



Metagenome Assembly and Metagenome-Assembled Genome of “*Candidatus Aramenus sulfurataquae*” from Thermal Sediments from the Los Azufres Volcanic Complex

Roberto Marín-Paredes,^a Yunuen Tapia-Torres,^{a,b}  Esperanza Martínez-Romero,^c Mauricio Quesada,^{b,d}  Luis E. Servín-Garcidueñas^{a,b}

^aLaboratorio de Microbiómica, Escuela Nacional de Estudios Superiores Unidad Morelia, Universidad Nacional Autónoma de México, Morelia, Michoacán, México

^bLaboratorio Nacional de Análisis y Síntesis Ecológica, Escuela Nacional de Estudios Superiores Unidad Morelia, Morelia, Michoacán, México

^cCentro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

^dInstituto de Investigaciones en Ecosistemas y Sustentabilidad, Universidad Nacional Autónoma de México, Morelia, Michoacán, México

ABSTRACT A plethora of hot springs are found at the Los Azufres volcanic complex in Mexico, and studies are needed to determine their microbial genomic diversity. Here, we report a metagenome of hot spring sediments and a metagenome-assembled genome of “*Candidatus Aramenus sulfurataquae*.” This study reveals novel genomic sequences of *Sulfolobales* archaea.

Archaea from the order *Sulfolobales* are commonly found in acidic high-temperature environments (1, 2). A large diversity of environmental *Sulfolobales* strains remains to be described. We are interested in characterizing the genomic diversity of *Sulfolobales* strains inhabiting extreme environments found in the Los Azufres volcanic area in western Mexico. Previously, we obtained a genome assembly for the “*Candidatus Aramenus*” genus from a hot spring metagenome from water samples (65°C, pH 3.6) (3). The phylogenetic relationship of the *Aramenus* genus within *Sulfolobales* has been investigated recently (4). Interestingly, current diversity studies have detected *Aramenus* archaea at the El Chichón volcano in Chiapas, Mexico (5). Additional genomes are required to better understand the phylogenetic position, genomic diversity, and evolutionary history of the *Aramenus* genus. We describe a new metagenome-assembled genome (MAG) for “*Candidatus Aramenus sulfurataquae*” strain AZ2 from thermal sediments within Los Azufres.

A sediment sample was collected from a hot spring (88.5°C, pH 2.8) at Los Azufres in May 2016. DNA was isolated using the DNeasy UltraClean microbial kit (Qiagen GmbH, Germany). Library preparation was performed using a DNA preparation kit (Illumina Inc., San Diego, CA, USA) according to the manufacturer's instructions. Sequencing was performed at the Laboratorio Nacional de Análisis y Síntesis Ecológica (Morelia, Mexico) on an Illumina NextSeq550 platform, producing 2 × 75-bp paired-end reads. Reads were assessed using FastQC v.0.11.8 (with default settings) (6) and were filtered for quality (scores of ≥Q30) and adaptor sequences using Trim Galore v.0.6.6 (7). Reads were assembled *de novo* with SPAdes v.3.8.2 (8, 9). Metagenomic annotation was performed using the Integrated Microbial Genome and Microbiomes (IMG/M) pipeline (10). The MAG of “*Candidatus Aramenus sulfurataquae*” was recovered by mapping the reads against the genome of “*Candidatus Aramenus sulfurataquae*” strain AZ1 (GenBank accession number [ASRH000000001](https://www.ncbi.nlm.nih.gov/nuccore/ASRH000000001)) using bowtie2 v.2.3.4.3 (11) with the -very-sensitive mode. Mapped reads were assembled with SPAdes v.3.12.0 (8, 9) using *k*-mer values of 21, 33, 43, 53, and 63. MAG quality was revised with CheckM v.1.0.13 (with default settings) (12–15). MAG annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v.5.2 (16). Average nucleotide identity (ANI) values were calculated as proposed previously (17),

Citation Marín-Paredes R, Tapia-Torres Y, Martínez-Romero E, Quesada M, Servín-Garcidueñas LE. 2021. Metagenome assembly and metagenome-assembled genome of “*Candidatus Aramenus sulfurataquae*” from thermal sediments from the Los Azufres volcanic complex. *Microbiol Resour Announc* 10:e00379-21. <https://doi.org/10.1128/MRA.00379-21>.

Editor Irene L. G. Newton, Indiana University, Bloomington

Copyright © 2021 Marín-Paredes et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Roberto Marín-Paredes, roberto.marin.ibq@gmail.com.

Received 15 April 2021

Accepted 28 August 2021

Published 30 September 2021

using the ANI calculator from the Konstantinidis laboratory (<http://enve-omics.ce.gatech.edu/ani>) (18).

