

Draft genome sequence of an uncultured archaeon from Antarctic endolithic communities

Claudia Coleine,¹ Diego Micheletti,² Massimo Pindo,² Simone Larger,² Erika Stefani,² Federico Biagioli,¹ Jason E. Stajich,^{3,4} Claudio Donati²

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT A draft genome sequence was assembled and annotated for an uncultured archaeon reconstructed from shotgun metagenomes obtained from Antarctic endoliths. The assembled genome is 1.99 megabases and encodes 2,405 predicted protein-coding genes. This genome sequence provides insights into the microbial diversity and functional potential of extremophiles inhabiting Antarctic rock environments.

KEYWORDS Antarctica, extremophiles, archaea, microbes, metagenomics

Endolithic microbial communities in Antarctica thrive in one of the most extreme environments on Earth (1, 2). These communities inhabit rock airspaces, exploiting microenvironments that allow them to survive at the limits of habitability (3). They host highly adapted microbes that sustain metabolism through trace gas oxidation and atmospheric chemosynthesis under extreme oligotrophic conditions (4–7). Here, we report the draft genome sequence of an uncultured *Nitrosocosmicus* archaeon reconstructed from shotgun metagenomes from Antarctic endolithic communities.

Nitrosocosmicus belongs to the phylum Thaumarchaeota, a group of ammonia-oxidizing archaea (8, 9), which is known for the ability to oxidize ammonia under extreme conditions (10, 11). The discovery of *Nitrosocosmicus* in Antarctic endoliths highlights ammonia oxidation as a key survival strategy for Thaumarchaeota in nutrient-poor cold deserts (12). Sequencing this *Nitrosocosmicus* archaeon might reveal insights into microbial diversity, survival, and biogeochemical roles in Antarctic ecosystems (13, 14).

A quartz was collected on Kay Island (Northern Victoria Land, Antarctica; 74.07°S, 165.31°E) in sterile bags, stored at –20°, and shipped to the University of Tuscia. Environmental DNA was extracted from 1 g of crushed rock using the DNeasy PowerSoil Pro Kit (Qiagen, Germany) (15). DNA quality was assessed by the 4150 TapeStation system and quantified using the Qubit dsDNA HS Assay Kit (Life Technologies, USA) (16). Illumina libraries were constructed using the Kapa Hyperplus Kit (Roche). Full shotgun metagenomic sequencing was performed using the Illumina NovaSeq6000 platform, generating paired-end 2 × 150 bp reads. The total number of reads was 82,718,564 reads. Raw reads were quality trimmed and filtered using BBDuk (BBMap v38.79). During the quality filtering procedure: (i) raw reads were quality trimmed to Q6 using the Phred algorithm; (ii) reads that contained 4 or more “N” bases, had an average quality below 10, and shorter than 50 bp or under 50% of the original length were removed. Assembly was performed using SPAdes v3.15.1 (metaSPAdes mode, k-mers: 21, 33, 55, 77, 99) (17). Contigs were binned into metagenome-assembled genomes (MAGs) using MetaBAT2 v2.12.1, with coverage depth estimated via Bowtie2 and Samtools (18, 19). Genomes completeness and contamination were assessed using CheckM v1.1.2 (20). MAGs were classified as high quality (>90% completeness, <5% contamination) or medium quality (≥50% completeness, <10% contamination) (6). A high-quality archaeal MAG was taxonomically classified using the metashot/prok-classify v1.2.1 workflow

Editor Kenneth M. Stedman, Portland State University, Portland, Oregon, USA

Address correspondence to Claudia Coleine, coleine@unitus.it, or Claudio Donati, claudio.donati@fmach.it.

The authors declare no conflict of interest.

Received 25 April 2025

Accepted 25 June 2025

Published 23 July 2025

Copyright © 2025 Coleine 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/).

(<https://github.com/metashot/prok-classify>, parameters: `-gtdbtk_db release202`), which integrates the Genome Taxonomy Database Toolkit (GTDB-Tk) v1.5.0 (21).

