

Draft genome sequence of a member of a putatively novel *Rubrobacteraceae* genus from lava tubes in Lava Beds National Monument

Catherine Maggiori,¹ Zachariah John,¹ Dina M. Bower,^{2,3} Maëva Millan,⁴ Aria S. Hahn,⁵ Amy McAdam,⁶ Sarah Stewart Johnson¹

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We report the draft genome sequence of a member of a potentially novel genus of *Rubrobacteraceae* isolated from Golden Dome Cave in Lava Beds National Monument. Members of this family are known to inhabit thermophilic environments. The metagenome-assembled genome presented here helps illuminate the genetic capacity of basaltic lava tube environments.

KEYWORDS microbiology, microbial ecology, astrobiology, metagenomics, environmental microbiology

Lava Beds National Monument contains numerous volcanic formations, such as basaltic lava tubes analogous to those on Mars. Golden Dome is one such cave, composed of lava flow from the Medicine Lake volcano ~36,000 years ago (1). Rock samples from the walls of Golden Dome cave were collected aseptically on 10 May 2022, frozen, and shipped to Georgetown University. The samples were stored at –80°C until analysis.

Samples underwent a pre-extraction chemical lysis: ~250 mg of sample was mixed with 400 µL 2× PBS, 25 µL metapolyzyme (cat. no. MAC4LDF-5 X1VL), and 250 µL ZymoBIOMICS DNA Miniprep Kit Lysis Solution and vortexed for 10 seconds. The mixture was incubated at 35°C for 2 hours, with intervals of 10 second vortexing every 30 minutes. 15 µL of proteinase K was then added, followed by another incubation at 55°C for 1 hour. DNA was extracted from the lysed samples using steps 4–11 of the ZymoBIOMICS DNA Miniprep Kit (cat. no. D6035).

A paired-end library was constructed using 3.12 ng of DNA and Illumina's Nextera XT DNA Library Preparation Kit with Illumina DNA/RNA UD Indexes and sequenced on a NovaSeq platform to produce 5,017,368 150 bp reads. Reads were trimmed to remove adapters and low-quality bases using Trimmomatic version 0.39 (2). The remaining 5,015,316 reads were assembled using SPAdes genome assembler version 4.0.0 (metaSPAdes mode) (3) and filtered for a minimum length of 500 bp using bbtools version 39.08. Bins were generated using MetaBat2 version 2.17 (4), quality checked with CheckM version 1.2.3 (5), and analyzed with anvi'o version 8 (6). KEGG Kofam and COG databases were used for annotation, and GTDb-tk version 2.4.0 was used to assign taxonomy (7). All tools were used with default parameters, unless otherwise specified.

One 2,434,521 bp bin showed a 7.10% alignment rate with the metagenome, comprising 147 contigs with an N_{50} of 24,326 bp and 2,578 open reading frames (ORFs) as identified by Prodigal, a GC content of 64.985%, and completeness and contamination values of 89.04% and 1.84%, respectively (Fig. 1). A total of 1 16S rRNA and 39 tRNA genes were identified, and GTDb-tk classified this metagenome-assembled genome (MAG) as a member of *Rubrobacteraceae*. Its 16S rRNA gene was unclassified when aligned against the SILVA database (8), and NCBI's 16S rRNA database showed a closest

Editor Simon Roux, DOE Joint Genome Institute, Berkeley, California, USA

Address correspondence to Catherine Maggiori, catherine.maggiori@gmail.com.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 20 December 2024

Accepted 27 March 2025

Published 21 April 2025

Copyright © 2025 Maggiori 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/).

