doi.org/10.1128/microbe.9.111.1>
- Rodríguez-R., L., & Konstantinidis, K. (2016). The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints*. <https://doi.org/10.7287/peerj.preprints.1900v1>
- Saier, M. H., Beatty, J. T., Goffeau, A., Harley, K. T., Heijne, W. H. M., Huang, S. C., ... Virk, P. S. (1999). The major facilitator superfamily. *Journal of Molecular Microbiology and Biotechnology*, 1(2), 257–279. <https://doi.org/10.1128/mmbr.62.1.1-34.1998>
- Sakai, H. D., & Kurosawa, N. (2018). *Saccharolobus caldissimus* gen. nov., sp. nov., a facultatively anaerobic iron-reducing hyperthermophilic archaeon isolated from an acidic terrestrial hot spring, and reclassification of *sulfolobus solfataricus* as *saccharolobus solfataricus* comb. nov. and *sulfolobus shibatae* as *saccharolobus shibatae* comb. nov. *International Journal of Systematic and Evolutionary Microbiology*, 68(4), 1271–1278. <https://doi.org/10.1099/ijsem.0.002665>
- Schleper, C., Puehler, G., Holz, I., Gambacorta, A., Janekovic, D., Santarius, U., ... Zillig, W. (1995). *Picrophilus* gen. nov., fam. nov.: A novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0. *Journal of Bacteriology*, 177(24), 7050–7059. <https://doi.org/10.1128/jb.177.24.7050-7059.1995>
- Segerer, A. H., Trincone, A., Gahrtz, M., & Stetter, K. O. (1991). *Stygiolobus azoricus* gen. nov., sp. nov. represents a novel genus of anaerobic, extremely thermoacidophilic archaeobacteria of the order Sulfolobales. *International Journal of Systematic Bacteriology*, 41(4), 495–501. <https://doi.org/10.1099/00207713-41-4-495>
- Segerer, A., Neuner, A., Kristjansson, J. K., & Stetter, K. O. (1986). *Acidianus infernus* gen. nov., sp. nov., and *Acidianus brierleyi* comb. nov.: Facultatively aerobic, extremely acidophilic thermophilic sulfur-metabolizing archaeobacteria. *International Journal of Systematic Bacteriology*, 36(4), 559–564. <https://doi.org/10.1099/00207713-36-4-559>
- Servín-Garcidueñas, L. E. (2015). Estudio metagenómico de microorganismos extremófilos del campo geotérmico de Los Azufres y diversidad genómica de simbioses de *Phaseolus* e insectos nativos de México. Universidad Nacional Autónoma de México.
- Servín-Garcidueñas, L. E., & Martínez-Romero, E. (2014). Draft genome sequence of the Sulfolobales archaeon AZ1, obtained through metagenomic analysis of a Mexican hot spring. *Genome Announcements*, 2(2). <https://doi.org/10.1128/genomeA.00164-14>
- Servín-Garcidueñas, L. E., Peng, X., Garrett, R. A., & Martínez-Romero, E. (2013a). Genome sequence of a novel archaeal fusellovirus assembled from the metagenome of a Mexican hot spring. *Genome Announcements*, 1(2). <https://doi.org/10.1128/genomeA.00164-13>