The assembled genome of *Nitrosocosmicus* sp. Mars-44.bin.73 is 1.99 Mb in size, with 72.49% completeness and 3.4% contamination. The genome consists of 371 contigs, with an N50 contig length of 6,024 bp and an N50 scaffold length of 6,098 bp. The longest contig and scaffold measure 23,677 bp, while the mean contig and scaffold lengths are 5,363 and 5,606 bp, respectively. The GC content is 34.4% with a coding density of 72.45%. Gene prediction and functional annotation were performed using Prokka v1.14.5 (22), which identified 2,405 predicted protein-coding genes, 35 tRNA genes, and 1 5S rRNA gene. Functional annotation was further refined using EggNOG-mapper v2.1.4 (23), which assigned functions to the predicted genes.

ACKNOWLEDGMENTS

C.C. is supported by the European Commission under the H2020 Marie Skłodowska-Curie Actions Grant Agreement No. 702057 (DRYLIFE). C.C. and C.D. thank the Italian National Program for Antarctic Research for funding sampling campaigns and research activities in Italy in the frame of PNRA projects. The Italian Antarctic National Museum (MNA) is kindly acknowledged for financial support to the Mycological Section of the MNA and for providing rock samples used in this study stored in the Culture Collection of Antarctic fungi (MNA-CCFEE), University of Tuscia.

AUTHOR AFFILIATIONS

¹Department of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy

²Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Italy

³Department of Microbiology and Plant Pathology, Institute of Integrative Genome Biology, University of California, Riverside, Riverside, California, USA

⁴Institute for Integrative Genome Biology, University of California, Riverside, Riverside, California, USA

AUTHOR ORCIDs

Claudia Coleine  <http://orcid.org/0000-0002-9289-6179>

Jason E. Stajich  <http://orcid.org/0000-0002-7591-0020>

Claudio Donati  <http://orcid.org/0000-0001-8688-1651>

DATA AVAILABILITY

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession [JBNJLU00000000](https://doi.org/10.1128/JBNJLU00000000). The version described in this paper is version [JBNJLU010000000](https://doi.org/10.1128/JBNJLU010000000), and it is associated with BioProject [PRJNA905198](https://doi.org/10.1128/PRJNA905198) (Biosample [SAMN31867300](https://doi.org/10.1128/SAMN31867300)).

REFERENCES

- Friedmann EI. 1982. Endolithic microorganisms in the Antarctic cold desert. *Science* 215:1045–1053. <https://doi.org/10.1126/science.215.453.6.1045>
- Archer SD, de los Ríos A, Lee KC, Niederberger TS, Cary SC, Coyne KJ, Douglas S, Lacap-Bugler DC, Pointing SB. 2017. Endolithic microbial diversity in sandstone and granite from the McMurdo Dry Valleys, Antarctica. *Polar Biol* 40:997–1006. <https://doi.org/10.1007/s00300-016-2024-9>
- Wierzbos J, de los Ríos A, Ascaso C. 2012. Microorganisms in desert rocks: the edge of life on Earth. *Int Microbiol* 15:173–183. <https://doi.org/10.2436/20.1501.01.170>
- Cowan DA, Ferrari BC, McKay CP. 2022. Out of thin air? astrobiology and atmospheric chemotrophy. *Astrobiology* 22:225–232. <https://doi.org/10.1089/ast.2021.0066>
- Ji M, Greening C, Vanwonterghem I, Carere CR, Bay SK, Steen JA, Montgomery K, Lines T, Beardall J, van Dorst J, Snape I, Stott MB, Hugenholtz P, Ferrari BC. 2017. Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature* 552:400–403. <https://doi.org/10.1038/nature25014>
- Albanese D, Coleine C, Rota-Stabelli O, Onofri S, Tringe SG, Stajich JE, Selbmann L, Donati C. 2021. Pre-cambrian roots of novel Antarctic cryptoendolithic bacterial lineages. *Microbiome* 9:1–15. <https://doi.org/10.1186/s40168-021-01021-0>
- Coleine C, Gevi F, Fanelli G, Onofri S, Timperio AM, Selbmann L. 2020. Specific adaptations are selected in opposite sun exposed Antarctic cryptoendolithic communities as revealed by untargeted metabolomics. *PLoS One* 15:e0233805. <https://doi.org/10.1371/journal.pone.0233805>
- Lehtovirta-Morley LE. 2018. Ammonia oxidation: ecology, physiology, biochemistry and why they should matter to all microbial ecologists. *FEMS Microbiol Lett* 365:fny057. <https://doi.org/10.1093/femsle/fny058>
- Alves RJE, Minh BQ, Urich T, et al. 2018. A unified catalog of 2,776 archaeal genomes provides insights into the evolutionary history of the

