

Draft Genome Sequence of *Bacillus* sp. Strain CMAA 1185, a Cellulolytic Bacterium Isolated from Stain House Lake, Antarctic Peninsula

Suikinaí Nobre Santos,^a Vanessa Nessner Kavamura,^a Rodrigo Gouvêa Taketani,^a Rafael L. F. Vasconcellos,^a Tiago Domingues Zucchi,^b Itamar Soares Melo^a

Laboratory of Environmental Microbiology, Embrapa Environment, Brazilian Agricultural Research Corporation, EMBRAPA, Jaguariúna, São Paulo, Brazil^a; Agrivalle, Pouso Alegre, Minas Gerais, Brazil^b

The aim of this study was to report the genome sequence of the cellulolytic *Bacillus* sp. strain CMAA 1185, isolated from Stain House Lake, Antarctica.

Received 28 March 2015 Accepted 16 April 2015 Published 28 May 2015

Citation Santos SN, Kavamura VN, Taketani RG, Vasconcellos RLF, Zucchi TD, Melo IS. 2015. Draft genome sequence of *Bacillus* sp. strain CMAA 1185, a cellulolytic bacterium isolated from Stain House Lake, Antarctic Peninsula. *Genome Announc* 3(3):e00436-15. doi:10.1128/genomeA.00436-15.

Copyright © 2015 Santos et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Itamar Soares Melo, itamar.melo@embrapa.br.

Antarctica has become an interesting site for bioprospecting microorganisms with biotechnological purposes. This extremely harsh environment usually poses numerous threats to living beings (1) and, notwithstanding, may select organisms that held unique biosynthetic pathways (2). In our survey at Stain House Lake, we isolated bacteria able to degrade cellulose (3). One of these bacterial isolates, CMAA 1185, was revealed to produce a cellulolase active at low temperatures, which may play a crucial, active at low temperatures, which may play a crucial role in this environment, and it also has many industrial applications. Thus, to extend the knowledge of the genes related to hydrolases of biopolymers of strain CMAA 1185, whole-genome sequencing was performed using the Ion Torrent (PGM) platform. Genomic DNA was extracted from a pure culture grown overnight on LB medium using the PureLink genomic DNA kit (Life Technologies). Sequencing was carried out on the Ion 316 Chip sequencer provided in the Ion sequencing kit 200-bp version 2.0, according to the manufacturer's protocol. The genome sequence was *de novo* assembled using the MIRA version 4, CLC Genomics Workbench version 5.5.1, and SeqMan NGen version 4.0.0 packages, and the contigs obtained from the assembly were integrated using CISA (4, 5). The taxonomic position of strain CMAA 1185 was further evaluated using the JSpecies package (6).

The number of reads (<Q20) generated using *Bacillus subtilis* subsp. Natto BEST195 (GenBank accession no. PRJDB2126) as a reference was 2,816,934, which were allocated into 10 contigs, with 136× coverage and a mean length of 1,046,675 bp. The assembled data were analyzed by RAST annotation (7), and the genome size was found to be 3,907,872 bp, comprising 4,263 open reading frames (ORFs). The G+C content was estimated to be 43.3 mol%.

A total of 4,051 genes were identified. Of these, there are 176 pseudogenes, 12 rRNAs genes (5S, 16S, and 23S), 65 tRNA genes, and one noncoding RNA (ncRNA) gene. The genome contains 17 genes copies related to hydrolyzes, mainly glycosyltransferases, polysaccharide lyase, peptidoglycan chitinase

and amylase (8). 16S rRNA analyses revealed that this isolate is most closely related to the type strain of *Bacillus subtilis* (3). These organisms share an average nucleotide identity of 99%.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [LAKC000000000](https://www.ncbi.nlm.nih.gov/nuccore/LAKC000000000). The version described in this paper is version LAKC01000000.

ACKNOWLEDGMENTS

This study was made possible with the financial and logistic support of PROANTAR, Marinha do Brazil. This is part of the activity of INCT Criosfera project. We thank FAPESP for providing the funding for developing this research (grant 2013/16037-0). R.G.T. was funded by FAPESP with a Young Investigators Fellowship (grants 2013/23470-2 and 2013/03158-4).

REFERENCES

- Gouvêa Taketani R, Domingues Zucchi T, Soares Melo I, Mendes R. 2013. Whole-genome shotgun sequencing of *Rhodococcus erythropolis* strain P27, a highly radiation-resistant actinomycete from Antarctica. *Genome Announc* 1(5):e00763-13. <http://dx.doi.org/10.1128/genomeA.00763-13>.
- D'Amico S, Collins T, Marx JC, Feller G, Gerday C. 2006. Psychrophilic microorganisms: challenges for life. *EMBO Rep* 7:385–389. <http://dx.doi.org/10.1038/sj.embor.7400662>.
- Melo IS, Zucchi TD, Silva RE, Vilela ES, Sáber ML, Rosa LH, Pellizari VH. 2014. Isolation and characterization of cellulolytic bacteria from the Stain House Lake, Antarctica. *Folia Microbiol (Praha)* 59:303–306. <http://dx.doi.org/10.1007/s12223-013-0295-x>.
- Vasconcellos RL, Mendes R, Taketani RG, Zucchi TD, Melo IS. 2013. Draft genome sequence of *Pseudomonas* sp. strain CMAA 1215, a plant growth-promoting bacterium isolated from a Brazilian mangrove. *Genome Announc* 1(6):e00995-13. <http://dx.doi.org/10.1128/genomeA.00995-13>.
- Lin SH, Liao YC. 2013. CISA: contig integrator for sequence assembly of bacterial genomes. *PLoS One* 8:. <http://dx.doi.org/10.1371/journal.pone.0060843>.

6. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* **106**: 19126–19131. <http://dx.doi.org/10.1073/pnas.0906412106>.
7. 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>.
8. Park BH, Karpinets TV, Syed MH, Leuze MR, Uberbacher EC. 2010. CAZymes analysis toolkit (CAT): Web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. *Glycobiology* **20**:1574–1584. <http://dx.doi.org/10.1093/glycob/cwq106>.