AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

A Hybrid Sequencing Approach Completes the Genome Sequence of *Thermoanaerobacter ethanolicus* JW 200

Richard K. Tennant,[a] Monica L. Ayine,[a] Ann L. Power,[a] James A. Gilman,[a] Mark Hewlett,[a] Paul B. C. James,[a] Chloe Singleton,[a] David A. Parker,[a] John Love[a]

[a]Biosciences, College of Environmental and Life Sciences, The BioEconomy Centre, University of Exeter, Exeter, United Kingdom

ABSTRACT *Thermoanaerobacter ethanolicus* JW 200 has been identified as a potential sustainable biofuel producer due to its ability to readily ferment carbohydrates to ethanol. A hybrid sequencing approach, combining Oxford Nanopore and Illumina DNA sequence reads, was applied to produce a single contiguous genome sequence of 2,911,280 bp.

Thermoanaerobacter ethanolicus JW 200 is a thermophilic Gram-positive obligate anaerobe originally isolated from Yellowstone National Park (1). *T. ethanolicus* JW 200 readily ferments glucose to ethanol and has received attention as a sustainable biofuel producer (2) and for its bioconversion capabilities (3).

T. ethanolicus JW 200 was obtained from the DSMZ (DSMZ 2246) and was cultured in *Thermoanaerobacter* medium (DSMZ 61) at 65°C. DNA was purified in an anaerobic cabinet using a combination of the MP Bio FastDNA spin kit and a high-molecular-weight genomic DNA (gDNA) extraction protocol (4). An overnight culture was resuspended in CLS-TC buffer (MP Bio, USA) and transferred to a tube containing lysing matrix A (MP Bio, USA). Lysis tubes were placed into MP Bio FastPrep and homogenized at 6.0 $m \cdot s^{-1}$ for 40 s. Centrifuged lysates were treated with RNase A, and genomic DNA was isolated using AMPure XP beads. Purified DNA was quantified using a Qubit fluorometer (Thermo Fisher Scientific, UK), and DNA quality was measured using the genomic TapeStation assay tapes (Agilent Technologies, USA). Illumina DNA sequencing libraries were prepared with the Nextera XT library prep kit (Illumina, USA) and sequenced by an Illumina MiSeq instrument, using 150-bp paired-end sequencing. Oxford Nanopore libraries were prepared using the SQK-LSK108 1D genomic DNA ligation kit (Oxford Nanopore, UK) and sequenced on a MinION instrument using an R9.4a flow cell (Oxford Nanopore). Oxford Nanopore data were base called using Albacore v2.3.3, and DNA sequence data were verified using Porechop (5), yielding 11.37 Gbp of data, with an N_{50} value of 6,401 bp. Illumina DNA sequence data were quality controlled and filtered using TrimGalore v0.3.3 (6) with the "paired" parameter selected, which yielded 2,031,855 paired-end reads with an N_{90} value of 138 bp. A hybrid assembly of the Oxford Nanopore and Illumina data was performed by MaSuRCA v3.2.3 (7) using the default parameters. The genome was assembled to a single 2,911,280-bp contiguous sequence with a GC content of 34.2% and more than 3,500-fold genome coverage. The assembled genome was verified by aligning the reads against the *de novo* assembled genome using BWA-MEM v0.7.15 (8) with the default parameters. The alignment was visualized in Tablet v1.17 (9) to ensure complete coverage. The completed *T. ethanolicus* genome was annotated using Prokka v1.12 (10), and 2,818 coding sequences (CDS) were identified.

Data availability. The complete genome sequence of *T. ethanolicus* JW 200 is deposited in GenBank under the accession number CP033580. Illumina and Oxford

Citation Tennant RK, Ayine ML, Power AL, Gilman JA, Hewlett M, James PBC, Singleton C, Parker DA, Love J. 2019. A hybrid sequencing approach completes the genome sequence of *Thermoanaerobacter ethanolicus* JW 200. Microbiol Resour Announc 8:e01530-18. https://doi.org/10.1128/MRA.01530-18.

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

Address correspondence to John Love, J.Love@exeter.ac.uk.

Received 13 November 2018
Accepted 10 December 2018
Published 17 January 2019

Nanopore DNA sequence reads have been deposited in the NCBI Sequence Read Archive (accession numbers SRR8113455 and SRR8113456).

ACKNOWLEDGMENTS

This work was supported by a grant from Shell Research Ltd.

We acknowledge the Exeter Sequencing Service for its assistance in sequencing the Illumina libraries.

We declare no competing interests.

REFERENCES

1. Wiegel J, Ljungdahl LG. 1981. *Thermoanaerobacter ethanolicus* gen. nov., spec. nov., a new, extreme thermophilic, anaerobic bacterium. Arch Microbiol 128:343–348. https://doi.org/10.1007/BF00405910.
2. Shao X, Zhou J, Olson DG, Lynd LR. 2016. A markerless gene deletion and integration system for *Thermoanaerobacter ethanolicus*. Biotechnol Biofuels 9:100. https://doi.org/10.1186/s13068-016-0514-1.
3. Hild HM, Stuckey DC, Leak DJ. 2003. Effect of nutrient limitation on product formation during continuous fermentation of xylose with *Thermoanaerobacter ethanolicus* JW200 Fe(7). Appl Microbiol Biotechnol 60:679–686. https://doi.org/10.1007/s00253-002-1175-5.
4. Mayjonade B, Gouzy J, Donnadieu C, Pouilly N, Marande W, Callot C, Langlade N, Muños S. 2016. Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. Biotechniques 61:203–205. https://doi.org/10.2144/000114460.
5. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom 3:e000132. https://doi.org/10.1099/mgen.0.000132.
6. Kruegar F. 2014. Trim Galore! Babraham Bioinformatics, Cambridge, United Kingdom. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
7. Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. Bioinformatics 29:2669–2677. https://doi.org/10.1093/bioinformatics/btt476.
8. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595. https://doi.org/10.1093/bioinformatics/btp698.
9. Milne I, Bayer M, Stephen G, Cardle L, Marshall D. 2016. Tablet: visualizing next-generation sequence assemblies and mappings. Methods Mol Biol 1374:253–268. https://doi.org/10.1007/978-1-4939-3167-5_14.
10. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.