Title

Complete Genome Sequence of Bacillus cereus Strain CPT56D-587-MTF, Isolated from a Nitrate- and Metal-Contaminated Subsurface Environment

Journal

Microbiology Resource Announcements, 11(5)

ISSN

2576-098X

Authors

Goff, Jennifer L
Lui, Lauren M
Nielsen, Torben N
et al.

Publication Date

2022-05-19

DOI

10.1128/mra.00145-22

Peer reviewed

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Bacillus cereus* Strain CPT56D-587-MTF, Isolated from a Nitrate- and Metal-Contaminated Subsurface Environment

Jennifer L. Goff,[a] Lauren M. Lui,[b] Torben N. Nielsen,[b] Michael P. Thorgersen,[a] Elizabeth G. Szink,[a] John-Marc Chandonia,[b] Farris L. Poole II,[a] Jizhong Zhou,[c,d] Terry C. Hazen,[e,f] Adam P. Arkin,[b,g] Michael W. W. Adams[a]

[a]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia, USA
[b]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA
[c]Institute for Environmental Genomics, University of Oklahoma, Norman, Oklahoma, USA
[d]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma, USA
[e]Oak Ridge National Lab, Oak Ridge, Tennessee, USA
[f]Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, Tennessee, USA
[g]Department of Bioengineering, University of California, Berkeley, Berkeley, California, USA

Jennifer L. Goff and Lauren M. Lui contributed equally to this article. Equal contribution author order was determined alphabetically.

ABSTRACT *Bacillus cereus* strain CPT56D-587-MTF was isolated from nitrate- and toxic metal-contaminated subsurface sediment at the Oak Ridge Reservation (ORR) (Oak Ridge, TN, USA). Here, we report the complete genome sequence of this strain to provide genomic insight into its strategies for survival at this mixed-waste site.

Strain CPT56D-587-MTF was isolated from subsurface sediment (long −84.27335°, lat 35.977268°, depth 535.94 cm) from the Oak Ridge Reservation (Oak Ridge, TN, USA), which is contaminated with legacy uranium and nitrate waste (1). Sediment (1 g) was inoculated into anoxic modified Reasoner's 2A (R2A) medium (2) (10 mM nitrite, 100 mM KH_2PO_4, pH 5.5). Following room temperature incubation for a week, strain CPT56D-587-MTF was recovered by streak-plating onto LB agar. A complete genome sequence was obtained to provide insight into its survival strategies. CPT56D-587-MTF was grown for 24 h in R2A medium (30°C, 200 rpm). The cell pellet was digested by resuspension in 750 μL phosphate-buffered saline (PBS) and incubated at 37°C for 30 min with 25 μL MetaPolyzyme (Sigma-Aldrich) and 25 μL lytic enzyme solution (Qiagen), followed by digestion in 167 μL 6× buffer B1 (300 mM Tris-Cl [pH 8.0], 300 mM EDTA [pH 8.0], 3% Tween 20, 3% Triton X-100) (Qiagen), 35 μL proteinase K, and 2 μL RNase A with incubation at 50°C and 50 rpm for 30 min. The lysate was processed using the Genomic-tip 20/G kit (Qiagen) per the manufacturer's directions.

High-molecular-weight (HMW) DNA was prepped for Nanopore sequencing. End repair was performed using the NEBNext companion module for Oxford Nanopore Technologies ligation sequencing (New England BioLabs) according to the manufacturer's instructions. The native barcoding expansion (EXP-NBD104; Oxford Nanopore Technologies) and ligation sequencing (LSK-SQK109; Oxford Nanopore Technologies) kits were used for barcoding and adapter ligation, respectively. The library was sequenced on a R9.4.1 flow cell on a MinION device (Oxford Nanopore Technologies). HMW DNA was prepped for Illumina library creation by needle shearing. The DNA was not size selected. The Illumina library was generated using the Illumina DNA prep kit (catalog number 20018705; previously, Nextera) with indices from set A primers (catalog number 20027213) according to the manufacturer's instructions and sequenced using 2 × 150-bp reads on a NovaSeq 6000 instrument by Novogene.

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

Address correspondence to Michael W. W. Adams, adamsm@uga.edu.

The authors declare no conflict of interest.

