

Draft genome sequences of *Arthrobacter* sp. AZCC_0090 and *Mycobacterium* sp. AZCC_0083 isolated from oligotrophic subsurface forest soil in the Santa Catalina mountains of Southern Arizona

Melanie R. Kridler,¹ Isabella A. Viney,¹ Joy M. Custer,¹ Bradley Schlottman,¹ Ryan Bartelme,¹ Paul Carini^{1,2}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Here, we present the genomes of two soil actinobacteria: *Arthrobacter* sp. strain AZCC_0090 and *Mycobacterium* sp. strain AZCC_0083, isolated from oligotrophic subsurface soils in Southern Arizona, USA.

KEYWORDS environmental microbiology, soil microbiology, genomics, microbial ecology, genome sequences

Deep soil horizons contain microbes crucial for subsurface biogeochemistry and the global carbon cycle. While microbial communities have been extensively studied in surface horizons, we know far less about the microbial communities residing in deeper soil horizons despite their important roles in the mineralization of nutrients necessary for plant life (1–3).

Arizona Culture Collection (AZCC) cultures *Arthrobacter* sp. AZCC_0090 and *Mycobacterium* sp. AZCC_0083 were isolated from 50 cm subsurface soil at Oracle Ridge in Tucson, AZ, USA as described elsewhere (3) (Table 1). The cultures were propagated from isolated colonies on yeast mannitol medium (YM) consisting of (L⁻¹) 1.0 g yeast extract, 10.0 g mannitol, 0.5 g K₂HPO₄, 0.2 g MgSO₄, and 0.1 g NaCl. Solid YM medium was prepared by including 15 g L⁻¹ Noble agar. The cultures were incubated at 20°C. Broth cultures were shaken at 180 RPM. Growth appeared within 4 days.

Genomic DNA (gDNA) was isolated from a broth culture originating from a single colony using the Joint Genome Institute CTAB protocol for bacterial DNA isolation (6). For AZCC_0083, 100 ng of gDNA was sheared to 511 bp using the Covaris LE220 and size selected with SPRI using TotalPure NGS beads (Omega Bio-tek) to enrich for 200–500 bp fragments. For AZCC_0090, 200 ng of gDNA was sheared to 461 bp. After size selection with double-SPRI to enrich for 200–500 bp fragments, DNA fragments were end-repaired, A-tailed, and ligated with Illumina-compatible sequencing adapters containing a unique molecular index barcode for each sample library using the KAPA-HyperPrep kit (KAPA Biosystems).

The libraries were quantified using KAPA Biosystems' next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. Genomes were generated from Illumina 2 × 151 bp libraries using the Illumina NovaSeq XP V1 reagent kits and the S4 flowcell. Illumina sequences were quality filtered using BBTools: BBDuk (version 38.75) and BBDuk and assembled with SPAdes (version v3.13.0; -phred-offset 33 -cov-cutoff auto -t 16 -m 64 -careful -k 25,55,95) (7, 8). Contigs < 1 kb were discarded (BBTools reformat.sh: minlength = 1,000 ow = t). Annotations were made using the prokaryotic genome annotation pipeline (9) and are available in IMG/MER (10). Both genomes are nearly complete, with minimal contamination via CheckM2 (5) (Table

Editor Elinne Becket, California State University San Marcos, San Marcos, California, USA

Address correspondence to Paul Carini, paulcarini@arizona.edu.

The authors declare no conflict of interest.

See the funding table on p. 3.

See the companion article at <https://doi.org/10.1128/msphere.00024-20>.

Received 7 December 2023

Accepted 6 February 2024

Published 20 February 2024

Copyright © 2024 Kridler 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/).

TABLE 1 Isolation information, genome accessions, and summary statistics

Strain	<i>Arthrobacter</i> sp. AZCC_0090	<i>Mycobacterium</i> sp. AZCC_0083
Isolation site and soil depth	Oracle Ridge 50 cm	Oracle Ridge 50 cm
Isolation site coordinates	32°27'1" N 110°44'29" W	32°27'1" N 110°44'29" W
GenBank accession	JACHEE000000000	JACHGP000000000
Genome accession	PRJNA632064	PRJNA632065
SRA ID	SRX8890241	SRX8890242
Biosample accession	SAMN14915971	SAMN14915993
Number of raw reads generated	16,660,990	18,463,710
Number of raw bp generated	2,515,809,490	2,788,020,210
Number of reads after filtering	16,539,064	18,410,218
Number of bp after filtering	2,479,728,727	2,759,045,573
<i>N50</i>	315,689	428,479
Scaffolds	34	40
Genome size (bp)	4,921,994	8,615,464
Percent GC	63.32	66.11
Average coverage (fold)	305	174
Contigs	41	47
Extrachromosomal elements	0	0
Protein-coding genes	4,653	8,566
5S rRNA	1	1
16S rRNA	1	3
23S rRNA	1	1
tRNAs	52	82
Putative CRISPR count	2	0
Putative biosynthetic gene clusters	0	40
Predicted viruses ^a	0	1
Completeness ^b	99.97	100
Contamination ^b	0.37	1.41

^aPredictions from geNomad (4).

