Genome Sequence of *Janthinobacterium* sp. CG23_2, a Violacein-Producing Isolate from an Antarctic Supraglacial Stream

Heidi J. Smith, a Christine M. Foreman, a Tatsuya Akiyama, a Michael J. Franklin, a Nicolas P. Devitt, b Thiruvarangan Ramarajb

Center for Biofilm Engineering, Montana State University, Bozeman, Montana, USA; a National Center for Genome Resources (NCGR), Santa Fe, New Mexico, USA

Here, we present the draft genome sequence for the violacein-producing *Janthinobacterium* sp. CG23_2 isolated from an Antarctic supraglacial stream. The genome is ~7.85 Mb, with a G+C content of 63.5%. The genome includes 7,247 candidate protein coding genes, which may provide insight into UV tolerance mechanisms.

Organisms inhabiting Antarctic supraglacial environments are subjected to a variety of environmental stresses. Ozone depletion over Antarctica has altered the spectral composition of solar radiation (1), and harmful UVC (100 to 280 nm) radiation is present on ice surfaces (2). Inescapable to microorganisms throughout the course of the austral summer, UV radiation damages biological macromolecules, including nucleic acids, lipids, and proteins and may lead to cell death. In order for microorganisms to inhabit environments with high UV radiation, efficient protective mechanisms and DNA and protein repair mechanisms are necessary. One mechanism for the protection from UV induced biological damage is for cells to produce protective pigments (3, 4).

The purple pigment violacein is produced by certain members of the β-proteobacteria, including some *Janthinobacterium* strains. Violacein has been investigated for its antimicrobial (5), antioxidant (6), and UV protection properties (4). Although the exact mechanism of violacein protection against UV damage is currently not well understood, evidence suggests that violacein can detoxify free radicals induced by UV radiation (4).

*Janthinobacterium* spp. are found in many different environments, including lakes, soils, and glaciers (7–9). We recently reported the genome sequence of a nonpigmented *Janthinobacterium* sp. isolated from a supraglacial stream. Here we present the genome sequence of a violacein producing strain *Janthinobacterium* sp. CG23_2, isolated from the same system in Antarctica. Genomic analyses of these strains will offer insight into UV tolerance mechanisms from environmentally relevant isolates.

*Janthinobacterium* sp. strain CG23_2 was isolated from the Cotton Glacier in the Antarctic Dry Valleys (77° 07’S, 161° 50’E). The organism was isolated on R2A agar medium incubated in the dark at 4°C for 12 days. *Janthinobacterium* sp. strain CG23_2 is a psychrotolerant, aerobic, violacein-pigmented, rod-shaped, Gram-negative, catalase-positive organism. Genomic DNA was isolated following standard cetyltrimethylammonium bromide (CTAB) isolation protocols (http://www.jgi.doe.gov).

Whole-genome DNA sequencing was performed using a Pacific Biosciences (PacBio, Menlo Park, CA) RS II instrument (10). A single molecule real-time (SMRT) cell library was constructed with 10 µg input DNA using the PacBio 20-kbp protocol. The library was then loaded onto one SMRT cell and sequenced using P5 polymerase and C3 chemistry with 180-min movie times. Sequencing yielded a total of 99,287 reads with mean read length of 5.9 kbps, totaling 581,513,481 bps (~85-fold coverage). *De novo* assembly was constructed using the hierarchical genome assembly process (HGAP2) protocol from SMRT Analysis v2.0, including consensus polishing with Quiver (11, 12). The final assembly consists of four contigs with a total genome size of ~7.85 Mbps. Approximately 93% of the genome is contained within two large contigs (4.2 and 3.1 Mbps). Remaining sequences were divided into two smaller contigs ranging from 420 to 153 kbps. A total of 7,247 candidate protein-coding genes were predicted using RAST (13) with a total G+C content of 63.5%. Upon comparison with genomes available within the RAST the closest relative to *Janthinobacterium* sp. CG23_2 was determined to be *Janthinobacterium* sp. Marseille (score 542).

**Nucleotide sequence accession numbers.** This genome sequence has been deposited in EMBL/GenBank under the accession number CYSS00000000. The version described in this paper is the first version, CYSS00000000.1.

**ACKNOWLEDGMENTS**

Funding for this research came from the National Science Foundation (OPP-0838970 and 1141978), and the Department of Microbiology at Montana State University, H.J.S. was supported by the NASA Earth and Space Science Fellowship (NESSF) program.

Logistical support was provided by Raytheon Polar Services and Petroleum Helicopters Incorporated.

Any opinions, findings, and conclusions or recommendations expressed in this paper do not necessarily reflect the views of the National Science Foundation.
FUNDING INFORMATION
NASA Earth and Space Science Fellowship provided funding to Heidi J. Smith. National Science Foundation (NSF) provided funding to Christine M. Foreman under grant number OPP-0838970. National Science Foundation (NSF) provided funding to Christine M. Foreman under grant number ANT-1141978.

REFERENCES