



Metagenomes in the Borderline Ecosystems of the Antarctic Cryptoendolithic Communities

 Claudia Coleine,^a Davide Albanese,^b Silvano Onofri,^a Susannah G. Tringe,^c Christa Pennacchio,^c Claudio Donati,^b
 Jason E. Stajich,^{d,e} Laura Selbmann^{a,f}

^aDepartment of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy

^bResearch and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Italy

^cDepartment of Energy Joint Genome Institute, Walnut Creek, California, USA

^dDepartment of Microbiology and Plant Pathology, University of California, Riverside, Riverside, California, USA

^eInstitute of Integrative Genome Biology, University of California, Riverside, Riverside, California, USA

^fMycological Section, Italian Antarctic National Museum, Genoa, Italy

Claudia Coleine and Davide Albanese contributed equally to the work. First author order was determined by drawing straws.

ABSTRACT Antarctic cryptoendolithic communities are microbial ecosystems dwelling inside rocks of the Antarctic desert. We present the first 18 shotgun metagenomes from these communities to further characterize their composition, biodiversity, functionality, and adaptation. Future studies will integrate taxonomic and functional annotations to examine the pathways necessary for life to evolve in the extremes.

Antarctic cryptoendolithic communities are microbial ecosystems dwelling inside airspaces of rocks under the superextreme conditions of the ice-free areas of continental Antarctica. In Antarctica, they were first described from the McMurdo Dry Valleys, which are considered one of the best analogues of the Martian environment on Earth and were thought to be devoid of life until the discovery of these cryptic life forms (1). The cryptoendolithic ecosystems are highly specialized, are adapted to exploit a narrow ecological niche, and represent excellent models to investigate how life can persist at the extremes of aridity, solar radiation, and temperature. Recent molecular and genomic studies have shed light on the structure and diversity of some biological functional groups found in these environments (2–7). Here, we report the first shotgun metagenomics of Antarctic endolithic communities to further examine the biodiversity, ecological functions, and potential stress response strategies of the community members. These genomic data can aid in developing models for the organisms from this extreme environment, including those to learn their identity, functional capabilities, ecosystem roles, and mechanisms of adaptation.

Eighteen colonized sandstones were aseptically collected, using a geological hammer and chisel, from sites in Victoria Land (continental Antarctica) along a latitudinal transect ranging from 74°10'44.0"S to 77°52'28.6"S, from 834 to 3,100 m above sea level, during the XXXI Italian Antarctic Expedition (2015 to 2016). Northern and southern sun-exposed rock surfaces at each site were sampled. Collected samples were immediately placed in sterile bags and were kept at –20°C throughout transport and storage at the University of Tuscia (Viterbo, Italy) until processing. Pieces of each sample were pulverized with a sterile hammer, and total DNA was extracted from 1 g of crushed rock using a PowerSoil kit (Mo Bio Laboratories, Carlsbad, CA, USA). DNA was used to prepare paired-end genomic libraries using Nextera DNA kits, at the U.S. Department of Energy (DOE) Joint Genome Institute (JGI), and was sequenced (2 × 151 cycles) on a NovaSeq system (Illumina, Inc., San Diego, CA).

Citation Coleine C, Albanese D, Onofri S, Tringe SG, Pennacchio C, Donati C, Stajich JE, Selbmann L. 2020. Metagenomes in the borderline ecosystems of the Antarctic cryptoendolithic communities. *Microbiol Resour Announc* 9:e01599-19. <https://doi.org/10.1128/MRA.01599-19>.

Editor Christina A. Cuomo, Broad Institute

Copyright © 2020 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/).

Address correspondence to Claudio Donati, claudio.donati@fmach.it, or Jason E. Stajich, jason.stajich@ucr.edu.