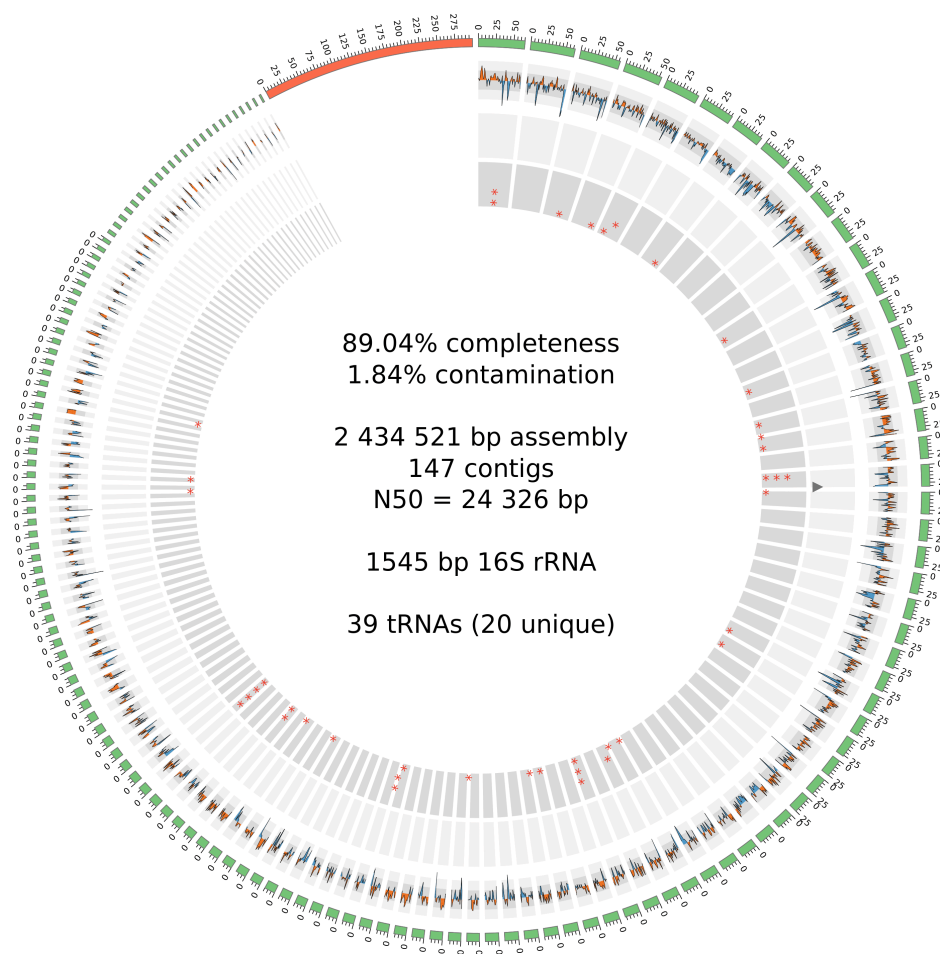


FIG 1 From the outer to inner ring of this Circos metagenome-assembled genome plot: green bars represent the contigs of this MAG, while the red bar is the approximate amount of DNA missing, with ticks indicating contig length in kbp; GC content of 1,000 bp windows from the mean GC of the entire MAG; the location of the 16S rRNA gene (triangle); and the location of tRNA genes (asterisk). This plot was generated with Koonkie Inc. *circos_mag* software (https://github.com/Koonkie-Cloud-Services/circos_mag) and Circos (9).

match to *Rubrobacter radiotolerans* strain P 1 (accession no. [NR_029191.2](#)) at 94.12% identity; thus, this MAG is likely a member of a novel genus of *Rubrobacteraceae*.

Based on gene annotations, this MAG is probably a facultatively anaerobic heterotroph, with complete glycolysis, pyruvate oxidation, and glyoxylate cycle pathways. Some chlorophyll production genes are present, plus an incomplete reductive pentose phosphate cycle module, similar to other *Rubrobacteraceae* (10).