-
- Servín-Garcidueñas, L. E., Peng, X., Garrett, R. A., & Martínez-Romero, E. (2013b). Genome sequence of a novel archaeal rudivirus recovered from a Mexican hot spring. *Genome Announcements*, *1*(1). <https://doi.org/10.1128/genomeA.00040-12>
- Suzuki, T., Iwasaki, T., Uzawa, T., Hara, K., Nemoto, N., Kon, T., ... Oshima, T. (2002). *Sulfolobus tokodaii* sp. nov. (f. *Sulfolobus* sp. strain 7), a new member of the genus *Sulfolobus* isolated from Beppu Hot Springs, Japan. *Extremophiles*, *6*(1), 39–44. <https://doi.org/10.1007/s007920100221>
- Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, *56*(4), 564–577. <https://doi.org/10.1080/10635150701472164>
- Tello-lopez, Mirna, R., & Suarez-Arriaga, Mario, C. (2000). geochemical evolution of the los azufres, mexico, geothermal reservoir. part i: water and salts. *Proceedings World Geothermal Congress 2000*, 2257–2262.
- Valdez Salas, B., Schorr Wiener, M., Rioseco de la Peña, L., & Navarrete Bedolla, M. (2000). Deterioration of materials in geothermal fields in Mexico. *Materials and Corrosion*, *51*(10), 698–704. [https://doi.org/10.1002/1521-4176\(200010\)51:10<698::aid-maco698>3.0.co;2-g](https://doi.org/10.1002/1521-4176(200010)51:10<698::aid-maco698>3.0.co;2-g)
- Van de Vossenberg, J. L. C. M., Driessen, A. J. M., & Konings, W. N. (1998). The essence of being extremophilic: The role of the unique archaeal membrane lipids. *Extremophiles*. <https://doi.org/10.1007/s007920050056>
- Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(12), 4576–4579. <https://doi.org/10.1073/pnas.87.12.4576>
- Wooley, J. C., Godzik, A., & Friedberg, I. (2010). A primer on metagenomics. *PLoS Computational Biology*. <https://doi.org/10.1371/journal.pcbi.1000667>
- Wu, S., Zhu, Z., Fu, L., Niu, B., & Li, W. (2011). WebMGA: A customizable web server for fast metagenomic sequence analysis. *BMC Genomics*, *12*. <https://doi.org/10.1186/1471-2164-12-444>
- Wu, Y. W., Simmons, B. A., & Singer, S. W. (2016). MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*, *32*(4), 605–607. <https://doi.org/10.1093/bioinformatics/btv638>
- Xiang, X., Dong, X., & Huang, L. (2003). *Sulfolobus tengchongensis* sp. nov., a novel thermoacidophilic archaeon isolated from a hot spring in Tengchong, China. *Extremophiles*, *7*(6), 493–498. <https://doi.org/10.1007/s00792-003-0355-2>
- Xie, C., Goi, C. L. W., Huson, D. H., Little, P. F. R., & Williams, R. B. H. (2016). RiboTagger: Fast and unbiased 16S/18S profiling using whole community shotgun metagenomic or metatranscriptome surveys. *BMC Bioinformatics*, *17*. <https://doi.org/10.1186/s12859-016-1378-x>
- Yeats, S., McWilliam, P., & Zillig, W. (1982). A plasmid in the archaeobacterium *Sulfolobus acidocaldarius*. *The EMBO Journal*, *1*(9), 1035–1038. <https://doi.org/10.1002/j.1460-2075.1982.tb01292.x>
- Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Bäckström, Di., Juzokaite, L., Vancaester, E., ... Ettema, T. J. G. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature*, *541*(7637), 353–358. <https://doi.org/10.1038/nature21031>
- Zillig, W., Stetter, K. O., Wunderl, S., Schulz, W., Priess, H., & Scholz, I. (1980). The *Sulfolobus*-"*Caldariella*" group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. *Archives of Microbiology*, *125*(3), 259–269. <https://doi.org/10.1007/BF00446886>

Anexos

Scripts utilizados de las herramientas bioinformáticas

Sedimentos fotosintéticos

Generación de una base de datos para realizar mapeos mediante bowtie2.

bowtie2-build genoma de referencia (*Cuniculiplasma divulgatum* PM4) nombre de la base de datos

Mapeos mediante bowtie2.

bowtie2 --no-unal --very-sensitive -x Nombre de la base de datos -U extendedfrags.fastq --al /directorio donde se realiza el mapeo -S Nombre de salida.sam

Ensamble del genoma de *Ca. Cuniculiplasma* sp. AZ01 mediante SPAdes.

spades.py -k 21,33,55,77,99,127 -s Archivo.fastq -o Directorio de salida

Análisis de comparación del ensamble contra el genoma de referencia mediante MeaQuast.

metaquast.py contigs.fasta -R genoma de *Cuniculiplasma divulgatum* PM4

Sedimentos no fotosintéticos.

Ensamble del metagenoma mediante SPAdes.

spades.py -k 21,33,43,53,63 --only-assembler -1 Archivo1.fastq -2 Archivo2.fastq -o Directorio de salida

Análisis estadístico de ensamblajes con Quast.

quast.py directorio/contigs.fasta -o directorio de salida

Generación de una base de datos para realizar mapeos mediante bowtie2.

bowtie2-build genoma de referencia (*Ca. Aramenus sulfurataquae* AZ1) nombre de la base de datos

Mapeos mediante bowtie2.

bowtie2 --no-unal --very-sensitive -x Nombre de la base de datos -1 Archivo1.fastq -2 Archivo2.fastq
--al-conc /directorio donde se realiza el mapeo -S Nombre de salida.sam

Ensamble del genoma de *Candidatus Aramenus sulfurataquae* mediante SPAdes.

spades.py -k 21,33,43,53,63 --only-assembler -1 Archivo1.fastq -2 Archivo2.fastq -o Directorio de salida

Análisis de comparación del ensamble contra el genoma de referencia mediante MetaQuast.

metaquast.py contigs.fasta -R genoma de *Candidatus Aramenus sulfurataquae*

Muestras adicionales de sedimentos del campo geotérmico de Los Azufres.