Received 5 January 2020

Accepted 5 February 2020

Published 5 March 2020

TABLE 1 Accession numbers and data for the metagenomes in this study

SRA accession no.	BioProject no.	GenBank accession no.	IMG/M identification no.	N_{50} (bp)	Sampling site
SRP176584	PRJNA513364	JAABUH000000000	3300030517	18,453	Battleship Promontory, north
SRP176586	PRJNA513365	JAABUI000000000	3300031909	25,796	Battleship Promontory, south
SRP176587	PRJNA513366	JAABUJ000000000	3300031452	23,579	Mount New Zealand, north
SRP176592	PRJNA513367	JAABUK000000000	3300032162	30,100	Pudding Butte, north
SRP176590	PRJNA513368	JAABUL000000000	3300031453	45,310	Pudding Butte, south
SRP176595	PRJNA513369	JAABUM000000000	3300031451	22,828	Siegfried Peak, north
SRP176596	PRJNA513370	JAABUN000000000	3300031454	16,381	Siegfried Peak, south
SRP176600	PRJNA513372	JAABUO00000000	3300031449	35,176	Finger Mountain, south
SRP176598	PRJNA513373	JAABUP000000000	3300031460	15,853	Linnaeus Terrace, north
SRP176601	PRJNA513371	JAABUQ000000000	3300031520	45,912	Finger Mountain, north
SRP176608	PRJNA513375	JAABUR000000000	3300031450	81,671	University Valley, south
SRP176612	PRJNA513376	JAABUS000000000	3300031473	53,362	Trio Nunatak, north
SRP176609	PRJNA513359	JAABUT000000000	3300031447	18,466	Ricker Hills, north
SRP176611	PRJNA513360	JAABUU000000000	3300030523	16,877	Richard Nunatak, white sandstone
SRP176669	PRJNA513362	JAABUV000000000	3300031448	50,022	Knobhead, north
SRP176664	PRJNA513363	JAABUW000000000	3300031471	49,126	Knobhead, south
SRP176667	PRJNA513361	JAABUX000000000	3300031472	75,065	Richard Nunatak, red sandstone
SRP176606	PRJNA513374	JAABUY000000000	3300031470	43,747	University Valley, north

The BBDuk v.38.25 tool was used to remove contaminants and to trim adapters and low-quality sequences. The procedure removed reads that contained ≥ 4 "N" bases, had an average quality score across the read of < 3 , or had a minimum length of ≤ 51 bp or 33% of the full read length. Trimmed, screened, paired-end reads were corrected using BFC v.r181 (8) (with parameters "-1 -s 10g -k 21"). Reads lacking mate pairs after trimming and quality control were also removed. The trimmed, corrected reads were assembled with metaSPAdes v.3.12.0 (9) (with parameters "-m 2000 --only-assembler -k 68 33,55,77,99,127 --meta"). Coverage was calculated by mapping the filtered sequence reads to the assembly using BMap v.38.25 (<https://github.com/BioInfoTools/BMap>) (with default parameters and "ambiguous=random").

A total of 3,817,654,184 filtered reads were obtained after quality control, with a mean of 212,091,899 reads per sample (minimum, 135,633,610 reads; maximum, 266,930,788 reads), which were assembled into more than 10 million contigs across all samples, with a GC content of $58.4\% \pm 3.2\%$ (mean \pm standard deviation) and N_{50} of $37,096 \pm 19,919$ bp. The DOE JGI Metagenome Annotation Pipeline v.4.16.5 (10), part of the Integrated Microbial Genomes with Microbiome Samples (IMG/M) system v.4 (11), predicted a total of 21,647,468 protein-coding genes across all assemblies.

Data availability. The reads and assemblies were deposited under the NCBI accession numbers listed in Table 1. Assembly, gene prediction, and annotation data sets are available at the IMG/M website (<https://img.jgi.doe.gov>) and in the Zenodo repository (12).

ACKNOWLEDGMENTS

The Italian Antarctic National Museum (MNA) is gratefully acknowledged for financial support to the Mycological Section of the MNA for preserving Antarctic rock samples analyzed in this study and stored in the Culture Collection of Fungi from Extreme Environments, University of Tuscia, Italy. C.C. and L.S. thank the Italian National Program for Antarctic Researches (PNRA) for funding sampling campaigns and research activities in Italy in the frame of PNRA projects. J.E.S. is a CIFAR fellow in the program Fungal Kingdom: Threats and Opportunities. The work conducted by the U.S. DOE JGI, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. DOE under contract DE-AC02-05CH11231. Data analyses performed at the High-Performance Computing Cluster at the University of California, Riverside in the Institute of Integrative Genome Biology were supported by NSF grant DBI-1429826 and NIH grant S10-OD016290.