- domain. *Nat Commun* 9:1513. <https://doi.org/10.1038/s41467-018-03861-1>
10. Kerou M, Offre P, Valledor L, Abby SS, Melcher M, Nagler M, Weckwerth W, Schleper C. 2016. Proteomics and comparative genomics of *Nitrososphaera viennensis* reveal the core genome and adaptations of archaeal ammonia oxidizers. *Proc Natl Acad Sci USA* 113. <https://doi.org/10.1073/pnas.1601212113>
 11. Bayer B, Vojvoda J, Offre P, Alves RJE, Elisabeth NH, Garcia JA, Volland J-M, Srivastava A, Schleper C, Herndl GJ. 2016. Physiological and genomic characterization of two novel marine *Nitrosopumilus* species indicate niche differentiation. *ISME J* 10:1051–1063. <https://doi.org/10.1038/ismej.2015.200>
 12. Hayatsu M, Tago K, Uchiyama I, Toyoda A, Wang Y, Shimomura Y, Okubo T, Kurisu F, Hirono Y, Nonaka K, Akiyama H, Itoh T, Takami H. 2017. An acid-tolerant ammonia-oxidizing γ -proteobacterium from soil. *ISME J* 11:1130–1141. <https://doi.org/10.1038/ismej.2016.191>
 13. Merino N, Aronson HS, Bojanova DP, Feyhl-Buska J, Wong ML, Zhang S, Giovannelli D. 2019. Living at the extremes: extremophiles and the limits of life in a planetary context. *Front Microbiol* 10:780. <https://doi.org/10.3389/fmicb.2019.00780>
 14. Ettinger CL, Saunders M, Selbmann L, Delgado-Baquerizo M, Donati C, Albanese D, Roux S, Tringe S, Pennacchio C, Del Rio TG, Stajich JE, Coleine C. 2023. Highly diverse and unknown viruses may enhance Antarctic endoliths' adaptability. *Microbiome* 11:103. <https://doi.org/10.1186/s40168-023-01554-6>
 15. Coleine C, Albanese D, Ray AE, Delgado-Baquerizo M, Stajich JE, Williams TJ, Larsen S, Tringe S, Pennacchio C, Ferrari BC, Donati C, Selbmann L. 2024. Metagenomics untangles potential adaptations of Antarctic endolithic bacteria at the fringe of habitability. *Sci Total Environ* 917:170290. <https://doi.org/10.1016/j.scitotenv.2024.170290>
 16. Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. 2017. Nextflow enables reproducible computational workflows. *Nat Biotechnol* 35:316–319. <https://doi.org/10.1038/nbt.3820>
 17. 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>
 18. Albanese D, Donati C. 2021. Large-scale quality assessment of prokaryotic genomes with metashot/prok-quality. *F1000Res* 10:822. <https://doi.org/10.12688/f1000research.54418.1>
 19. Kang DD, Li F, Kirton 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>
 20. 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>
 21. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P-A, Hugenholtz P. 2022. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 50:D785–D794. <https://doi.org/10.1093/nar/gkab776>
 22. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
 23. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* 38:5825–5829. <https://doi.org/10.1093/molbev/msab293>