Draft Genome Sequence of *Sphingomonas* sp. Strain Ant20, Isolated from Oil-Contaminated Soil on Ross Island, Antarctica

Sandra Ronca,[a] Aline Frossard,[a] Leandro D. Guerrero,[a] Thulani P. Makhalanyane,[a] Jackie M. Aislabie,[b] Don A. Cowan[a]

Centre for Microbial Ecology and Genomics, University of Pretoria, Pretoria, South Africa[a]; Landcare Research, Hamilton, New Zealand[b]

Here, we present the draft genome of *Sphingomonas* sp. strain Ant20, isolated from oil-polluted soil near Scott Base, Ross Island, Antarctica. The genome of this aromatic hydrocarbon-degrading bacterium provides valuable information on the microbially mediated biodegradation of aromatic compounds in cold-climate systems.

Address correspondence to Don A. Cowan, don.cowan@up.ac.za.

Antarctic exploration has led to localized, although significant, impacts on the environment. For example, fuel spills can have substantial environmental impacts on Antarctic ice-free soils through the accumulation of aliphatic and aromatic compounds (1, 2). The protocol for contaminated-soil recovery requires that soil remediation not result in greater environmental impact than that with the "do-nothing" approach (3). To achieve this goal, bioremediation is viewed as an alternative technology for cold-climate soils (4, 5). As the Antarctic Treaty prevents the introduction of foreign organisms into the continent, the isolation of indigenous bacterial strains with bioremediation capabilities is the preferred approach. The Gram-negative *Sphingomonas* sp. strain Ant20, isolated from oil-contaminated soil near Scott Base, Antarctica, was observed to utilize aromatic hydrocarbons, such as phenanthrene and 1-methylnaphthalene, as growth substrates (6). We sequenced the genome of this bacterial strain to unravel its genetic potential as a biodegrader.

Sphingomonas sp. Ant20 was grown on Reasoner's 2A (R2A) medium at 15°C. DNA was extracted according to a modified protocol from Miller and colleagues (7), and the genome was sequenced on an Ion Torrent PGM sequencer (318 Chip) with 400-bp chemistry (Life Technologies) at the DNA Sequencing Facility of the University of Pretoria, Pretoria, South Africa. After quality filtering using in-house scripts, 2,403,622 reads with an average size of 241 bp were assembled using MIRA 4.Orc4 (8). The connections between the resulting contigs with a minimum length of 1,000 bp were checked on cytoscape 3.x (9) and assembled using GAP5 (10). The final draft genome contained 299 contigs, with a mean size of 15,009 bp and a maximum length of 191,481 bp. The total genome size was 4,487,680 bp, with a mean G+C content of 65.6% and an average coverage of 132×. The genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (http://www.ncbi.nlm.nih.gov/genome/annotation_prok) and Rapid Annotations using Subsystems Technology (RAST) (11), which identified 4,580 protein-coding sequences (1,374 of annotated function and 3,206 hypothetical proteins). A comparison of the 16S rRNA gene sequences on NCBI indicated 96% identity with *Sphingomonas taxi* strain ATCC 5566 (GenBank accession no. CP009571.1). Annotation in RAST showed that interesting genes/operons were present in the *Sphingomonas* sp. Ant20 genome, including genes related to resistance to antibiotics and toxic compounds, such as genes encoding metal efflux pump proteins (e.g., for arsenic, cobalt-zinc-cadmium, and copper). Several stress-related proteins were identified, including cold and heat shock, oxidative, and carbon starvation proteins. Moreover, a number of proteins involved in various aromatic compound degradation pathways were found, including pathways for quinate, benzoate, beta-ketoadipate, *N*-heterocyclic aromatic, and aromatic amine compounds. However, no protein was predicted for phenanthrene and naphthalene degradation, despite the fact that Ant20 strain has been shown experimentally to degrade these compounds (6).

The analysis of the draft genome of *Sphingomonas* sp. Ant20 thus indicates that this organism may be considered a candidate for bioremediation in cold soils. In addition, the availability of previously annotated genomes of bacterial strains also isolated from oil-contaminated soils in Antarctica will allow comparative genomic studies and the comprehension of factors underlying the ability to degrade aromatic compounds.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession no. JRVI00000000. The version described in this paper is version JRVI01000000.

REFERENCES

1. **Aislabie JM, Balks MR, Foght JM, Waterhouse EJ.** 2004. Hydrocarbon spills on Antarctic soils: effects and management. Environ Sci Technol **38:**1265–1274. http://dx.doi.org/10.1021/es0305149.

2. **Cowan DA, Chown SL, Convey P, Tuffin M, Hughes K, Pointing S, Vincent WF.** 2011. Non-indigenous microorganisms in the Antarctic: assessing the risks. Trends Microbiol **19:**540–548. http://dx.doi.org/10.1016/j.tim.2011.07.008.

3. **Stark JS, Snape I, Riddle MJ.** 2006. Abandoned Antarctic waste disposal sites: monitoring remediation outcomes and limitations at Casey Station. Ecol Manag Restor **7:**21–31. http://dx.doi.org/10.1111/j.1442-8903.2006.00243.x.

4. **Kauppi S, Sinkkonen A, Romantschuk M.** 2011. Enhancing bioremediation of diesel-fuel-contaminated soil in a boreal climate: comparison of biostimulation and bioaugmentation. Int Biodeterior Biodegrad **65:**359–368. http://dx.doi.org/10.1016/j.ibiod.2010.10.011.

5. **Vázquez S, Nogales B, Ruberto L, Hernández E, Christie-Oleza J, Lo Balbo A, Bosch R, Lalucat J, Mac Cormack W.** 2009. Bacterial community dynamics during bioremediation of diesel oil-contaminated Antarctic soil. Microb Ecol **57:**598–610. http://dx.doi.org/10.1007/s00248-008-9420-9.

6. **Aislabie J, Foght J, Saul D.** 2000. Aromatic hydrocarbon-degrading bacteria from soil near Scott Base, Antarctica. Polar Biol **23:**183–188. http://dx.doi.org/10.1007/s003000050025.

7. **Miller DN, Bryant JE, Madsen EL, Ghiorse WC.** 1999. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. Appl Environ Microbiol **65:**4715–4724.

8. **Chevreux B, Wetter T, Suhai S.** 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. In Computer science and biology. Proceedings of the German Conference on Bioinformatics, GCB '99. GCB, Hannover, Germany.

9. **Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T.** 2011. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics **27:**431–432. http://dx.doi.org/10.1093/bioinformatics/btq675.

10. **Bonfield JK, Whitwham A.** 2010. Gap5—editing the billion fragment sequence assembly. Bioinformatics **26:**1699–1703. http://dx.doi.org/10.1093/bioinformatics/btq268.

11. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics **9:**75. http://dx.doi.org/10.1186/1471-2164-9-75.