^bPredictions from CheckM2 (5).

1). Strain AZCC_0083 codes for 40 predicted biosynthetic gene clusters, and AZCC_0090 codes for two putative CRISPR genes (Table 1).

ACKNOWLEDGMENTS

Funding from this work came from startup funds provided to P.C. from the University of Arizona's Technology and Research Initiative Fund (the Water, Environmental, and Energy Solutions initiative), seed grants from the University of Arizona Center for Environmentally Sustainable Mining, and College of Agriculture and Life Sciences. With outside funding from the National Science Foundation (awards IOS-2141605 to P.C. and DGE-2137419 to I.A.V.), and a Simons Foundation Early Career Investigator in Marine Microbial Ecology and Evolution Award to P.C. The work (proposal: 10.46936/10.25585/60001177) conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231.

M.K., I.A.V., and P.C. wrote the announcement and analyzed genomes, and B.S. and R.B. performed the experiments.

AUTHOR AFFILIATIONS

¹Department of Environmental Science, University of Arizona, Tucson, Arizona, USA

²BIO5 Institute, University of Arizona, Tucson, Arizona, USA

AUTHOR ORCID*s*

Melanie R. Kridler  <http://orcid.org/0009-0007-8965-018X>
 Ryan Bartelme  <http://orcid.org/0000-0002-6178-2603>
 Paul Carini  <http://orcid.org/0000-0002-9653-7309>

FUNDING

Funder	Grant(s)	Author(s)
National Science Foundation (NSF)	IOS-2141605	Paul Carini
National Science Foundation (NSF)	DGE-2137419	Isabella A. Viney
Technology and Research Initiative Fund-WEES		Paul Carini

AUTHOR CONTRIBUTIONS

Melanie R. Kridler, Formal analysis, Writing – original draft, Writing – review and editing | Isabella A. Viney, Formal analysis, Writing – original draft, Writing – review and editing | Bradley Schlottman, Investigation, Methodology | Ryan Bartelme, Investigation, Methodology | Paul Carini, Formal analysis, Writing – original draft, Writing – review and editing.

DATA AVAILABILITY

The draft genome sequences of *Arthrobacter* sp. AZCC_0090 and *Mycobacterium* sp. AZCC_0083 have been deposited at NCBI. Their accessions are listed in Table 1.

REFERENCES

- Brewer TE, Aronson EL, Arogyaswamy K, Billings SA, Botthoff JK, Campbell AN, Dove NC, Fairbanks D, Gallery RE, Hart SC, Kaye J, King G, Logan G, Lohse KA, Maltz MR, Mayorga E, O'Neill C, Owens SM, Packman A, Pett-Ridge J, Plante AF, Richter DD, Silver WL, Yang WH, Fierer N. 2019. Ecological and genomic attributes of novel bacterial taxa that thrive in subsurface soil horizons. *mBio* 10. <https://doi.org/10.1128/mBio.01318-19>
- Gabor RS, Eilers K, McKnight DM, Fierer N, Anderson SP. 2014. From the litter layer to the saprolite: chemical changes in water-soluble soil organic matter and their correlation to microbial community composition. *Soil Biology and Biochemistry* 68:166–176. <https://doi.org/10.1016/j.soilbio.2013.09.029>
- Bartelme RP, Custer JM, Dupont CL, Espinoza JL, Torralba M, Khalili B, Carini P. 2020. Influence of substrate concentration on the culturability of heterotrophic soil microbes isolated by high-throughput dilution-to-extinction cultivation. *mSphere* 5:e00024-20. <https://doi.org/10.1128/mSphere.00024-20>
- Camargo AP, Nayfach S, Chen I-MA, Palaniappan K, Ratner A, Chu K, Ritter SJ, Reddy TBK, Mukherjee S, Schulz F, Call L, Neches RY, Woyke T, Ivanova NN, Elie-Fadrosh EA, Kyrpides NC, Roux S. 2023. IMG/VR v4: an expanded database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata. *Nucleic Acids Res* 51:D733–D743. <https://doi.org/10.1093/nar/gkac1037>
- Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. 2023. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods* 20:1203–1212. <https://doi.org/10.1038/s41592-023-01940-w>
- William S, Helene Feil AC. 2012. Bacterial Genomic DNA isolation using CTAB. Available from: <https://jgi.doe.gov/wp-content/uploads/2014/02/JGI-Bacterial-DNA-isolation-CTAB-Protocol-2012.pdf>
- BBTools. DOE joint genome Institute. 2016. Available from: <https://jgi.doe.gov/data-and-tools/software-tools/bbtools>
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>
- 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>
- Markowitz VM, Mavromatis K, Ivanova NN, Chen I-MA, Chu K, Kyrpides NC. 2009. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 25:2271–2278. <https://doi.org/10.1093/bioinformatics/btp393>