Metagenomic reads (60,553,942 reads) were assembled into 1,718 contigs (N_{50} , 7,585 bp) with a minimum length of 1,000 bp, containing 11,279,257 bp. The metagenome revealed *Sulfolobales* sequences related to the *Acidianus*, *Sulfolobus*, and *Aramenus* genera. A 1,681,874-bp draft MAG comprising 379 contigs (N_{50} , 11,611 bp), with a GC content of 47.8% and genome coverage of 84.31 \times , could be obtained from the metagenome. The MAG was estimated to be 99.29% complete, with no detected contamination. MAG annotation predicted 2,103 genes and revealed a 1,498-bp 16S rRNA gene that shares 100% sequence identity with the gene from “*Candidatus Aramenus sulfurataquae*” strain AZ1. Genes involved in carbon fixation and sulfur metabolism could be detected. The MAG exhibited an ANI value of 97.91% against the “*Candidatus Aramenus sulfurataquae*” strain AZ1 genome, suggesting that they correspond to the same species. We identified a “*Candidatus Aramenus sulfurataquae*” population in a new niche within Los Azufres with higher temperature and lower pH.

Data availability. Raw sequencing data are available in the NCBI Sequence Read Archive (SRA) under accession number [SRR7716004](https://www.ncbi.nlm.nih.gov/sra/SRR7716004). The metagenome assembly is available in GenBank under accession number [JAFNEM000000000](https://www.ncbi.nlm.nih.gov/genbank/JAFNEM000000000). The metagenome functional annotation is available from the JGI Genome Portal under accession number [3300029921](https://www.jgi.doe.gov/data/record/3300029921). The annotated MAG for “*Candidatus Aramenus sulfurataquae*” strain AZ2 is available in GenBank under accession number [JAFLRL000000000](https://www.ncbi.nlm.nih.gov/genbank/JAFLRL000000000).

ACKNOWLEDGMENTS

This work was supported by grants to L.E.S.-G. from PAPIIT-UNAM (grants IA210617 and IA208019) and by grants to M.Q. from the Laboratorio Nacional de Análisis y Síntesis Ecológica, CONACYT-UNAM-UAGro-UMSNH (grants 2015-LN250996, 2016-LN271449, 2017-LN280505, and 2018-LN293701). R.M.-P. is a Ph.D. student from the Programa de Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México (UNAM), supported by a fellowship from CONACYT-México. This work is a prerequisite for R.M.-P. to obtain a Ph.D. degree in biological sciences.

We thank Luisa Falcón-Álvarez for critically reading the paper. We thank Gumersindo Sánchez-Montoya, Ulises Olivares-Pinto, and Violeta Patiño-Conde for their support.

REFERENCES

- Huber H, Prangishvili D. 2006. *Sulfolobales*, p 23–51. In Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (ed), *The prokaryotes*, 3rd ed. Springer-Verlag, New York, NY.
- Stetter KO. 1989. Order III: *Sulfolobales* ord. nov. family *Sulfolobaceae* fam. nov., p 2250–2251. In Staley JT, Bryant MP, Pfennig N, Holt JG (ed), *Bergey's manual of systematic bacteriology*, vol 3. Williams & Wilkins Co, Baltimore, MD.
- Servín-Garcidueñas LE, Martínez-Romero E. 2014. Draft genome sequence of the *Sulfolobales* archaeon AZ1, obtained through metagenomic analysis of a Mexican hot spring. *Genome Announc* 2:e00164-14. <https://doi.org/10.1128/genomeA.00164-14>.
- Counts JA, Willard DJ, Kelly RM. 2021. Life in hot acid: a genome-based reassessment of the archaeal order *Sulfolobales*. *Environ Microbiol* 23: 3568–3584. <https://doi.org/10.1111/1462-2920.15189>.
- Peña-Ocaña BA, Ovando-Ovando CI, Puente-Sánchez F, Tamames J, Servín-Garcidueñas LE, González-Toril E, Gutiérrez-Sarmiento W, Jasso-Chávez R, Ruiz-Valdiviezo VM. 2021. Metagenomic and metabolic analyses of poly-extreme microbiome from an active crater volcano lake. *Environ Res* 203:111862. <https://doi.org/10.1016/j.envres.2021.111862>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Krueger F. 2015. Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore.
- Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol* 20:714–737. <https://doi.org/10.1089/cmb.2013.0084>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Chen IA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, Huntemann M, Varghese N, White JR, Seshadri R, Smirnova T, Kirton E, Jungbluth SP, Woyke T, Eloe-Fadrosh EA, Ivanova NN, Kyrpides NC. 2019. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res* 47:D666–D677. <https://doi.org/10.1093/nar/gky901>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- Matsen FA, Kodner RB, Armbrust EV. 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 11:538. <https://doi.org/10.1186/1471-2105-11-538>.
- Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
- Eddy SR. 1995. Multiple alignment using hidden Markov models. *Proc Int Conf Intell Syst Mol Biol* 3:114–120.

16. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
17. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57:81–91. <https://doi.org/10.1099/ijs.0.64483-0>.
18. Rodriguez-R LM, Konstantinidis KT. 2016. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ* 4:e1900v1. <https://doi.org/10.7287/peerj.preprints.1900v1>.