AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Clostridium* sp. Strain E02, Isolated from an Estuarine Environment

Michael Christopher Macey,[a] Elliot Curtis-Harper,[b] Karen Olsson-Francis[a]

[a]School of Environment, Earth and Ecosystem Sciences, Faculty of Science, Technology, Engineering and Mathematics, The Open University, Milton Keynes, United Kingdom

[b]School of Physical Sciences, Faculty of Science, Technology, Engineering and Mathematics, The Open University, Milton Keynes, United Kingdom

ABSTRACT Here, we report the draft genome sequence of a strain of *Clostridium* isolated from sediment collected from an estuarine environment. The strain was isolated using a minimal medium designed to select for chemoautotrophic microorganisms. The strain may represent a novel species within the genus *Clostridium*, and this genome sequence enables further investigation into the genetic and metabolic diversity of this organism.

Anoxic sediments in estuarine systems are proposed as analogue sites to investigate the habitability of the ancient lake system at Gale Crater on Mars because of similarities in temperature, salinity, redox, and pH regimes (1–5). Characterization of microbes at these sites allows for hypothesis development with regard to potential viable Martian metabolisms. *Clostridium* sp. strain E02 was isolated from sediment from the River Dee, which flows into the Liverpool Bay, in the United Kingdom (53°21′15.40″ N, 3°10′24.95″ W). The isolation involved enrichment of sediment in a minimal growth medium (5). *Clostridium* sp. strain E02 is a rod-shaped, anaerobic, Gram-positive bacterium belonging to the family *Clostridiaceae* in the order *Clostridiales* (5).

The strain was cultured in anaerobic lysogeny broth (LB) at 25°C prior to DNA extraction using the Griffiths technique (6). Genome sequencing was performed by MicrobesNG (https://microbesng.uk/). DNA quantification and library preparation were carried out on a Hamilton Microlab Star automated liquid handling system. Libraries were quantified using the Kapa Biosystems library quantification kit for Illumina on a Roche light cycler 96 quantitative PCR (qPCR) machine. Libraries were sequenced on an Illumina HiSeq instrument, using a 250-bp paired-end protocol. A total of 553,783 trimmed reads were produced using Trimmomatic (v0.30) with a sliding window quality cutoff of Q15. *De novo* assembly was performed using SPAdes (v3.7; default settings) (7, 8). Coverage of 60× was achieved during sequencing, calculated using BWA, SAMtools (v0.1.19), and BEDTools genomecov (v2.2.7), all with default settings (9–11). Annotation was performed using the Rapid Annotations using Subsystems Technology (RAST) annotation server (v2.0) with the classic RAST pipeline (12). The presence of genes involved in nitrogen and sulfur cycling was investigated to assess the potential for growth in oligotrophic conditions. Gene screening was performed with BlastKOALA (v2.1; default settings) (13) and supported with BLAST searches against the genome sequence using BioEdit (v7.0.5; default settings) (14).

The draft genome of *Clostridium* sp. strain E02 is 4.08 Mb in size, with a G+C content of 40% and an N_{50} value of 146,559 bp. The genome is composed of 114 contigs, including 3,850 coding sequences, one 16S rRNA gene copy, and 68 tRNAs. Analysis of the 16S rRNA gene using the SILVA alignment, classification and tree service (15) identified *Clostridium* sp. strain E02 as closely related to *Clostridium saccharolyticum* WM1, with 98% sequence identity. The genomes of both strains were compared using

Citation Macey MC, Curtis-Harper E, Olsson-Francis K. 2019. Draft genome sequence of *Clostridium* sp. strain E02, isolated from an estuarine environment. Microbiol Resour Announc 8:e01579-18. https://doi.org/10.1128/MRA.01579-18.

Editor J. Cameron Thrash, Louisiana State University

Copyright © 2019 Macey et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Michael Christopher Macey, michael.macey@open.ac.uk.

