

# Complete genome sequence of *Cellulomonas* sp. strain ATA003

Alexander Bartholomäus,<sup>1</sup> Julia Mitzscherling,<sup>1</sup> Daniel Lipus,<sup>1</sup> Dirk Wagner,<sup>1,2</sup> Paris Lavin,<sup>3</sup> Roberto Contreras,<sup>4</sup> Rómulo Osés<sup>4</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 2.

**ABSTRACT** The Gram-positive, rod-shaped endophytic bacterium *Cellulomonas* sp. strain ATA003 was isolated from the endemic cactus *Maihueniopsis domeykoensis* seeds collected in the Coastal Atacama Desert, Chile. Here, we present a circular genome with a size of 4,084,881 bp and a GC content of 73.8% obtained by Nanopore sequencing.

**KEYWORDS** endophytes, genomes

The Atacama Desert of northern Chile is an extreme environment characterized by poor soils with low levels of organic matter and nutrients, high-temperature oscillations, UV radiation, and very low annual precipitations. However, the Atacama Desert harbors a rich microbial diversity that has only recently been discovered (1). Here, we identified a seed endophyte (SE) from the Atacama Desert. SEs can colonize endospheric compartments of seeds, providing benefits to host plants such as growth-promoting activities, improved seed germination, or resistance against stresses (2, 3). They are of potential interest for agriculture and medicine (4, 5).

Seeds of *Maihueniopsis domeykoensis* were collected by AELA SpA at Domeyko (28.952722 S, 70.894150 W) and were supplied by the Forestry Institute – INFOR Chile (La Serena, Chile). The surface of the seeds was sterilized by sequential washing in 50% ethanol for 1 min, 2% sodium hypochlorite for 3 min, and 50% ethanol for 30 s, and was rinsed twice with sterile distilled water, followed by grinding the seeds in a clean mortar and suspending in sterile phosphate buffer. The slurry was diluted and spread on “Combined carbon or Rennie semisolid N-free” medium (6) for 24 h at 30°C. After transfer to N-free agar (NFA) and 72 h of incubation at 30°C, individual bacterial colonies were selected and purified by restreaking five times on fresh NFA. All incubation and culturing steps took place under aerobic conditions. Genomic DNA was extracted from a pure culture grown on LB medium at 28°C using the UltraClean microbial DNA isolation kit (MoBio, Carlsbad, CA, USA).

High-molecular-weight DNA was prepared without specific size selection using the rapid sequencing kit SQK-RAD004 (Oxford Nanopore Technologies [ONT], Oxford, UK). The DNA was cleaned using AMPure XP beads (Beckman Coulter, Pasadena, CA, USA). The library was sequenced using the MinION device and the Flongle flow-cell R9.4.1 (ONT). The sequencing ran for 48 h. Default parameters were used for all software unless otherwise specified. The raw sequencing data were base called and demultiplexed with super high accuracy using Guppy v6.0.6 (ONT), resulting in 54,226 reads with an N50 value of 7,051 bp. No further read quality control, trimming, or correction was done. Assembly, polishing, and circularity assessment were performed using Flye v2.9.1 (7) (parameters –meta –nano-raw). The genome was used as-is from Flye without further rotation or flipping of the genome. The genome quality was assessed, and full-length 16S rRNA sequences were recovered using CheckM v1.2.1 (8).

The resulting circular genome was 4,084,881 bp with a GC content of 73.8% and a coverage of 52x. The genome was found to be 69.4% complete and 2.4% contaminated using the *Cellulomonas* marker set of CheckM v1.2.2 (8). The taxonomic affiliation and closest related species were inferred using GTDB-Tk v2.1.0 (9) for the genome and NCBI BLAST (10) (v2.14.0, nr/nt database) for 16S rRNA. The closest species were

**Editor** Leighton Pritchard, University of Strathclyde, Glasgow, United Kingdom

Address correspondence to Alexander Bartholomäus, abartho@gfz-potsdam.de.

The authors declare no conflict of interest.

See the funding table on p. 2.

**Received** 28 October 2023

**Accepted** 15 April 2024

**Published** 2 May 2024

Copyright © 2024 Bartholomäus 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/).