Durante el mes de abril de 2019 se tomaron muestras de sedimentos fotosintéticos de los sitios uno y tres (Figura 17). Se extrajo ADN a partir de estas muestras de sedimentos usando el “*kit UltraClean Mega (Prep) Soil DNA kit*” (MoBio Laboratories, Inc., Carlsbad, CA). La secuenciación se llevó a cabo mediante la plataforma de secuenciación Illumina HiSeq2500 paired-end (2x250) en la empresa Macrogen INC. Corea del Sur.

Los resultados de secuenciación masiva de ADN de sedimentos fotosintéticos se muestran en la **¡Error! No se encuentra el origen de la referencia..** El metagenoma de la muestra B del sitio de Marítaro es el primer metagenoma secuenciado de esa zona del campo geotérmico de Los Azufres.

Tabla 8: Resultados de secuenciación masiva de ADN de muestras de sedimentos fotosintéticos del campo geotérmico de Los Azufres.

Muestra	Total de bases (pb)	Total de lecturas	GC(%)
B (Sitio1, Marítaro)	21,579,531,228	85,974,228	52.54
C (Sitio 3, Spa Natural)	22,618,252,560	90,112,560	58.93

Para la muestra B (Zona de Marítaro) aún no ha sido posible ensamblar las lecturas debido al tamaño de los archivos. Al ser archivos tan grandes requiere de un mayor poder computacional del que se tiene. Por lo que se buscaron otras metodologías para identificar las especies de microorganismos presentes en la muestra de sedimentos fotosintéticos B. Se usó RiboTagger v 0.8.1 (Xie, Goi, Huson, Little, & Williams, 2016) para identificar las regiones v4 de genes ribosomales 16S ARNr contenidas en las secuencias de la muestra B. Una vez identificadas las secuencias de las regiones V4 se analizaron mediante el servidor Silva ACT: Alineamiento, clasificación y servicio de árbol (<https://www.arb-silva.de/aligner/>) (Pruesse, Peplies, & Glöckner, 2012). Mayormente se identificaron regiones v4 de ribosomales 16S ARNr de bacterias se encontraron 2 regiones v4 de ribosomales 18S ARNr de Eucariontes (Tabla 9). No se encontraron regiones v4 de ribosomales 16S ARNr que pertenezcan a arqueas.

Tabla 9: Anotación taxonómica de las regiones V4 de genes ribosomales 16S ARNr de sedimentos fotosintéticos de la zona de Marítaro.