Received 17 February 2022
Accepted 9 April 2022
Published 27 April 2022

TABLE 1 Contig assembly information

Contig	Length (bp)	Circularity	Coverage (×)	GenBank accession no.
1	5,668,734	Circular	112	CP090081.1
2	448,451	Circular	269	CP090082.1
3	99,643	Circular	260	CP090083.1
4	83,506	Circular	312	CP090084.1
5	83,392	Circular	236	CP090085.1
6	72,739	Linear	310	CP090086.1
7	70,091	Circular	250	CP090087.1
8	15,910	Circular	3,740	CP090088.1
9	5,876	Circular	1,242	CP090089.1

The sequencing read data were quality filtered and trimmed before assembly. For the Illumina data, adapters were removed in-house by Novogene. After quality filtering, there were 3,246,490 Illumina reads. These data were processed using BBTools v38.86. The processing was done in two passes (3): (i) bbduk.sh was run (*ktrim = r k = 23 mink = 11 hdist = 1 ref = adapters.fa tbo tpe 2*) to remove the remaining Illumina adapters given in adapters.fa (standard Illumina adapters); (ii) for quality filtering and trimming and to remove Illumina PhiX174 spike-ins given in the file phix174 Illumina.fa, bbduk.sh was run again (*bf1 k = 27 hdist = 1 qtrim = rl trimq = 17 cardinality = t ref = phix174_Illumina.fa*). The Nanopore sequencing yielded 136,213 raw reads (N_{50}/N_{90}, 11,659/2,738 bp). Nanopore base calling, adapter removal, demultiplexing, and quality filtering were performed using Guppy v4.0. Assembly was performed with the Nanopore and Illumina reads using the hybrid assembler Unicycler v0.4.8 (4) (default parameters). The Unicycler logs were checked to confirm that the assembly passed the quality thresholds and that the DNA elements were circularized. This is indicated in the "Rotating completed replicon" section of the Unicycler log. Only contig 6 was not circular.

The completed genome contains 6,548,342 bp in 9 contigs (G+C content, 35.37%). Contig 1 is the circularized chromosome, and contigs 2 to 9 are putative plasmids (Table 1). Genome annotation was performed using RASTtk v1.073 (5) in the DOE Systems Biology Knowledgebase (KBase) with default parameters (https://kbase.us/n/105874/55/) (6, 7). Taxonomic assignment performed using GTDB-Tk-v1.7.0 (8) in KBase (default parameters) identified the strain as a *Bacillus cereus* species (GenBank accession number GCA_000007825.1), with an average nucleotide identity of 98.58%. Finally, the assembled genome was deposited in GenBank and reannotated using the NCBI Prokaryotic Genome Annotation Pipeline v5.3 (9).

Data availability. This whole-genome sequencing project has been deposited at GenBank under the accession number GCA_021391515.1. The raw sequence reads have been deposited in the SRA under the accession numbers SRR17696030 (Illumina short reads) and SRR17696029 (Oxford Nanopore long reads). All KBase analyses are publicly available in the KBase static narrative (https://kbase.us/n/105874/55/).

ACKNOWLEDGMENTS

We thank Andrew Putt, Erin Kelly, Kenneth Lowe, and Miguel Rodriguez, Jr., for collecting soil samples.

This material by ENIGMA (Ecosystems and Networks Integrated with Genes and Molecular Assemblies) (http://enigma.lbl.gov), a Science Focus Area Program at Lawrence Berkeley National Laboratory, is based on work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under contract DE-AC02-05CH11231.

REFERENCES

1. Brooks SC. 2001. Waste characteristics of the former S-3 ponds and outline of uranium chemistry relevant to NABIR Field Research Center studies. NABIR Field Research Center, Oak Ridge, TN.
2. Reasoner DJ, Geldreich EE. 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl Environ Microbiol 49:1–7. https://doi.org/10.1128/aem.49.1.1-7.1985.
3. Lui LM, Nielsen TN, Arkin AP. 2021. A method for achieving complete microbial genomes and improving bins from metagenomics data.

PLoS Comput Biol 17:e1008972. https://doi.org/10.1371/journal.pcbi.1008972.
4. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
5. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.
6. Goff J. 2022. Complete genome of ORR isolate Bacillus cereus CPT56D-587-MTF. DOE Systems Biology Knowledgebase. https://kbase.us/n/105874/55/. Accessed 14 February 2022.
7. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, et al. 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569. https://doi.org/10.1038/nbt.4163.
8. 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.
9. 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.