*Cellulomonas cellasea* strain NBRC 3753 (ANI 80.35%, accession [GCF\\_006538745.1](https://doi.org/10.1093/mra/01037-23)) based on the genome and *Cellulomonas fimi* strain NCTC7547 (97.65% identity and 100% query coverage, accession [LR134387.1](https://doi.org/10.1093/mra/01037-23)) based on 16S rRNA sequence. Gene annotation was performed automatically by NCBI using the PGAP pipeline v6.6 (11).

## ACKNOWLEDGMENTS

We acknowledge the support of the following grants: Fondecyt Iniciación grant No. 11190754 (National Agency of Research and Development, ANID) and Concurso Anillo en Areas Tematicas Especificas 2022 ATE220038.

## AUTHOR AFFILIATIONS

<sup>1</sup>GFZ German Research Centre for Geosciences, Section Geomicrobiology, Potsdam, Germany

<sup>2</sup>University of Potsdam, Institute of Geoscience, Potsdam, Germany

<sup>3</sup>Departamento de Biotecnología, Facultad de Ciencias del Mar y Recursos Biológicos, Universidad de Antofagasta, Antofagasta, Chile

<sup>4</sup>Centro Regional de Investigación y Desarrollo Sustentable de Atacama (CRIDESAT), Universidad de Atacama, Copiapó, Chile

## AUTHOR ORCIDS

Alexander Bartholomäus  <http://orcid.org/0000-0003-0970-7304>

Daniel Lipus  <http://orcid.org/0000-0002-5706-9143>

Dirk Wagner  <http://orcid.org/0000-0001-5064-497X>

## FUNDING

Funder	Grant(s)	Author(s)
Agencia Nacional de Investigación y Desarrollo (ANID)	11190754	Rómulo Osés

## DATA AVAILABILITY

The genome of *Cellulomonas* sp. ATA003 was deposited at GenBank under the accession number [GCF\\_031582905.1](https://doi.org/10.1093/mra/01037-23) and the BioProject accession number [PRJNA1010543](https://doi.org/10.1093/mra/01037-23). The raw reads are accessible via the SRA accession number [SRX21545741](https://doi.org/10.1093/mra/01037-23). The 16S full-length rRNA is also available on GenBank under accession number [OR498076](https://doi.org/10.1093/mra/01037-23).

## REFERENCES

- Bull AT, Asenjo JA, Goodfellow M, Gómez-Silva B. 2016. The Atacama desert: technical resources and the growing importance of novel microbial diversity. *Annu Rev Microbiol* 70:215–234. <https://doi.org/10.1146/annurev-micro-102215-095236>
- Berg G, Grube M, Schlöter M, Smalla K. 2014. The plant microbiome and its importance for plant and human health. *Front Microbiol* 5:491. <https://doi.org/10.3389/fmicb.2014.00491>
- White JF, Kingsley KL, Zhang Q, Verma R, Obi N, Dvinskikh S, Elmore MT, Verma SK, Gond SK, Kowalski KP. 2019. Review: endophytic microbes and their potential applications in crop management. *Pest Manag Sci* 75:2558–2565. <https://doi.org/10.1002/ps.5527>
- Mascot-Gómez E, Flores J, López-Lozano NE. 2021. The seed-associated microbiome of four cactus species from Southern Chihuahuan desert. *J Arid Environ* 190:104531. <https://doi.org/10.1016/j.jaridenv.2021.104531>
- Puente ME, Li CY, Bashan Y. 2009. Endophytic bacteria in cacti seeds can improve the development of cactus seedlings. *Environ Exp Bot* 66:402–408. <https://doi.org/10.1016/j.envexpbot.2009.04.007>
- Baldani JL, Reis VM, Videira SS, Boddey LH, Baldani VLD. 2014. The art of isolating nitrogen-fixing bacteria from non-leguminous plants using N-free semi-solid media: a practical guide for microbiologists. *Plant Soil* 384:413–431. <https://doi.org/10.1007/s11104-014-2186-6>
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>
- 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>
- Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-TK: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>