Dominio	Filo	Clase	Orden
Bacteria	<i>Acidobacteria</i>	<i>Acidobacteriia</i>	<i>Acidobacteriales</i>
Bacteria	<i>Actinobacteria</i>	<i>Acidimicrobiia</i>	<i>Acidimicrobiales</i>
Bacteria	<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Frankiales</i>
Bacteria	<i>Acidobacteria</i>	<i>Holophagae</i>	<i>Holophagales</i>
Bacteria	<i>Actinobacteria</i>	<i>Thermoleophilia</i>	<i>Gaiellales</i>
Bacteria	<i>Bacteroidetes</i>	<i>Bacteroidia</i>	<i>Bacteroidales</i>
Bacteria	<i>Cyanobacteria</i>	<i>Oxyphotobacteria</i>	<i>Chloroplast</i>
Bacteria	<i>ε-bacteraeota</i>	<i>Desulfurellia</i>	<i>Desulfurellales</i>
Bacteria	<i>ε-bacteraeota</i>	<i>Campylobacteria</i>	<i>Campylobacteriales</i>
Bacteria	<i>Firmicutes</i>	<i>Bacilli</i>	<i>Bacillales</i>
Bacteria	<i>Firmicutes</i>	<i>Clostridia</i>	<i>Thermoanaerobacteriales</i>
Bacteria	<i>Firmicutes</i>	<i>Clostridia</i>	<i>Clostridiales</i>
Bacteria	<i>Nitrospirae</i>	<i>Nitrospira</i>	<i>Nitrospirales</i>
Bacteria	<i>Patescibacteria</i>	<i>Saccharimonadia</i>	<i>Saccharimonadales</i>
Bacteria	<i>Planctomycetes</i>	<i>Phycisphaerae</i>	<i>Tepidisphaerales</i>
Bacteria	<i>Proteobacteria</i>	<i>α-proteobacteria</i>	<i>Acetobacteriales</i>
Bacteria	<i>Proteobacteria</i>	<i>α-proteobacteria</i>	<i>Micropepsales</i>
Bacteria	<i>Proteobacteria</i>	<i>α-proteobacteria</i>	<i>Rickettsiales</i>
Bacteria	<i>Proteobacteria</i>	<i>α-proteobacteria</i>	<i>Rhodospirillales</i>
Bacteria	<i>Proteobacteria</i>	<i>δ-proteobacteria</i>	<i>Oligoflexales</i>
Bacteria	<i>Proteobacteria</i>	<i>δ-proteobacteria</i>	<i>Desulfuromonadales</i>
Bacteria	<i>Proteobacteria</i>	<i>γ-proteobacteria</i>	<i>Betaproteobacteriales</i>
Bacteria	<i>Proteobacteria</i>	<i>γ-proteobacteria</i>	<i>Xanthomonadales</i>
Bacteria	<i>Proteobacteria</i>	<i>γ-proteobacteria</i>	<i>Enterobacteriales</i>
Bacteria	<i>Proteobacteria</i>	<i>γ-proteobacteria</i>	<i>Acidithiobacillales</i>
Bacteria	<i>Proteobacteria</i>	<i>γ-proteobacteria</i>	<i>Betaproteobacteriales</i>
Bacteria	<i>Proteobacteria</i>	<i>γ-proteobacteria</i>	<i>Cellvibrionales</i>
Bacteria	<i>Spirochaetes</i>	<i>Spirochaetia</i>	<i>Spirochaetales</i>
Bacteria	<i>Verrucomicrobia</i>	<i>Verrucomicrobiae</i>	<i>Chthoniobacteriales</i>
Eukaryota	<i>SAR</i>	<i>Stramenopiles</i>	<i>Ochrophyta</i>
Eukaryota	<i>Opisthokonta</i>	<i>Nucleomycea</i>	<i>Fungi</i>

Las lecturas del metagenoma de la muestra C (Los Azufres Spa Natural) se pudieron ensamblar. Es un ensamble preliminar ya que no fue posible realizar más ensambles para obtener un ensamble con valores de calidad aceptables. El ensamble del metagenoma se realizó mediante el programa SPAdes v 3.12.0 (Bankevich et al., 2012). El tamaño del metagenoma obtenido es de 203,293 pb contenido en 66,312 contigs. El contenido de GC y el valor de N50 es de 56.81% y 5,148, respectivamente.

Debido al tamaño del metagenoma obtenido no fue posible identificar genes ribosomales mediante algún servidor en línea. Por lo que se tomó la decisión de usar RiboTagger v 0.8.1 (Xie et al., 2016) para identificar las regiones V4 de genes ribosomales 16S ARNr contenidas en las secuencias de la muestra de sedimentos de Marítaro. Una vez identificadas las secuencias de las regiones V4 se analizaron mediante el servidor Silva ACT: Alineamiento, clasificación y servicio de árbol (<https://www.arb-silva.de/aligner/>)(Pruesse et al., 2012).

Mayormente se encontraron secuencias de regiones V4 de genes ribosomales 16S ARNr de arqueas y bacterias. Se encontraron dos regiones V4 de genes ribosomales 18S de eucariontes (Tabla 10). En este metagenoma resalta la presencia de una arquea ARMAN y una arquea del orden *Thermoplasmatales* del género *Cuniculiplasma*.

Tabla 10: Anotación taxonómica de las regiones V4 de genes ribosomales 16S ARNr de sedimentos fotosintéticos de la zona de Los Azufres Spa Natural.