This MAG also possesses some methylotrophic features, including the *mer* gene (5,10-methylenetetrahydromethanopterin reductase), used for methoxylated aromatic compound degradation in methanogenesis and methoxydotrophy in archaea (11), although other bacteria are known to use methoxylated aromatics for energy conservation (12). This MAG is also capable of synthesizing Factor 420, potentially to produce secondary metabolites/antibiotics (13).

We also identified various stress response genes in this MAG, including ORFs for oxidative stress, UV stress, osmotic stress, desiccation, and sporulation.

AUTHOR AFFILIATIONS

¹Department of Biology, Georgetown University, Washington, DC, USA

²Department of Astronomy, University of Maryland, College Park, Maryland, USA

³Planetary Systems Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

⁴LATMOS Laboratoire Atmosphères, Chargée de recherche (CNRS), Guyancourt, Île-de-France, France

⁵Koonkie Inc., Vancouver, British Columbia, Canada

⁶Planetary Environments Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

AUTHOR ORCIDS

Catherine Maggiori  <http://orcid.org/0000-0002-6138-1789>

FUNDING

| Funder | Grant(s) | Author(s) |
|---|----------|-----------------------|
| National Aeronautics and Space Administration | | Zachariah John |
| | | Dina M. Bower |
| | | Maëva Millan |
| | | Amy McAdam |
| | | Sarah Stewart Johnson |

AUTHOR CONTRIBUTIONS

Catherine Maggiori, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review and editing | Zachariah John, Formal analysis, Investigation, Methodology, Writing – review and editing | Dina M. Bower, Conceptualization, Funding acquisition, Project administration, Resources, Writing – review and editing | Maëva Millan, Conceptualization, Funding acquisition, Project administration, Resources, Writing – review and editing | Aria S. Hahn, Visualization, Writing – review and editing | Amy McAdam, Conceptualization, Funding acquisition | Sarah Stewart Johnson, Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

The data for this work are available at NCBI BioProject [PRJNA1189542](#). The whole shotgun metagenome reads are deposited under accession no. [SRR31455428](#), assembled reads under accession no. [JBLRVU000000000.1](#), and the MAG is accessible under accession no. [SRR31488317](#).

REFERENCES

- Lavoie KH, Winter AS, Read KJH, Hughes EM, Spilde MN, Northup DE. 2017. Comparison of bacterial communities from lava cave microbial mats to overlying surface soils from Lava Beds National Monument, USA. *PLoS One* 12:e0169339. <https://doi.org/10.1371/journal.pone.0169339>
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>
- Kang DD, Li F, Kirtan E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. <https://doi.org/10.7717/peerj.7359>
- 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>
- Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319. <https://doi.org/10.7717/peerj.1319>
- Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–6. <https://doi.org/10.1093/nar/gks1219>
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res* 19:1639–1645. <https://doi.org/10.1101/gr.092759.109>
- Freed JS, Ramaley RF, Kyndt JA. 2019. Whole-genome sequence of the novel *Rubrobacter taiwanensis* strain yellowstone, isolated from Yellowstone National Park. *Microbiol Resour Announc* 8:19. <https://doi.org/10.1128/MRA.00287-19>

11. Welte CU, de Graaf R, Dalcin Martins P, Jansen RS, Jetten MSM, Kurth JM. 2021. A novel methoxydotrophic metabolism discovered in the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *Environ Microbiol* 23:4017–4033. <https://doi.org/10.1111/1462-2920.15546>
12. Bache R, Pfennig N. 1981. Selective isolation of *Acetobacterium woodii* on methoxylated aromatic acids and determination of growth yields. *Arch Microbiol* 130:255–261. <https://doi.org/10.1007/BF00459530>
13. Hasan M, Schulze S, Berndt L, Palm GJ, Braga D, Richter I, Last D, Lammers M, Lackner G. 2022. Diversification by CofC and control by CofD govern biosynthesis and evolution of coenzyme F420 and its derivative 3PG-F420. *MBio*:e03501. <https://doi.org/10.1128/mbio.03501-21>