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,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 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.

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).


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.
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
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 BBMap v.38.25 (https://github.com/BioInfoTools/BBMap) (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% ± 3.2% (mean ± standard deviation) and \( N_{50} \) of 37,096 ± 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