Dominio	Filo	Clase	Orden	Familia	Género
Archaea	<i>Crenarchaeota</i>	<i>Thermoprotei</i>	<i>Sulfolobales</i>	<i>Sulfolobaceae</i>	<i>Acidianus</i>
Archaea	<i>Crenarchaeota</i>	<i>Thermoprotei</i>	<i>Sulfolobales</i>	<i>Sulfolobaceae</i>	<i>Sulfolobus</i>
Archaea	<i>Diapherotrites</i>	<i>Micrarchaeia</i>	\	\	\
Archaea	<i>Euryarchaeota</i>	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermoplasmataceae</i>	A-plasma
Archaea	<i>Euryarchaeota</i>	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermoplasmataceae</i>	E-plasma
Archaea	<i>Euryarchaeota</i>	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Ferroplasmaceae</i>	<i>Acidiplasma</i>
Archaea	<i>Euryarchaeota</i>	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Ferroplasmaceae</i>	<i>Ferroplasma</i>
Archaea	<i>Euryarchaeota</i>	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermoplasmataceae</i>	<i>Cuniculiplasma</i>
Bacteria	<i>Actinobacteria</i>	<i>Acidimicrobiia</i>	<i>Acidimicrobiales</i>	<i>Acidimicrobiaceae</i>	<i>Ferrimicrobium</i>
Bacteria	<i>Actinobacteria</i>	<i>Acidimicrobiia</i>	<i>Acidimicrobiales</i>	<i>Acidimicrobiales</i>	<i>Aciditerrimonas</i>
Bacteria	<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Frankiales</i>	<i>Acidothermaceae</i>	<i>Acidothermus</i>
Bacteria	<i>Actinobacteria</i>	<i>Acidimicrobiia</i>	<i>Acidimicrobiales</i>	<i>Acidimicrobiaceae</i>	<i>Acidimicrobium</i>
Bacteria	<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Corynebacteriales</i>	<i>Mycobacteriaceae</i>	<i>Mycobacterium</i>
Bacteria	<i>Actinobacteria</i>	<i>Thermoleophilia</i>	<i>Solirubrobacterales</i>	<i>Solirubrobacteraceae</i>	<i>Conexibacter</i>
Bacteria	<i>Cyanobacteria</i>	<i>Oxyphotobacteria</i>	<i>Chloroplast</i>		
Bacteria	<i>Firmicutes</i>	<i>Bacilli</i>	<i>Bacillales</i>	<i>Alicyclobacillaceae</i>	<i>Alicyclobacillus</i>
Bacteria	<i>Firmicutes</i>	<i>Bacilli</i>	<i>Bacillales</i>	<i>Alicyclobacillaceae</i>	<i>Acidibacillus</i>
Bacteria	<i>Firmicutes</i>	<i>Clostridia</i>	<i>Clostridiales</i>	<i>Family XVII</i>	<i>Sulfobacillus</i>
Bacteria	<i>Nitrospirae</i>	<i>Nitrospira</i>	<i>Nitrospirales</i>	<i>Nitrospiraceae</i>	<i>Leptospirillum</i>
Bacteria	<i>Proteobacteria</i>	<i>γ-proteobacteria</i>	<i>Acidithiobacillales</i>	<i>Acidithiobacillaceae</i>	<i>Acidithiobacillus</i>
Bacteria	<i>Proteobacteria</i>	<i>α-proteobacteria</i>	<i>Acetobacterales</i>	<i>Acetobacteraceae</i>	<i>Acidicaldus</i>
Eukaryota	<i>Archaeplastida</i>	<i>Rhodophyceae</i>	<i>Cyanidiales</i>		
Eukaryota	<i>Opisthokonta</i>	<i>Nucleomyces</i>	<i>Fungi</i>	<i>Dikarya</i>	<i>Ascomycota</i>

Por último, se lograron recuperar tres genomas parciales a partir del metagenoma de la muestra C mediante mapeos de lecturas contra genomas de referencia usando el programa bowtie2 v 2.3.4.3 (Langmead & Salzberg, 2012). En la Tabla 11 se muestran los estadísticos de los tres ensamblajes obtenidos. Estos ensamblajes son preliminares y aún requieren de un mayor trabajo bioinformático.

Tabla 11: Datos estadísticos de los tres ensamblajes obtenidos de arqueas a partir de la muestra de sedimentos fotosintéticos de la zona de Los Azufres, Spa Natural.

<i>Genoma</i>	Tamaño del genoma	Numero de contigs	Contenido de GC	Valor de N50	Genoma de referencia
Cuniculiplasma sp.	2,425,619	1,975	37.6%	3,183	<i>Cuniculiplasma divulgatum</i> PM4
Acidiplasma sp.	1,627,983	368	35.3%	8,504	<i>Acidiplasma cupricumulans</i> BH2
Arquea ARMAN	1,285,266	741	45.9%	3,708	<i>Candidatus Micrarchaeum</i> sp. AZ1