J.E.S. and L.S. are the co-principal investigator and principal investigator, respec-

tively, of JGI Community Sequencing Project 503708. S.G.T. and C.P. supported the project as leader and project manager, respectively, of the metagenomic group. Samples were collected by L.S. during the XXXI Italian Antarctic Expedition (2015 to 2016). C.C., J.E.S., and L.S. designed the research, C.C. performed DNA extraction and quality control checks, and C.C., D.A., J.E.S., and L.S. wrote the paper, with input from S.O., S.G.T., C.P., and C.D.

We declare no competing interests.

REFERENCES

- Friedmann EI. 1982. Endolithic microorganisms in the Antarctic Cold Desert. *Science* 215:1045–1053. <https://doi.org/10.1126/science.215.4536.1045>.
- 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>.
- Selbmann L, Onofri S, Coleine C, Buzzini P, Canini F, Zucconi L. 2017. Effect of environmental parameters on biodiversity of the fungal component in lithic Antarctic communities. *Extremophiles* 21:1069–1080. <https://doi.org/10.1007/s00792-017-0967-6>.
- Coleine C, Stajich JE, Zucconi L, Onofri S, Pombubpa N, Egidi E, Franks A, Buzzini P, Selbmann L. 2018. Antarctic cryptoendolithic fungal communities are highly adapted and dominated by lecanoromycetes and dothideomycetes. *Front Microbiol* 9:1392. <https://doi.org/10.3389/fmicb.2018.01392>.
- Coleine C, Zucconi L, Onofri S, Pombubpa N, Stajich JE, Selbmann L. 2018. Sun exposure shapes functional grouping of fungi in cryptoendolithic Antarctic communities. *Life* 8:19. <https://doi.org/10.3390/life8020019>.
- Coleine C, Stajich JE, Pombubpa N, Zucconi L, Onofri S, Canini F, Selbmann L. 2019. Altitude and fungal diversity influence the structure of Antarctic cryptoendolithic bacteria communities. *Environ Microbiol Rep* 11:718–726. <https://doi.org/10.1111/1758-2229.12788>.
- Lee J, Cho J, Cho YJ, Cho A, Woo J, Lee J, Hong SG, Sul WJ, Kim O-S. 2019. The latitudinal gradient in rock-inhabiting bacterial community compositions in Victoria Land, Antarctica. *Sci Total Environ* 657:731–738. <https://doi.org/10.1016/j.scitotenv.2018.12.073>.
- Li H. 2015. BFC: correcting Illumina sequencing errors. *Bioinformatics* 31:2885–2887. <https://doi.org/10.1093/bioinformatics/btv290>.
- 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>.
- Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Tennesen K, Palaniappan K, Szeto E, Pillay M, Chen I-M, Pati A, Nielsen T, Markowitz VMM, Kyrpides NC. 2016. The standard operating procedure of the DOE-JGI Metagenome Annotation Pipeline (MAP v.4). *Stand Genomic Sci* 11:17.
- Markowitz VM, Chen I-M, Chu K, Szeto E, Palaniappan K, Pillay M, Ratner A, Huang J, Pagani I, Tringe S, Huntemann M, Billis K, Varghese N, Tennesen K, Mavromatis K, Pati A, Ivanova NN, Kyrpides NC. 2014. IMG/M 4 version of the integrated metagenome comparative analysis system. *Nucleic Acids Res* 42:D568–D573. <https://doi.org/10.1093/nar/gkt919>.
- Coleine C, Stajich JE. 2020. Antarctic cryptoendolithic metagenomes. Zenodo. <https://doi.org/10.5281/zenodo.3610489>.