Received 30 November 2018
Accepted 11 December 2018
Published 10 January 2019

the genome to genome distance calculator (GGDC) (v2.1; default settings) (16) and the average nucleotide identity (ANI) calculator (default settings) (17). GGDC (23.1%) and ANI (76.5%) scores support that *Clostridium* sp. strain E02 is genetically distinct from the most related species. The genome was shown to contain genes encoding a molybdenum-iron nitrogenase (*nifDKH*), an assimilatory sulfate reductase (*cysND*), and an adenylsulfate reductase (*aprAB*).

Data availability. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number RJLQ00000000. The version described in this paper is version RJLQ01000000. The strain is available from the authors upon request. Raw sequencing reads for *Clostridium* sp. strain E02 are available in the NCBI Sequence Read Archive under accession number SRR8246104.

ACKNOWLEDGMENTS

This work was supported by an STFC-funded studentship awarded to Elliot Curtis-Harper.

Genome sequencing was provided by MicrobesNG (http://www.microbesng.uk), which is supported by the BBSRC (grant number BB/Lo24209/1).

REFERENCES

1. Bridges JC, Schwenzer SP, Leveille R, Westall F, Wiens RC, Mangold N, Bristow T, Edwards P, Berger G. 2015. Diagenesis and clay mineral formation at Gale Crater, Mars. J Geophys Res Planets 120:1–19. https://doi.org/10.1002/2014JE004757.
2. Arvidson RE, Squyres SW, Bell JF, Catalano JG, Clark BC, Crumpler LS, De Souza PA, Fairén AG, Farrand WH, Fox VK, Gellert R, Ghosh A, Golombek MP, Grotzinger JP, Guinness EA, Herkenhoff KE, Jolliff BL, Knoll AH, Li R, McLennan SM, Ming DW, Mittlefehldt DW, Moore JM, Morris RV, Murchie SL, Parker TJ, Paulsen G, Rice JW, Ruff SW, Smith MD, Wolff MJ. 2014. Ancient aqueous environments at Endeavour Crater, Mars. Science 343:1248097. https://doi.org/10.1126/science.1248097.
3. Telesh IV, Khlebovich VV. 2010. Principal processes within the estuarine salinity gradient: a review. Mar Pollut Bull 61:149–155. https://doi.org/10.1016/j.marpolbul.2010.02.008.
4. Bristow TF, Bish DL, Vaniman DT, Morris RV, Blake DF, Grotzinger JP, Rampe EB, Crisp JA, Achilles CN, Ming DW, Ehlmann BL, King PL, Bridges JC, Eigenbrode JL, Sumner DY, Chipera SJ, Moorokian JM, Treiman AH, Morrison SM, Downs RT, Farmer JD, Marais DD, Sarrazin P, Floyd MM, Mischna MA, McAdam AC. 2015. The origin and implications of clay minerals from Yellowknife Bay, Gale crater, Mars. Am Mineral 100:824–836. https://doi.org/10.2138/am-2015-5077CCBYNCND.
5. Curtis-Harper E, Pearson V, Summers S, Bridges J, Schwenzer SP, Olsson-Francis K. 2018. The microbial community of a terrestrial anoxic intertidal zone: a model for laboratory-based studies of potentially habitable ancient lacustrine systems on Mars. Microorganisms 6:E61. https://doi.org/10.3390/microorganisms6030061.
6. Griffiths RI, Whiteley AS, Donnell AGO, Bailey MJ. 2000. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. Appl Environ Microbiol 66:5488–5491. https://doi.org/10.1128/AEM.66.12.5488-5491.2000.
7. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
8. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
9. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
10. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
11. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033.
12. 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. https://doi.org/10.1186/1471-2164-9-75.
13. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol 428:726–731. https://doi.org/10.1016/j.jmb.2015.11.006.
14. Hall AT. 2011. BioEdit: an important software for molecular biology. GERF Bull Biosci 2:60–61.
15. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28:1823–1829. https://doi.org/10.1093/bioinformatics/bts252.
16. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60. https://doi.org/10.1186/1471-2105-14-60.
17. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81–91. https://doi.org/10.1099/ijs.0.64483-0.