Microbiology Resource Announcements

| Bacteriology | Announcement

Whole-genome sequence of *Paenibacillus* sp., isolated from soil, Fort Collins, Colorado, USA

Paris M. Kiehl,[1] Shelby M. Cagle,[1] Traci L. Kinkel,[1] Mark D. Stenglein[1]

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT A whole genome sequence of a *Paenibacillus* sp. bacterium isolated from soil in Fort Collins, Colorado was obtained. This bacterium was of particular interest due to its antibiotic potential against Gram-positive pathogen surrogates.

KEYWORDS bacteriology, genomics, *Paenibacillus*

This isolate was collected in 2022 as part of the Small World Initiative (1). Soil at Riverbend Natural Area, Colorado, USA (40.5719 N 105.0269 W) was serially diluted in phosphate-buffered saline and plated onto potato dextrose agar (PDA). A colony was streaked to purity on tryptic soy agar (TSA). Colonies were colorless on PDA but pink on TSA. The bacterium produced zones of inhibition against Gram-positive ESKAPE surrogates in a spread patch assay (2).

A TSA colony was used to inoculate a tryptic soy broth (TSB) culture, from which DNA was isolated by phenol-chloroform extraction and ethanol precipitation. 16S ribosomal RNA gene sequence was amplified using Promega PCR Master Mix and 27F/1492R primers (3). The 16S Sanger sequence produced equally scoring BLASTN alignments with 100% identity to >100 *Paenibacillus* species (e.g., MN629116.1). MALDI-TOF mass spectrometry failed to identify the microbe.

For genome sequencing, DNA was isolated from a second TSB culture using Qiagen QiaAmp Fast DNA Tissue Kit. Manufacturer's protocols were followed unless noted. For Illumina library preparation, DNA was fragmented to an average length of ~400 nt with a Covaris M220 ultrasonicator. DNA size distribution was determined using Agilent 2200 TapeStation. Library molecules were created with Quantabio sparQ DNA kit. Library preparation included end repair and A-tailing, adapter ligation, and library amplification. A BluePippin instrument was used to isolate library molecules of 300–600 bp. Library molarity was determined using Kapa Library Quantification kit (Roche). The library was sequenced on an Illumina MiSeq using a v2 500 cycle sequencing kit to generate 1.6×10^6 2×250 paired-end reads. Nanopore libraries were prepared using the Oxford Nanopore Technologies Rapid Barcoding Kit SQK-RBK004. Libraries were sequenced on a MinION Flongle 9.4.1 flow cell, producing 20,414 reads. Guppy v.6.3.4 was used for high accuracy mode basecalling (4). The nanopore read N50 was 13,371 (5).

Adapters were trimmed with cutadapt v3.5 (6). Default parameters were generally used; full command lines are available at the github repository linked below. Read quality and absence of adapters were assessed using FASTQC v.0.11.9 (7). A hybrid assembly using SPAdes v.3.15.5 (8) produced 26 contigs and 21 scaffolds (9). The largest scaffold was circular by inspection of reads mapped across the ends. Remaining contigs were ≤531 nt long, aligned with high-percent identity to the largest scaffold and were discarded from the assembly. QUAST v.5.2.0 was used to find the contig N50 of 5,045,338 nt (10). Short and long reads were remapped to the genome using bowtie2 v.2.4.5 and minimap2 (11, 12). Samtools v.1.14 was used to determine the coverage depth: 73× for

Editor Simon Roux, DOE Joint Genome Institute, Berkeley, California, USA

Address correspondence to Mark D. Stenglein, mark.stenglein@colostate.edu.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 20 June 2023
Accepted 11 August 2023
Published 25 September 2023

Copyright © 2023 Kiehl et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Downloaded from https://journals.asm.org/journal/mra on 04 March 2024 by 129.82.94.12.

Illumina and 13× for Nanopore (13). The assembly was annotated using Bakta v.1.6.1 (14). The GC content was 40.5%, and 4,536 genes were identified.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation award IOS 2048214. Nanopore sequencing was supported by a gift from Jim McDonald.

Thanks to Dan Sloan and students in MIP/BZ-565 for contributions to library preparation.

AUTHOR AFFILIATION

[1]Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, Colorado, USA

AUTHOR ORCIDs

Mark D. Stenglein http://orcid.org/0000-0002-0993-813X

FUNDING

Funder	Grant(s)	Author(s)
National Science Foundation (NSF)	IOS 2048214	Mark D. Stenglein

AUTHOR CONTRIBUTIONS

Paris M. Kiehl, Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Shelby M. Cagle, Investigation, Methodology, Supervision, Writing – review and editing | Traci L. Kinkel, Conceptualization, Supervision, Writing – review and editing | Mark D. Stenglein, Conceptualization, Project administration, Resources, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

This Whole Genome Shotgun sequence has been deposited in NCBI under accession CP126313.1. The deposited version was reannotated by NCBI using the prokaryotic genome annotation pipeline (15). Reads are available under BioProject accession PRJNA929261 and SRA run IDs SRR23337608 and SRR23337609. Analysis code is available at https://github.com/pkiehl2002/2022_MIP_280A4_final_project.

REFERENCES

1. Barral AM, Makhluf H, Soneral P, Gasper B. 2014. Small world initiative: crowdsourcing research of new antibiotics to enhance undergraduate biology teaching (618.41). FASEB J 28:618. https://doi.org/10.1096/fasebj.28.1_supplement.618.41
2. Marcolefas E, Leung T, Okshevsky M, McKay G, Hignett E, Hamel J, Aguirre G, Blenner-Hassett O, Boyle B, Lévesque RC, Nguyen D, Gruenheid S, Whyte L. 2019. Culture-dependent bioprospecting of bacterial isolates from the Canadian high Arctic displaying antibacterial activity. Front Microbiol 10:1836. https://doi.org/10.3389/fmicb.2019.01836
3. 16S Ribosomal DNA amplification for Phylogenetic study. 1991. Available from: https://journals.asm.org/doi/epdf/10.1128/jb.173.2.697-703.1991
4. Perešíni P, Boža V, Brejová B, Vinař T. 2021. Nanopore base calling on the edge. Bioinformatics 37:4661–4667. https://doi.org/10.1093/bioinformatics/btab528
5. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. Nanopack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149
6. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200
7. Babraham Bioinformatics. n.d. FastQC a quality control tool for high throughput sequence data. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc
8. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using spades *de novo* assembler. Curr Protoc Bioinformatics 70:e102. https://doi.org/10.1002/cpbi.102
9. Li H. 2023. Seqtk: A fast and lightweight tool for processing FASTA or FASTQ sequences. https://github.com/lh3/seqtk
10. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 34:i142–i150. https://doi.org/10.1093/bioinformatics/bty266
11. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191
12. Langmead B, Wilks C, Antonescu V, Charles R. 2019. Scaling read aligners to hundreds of threads on general-purpose processors. Bioinformatics 35:421–432. https://doi.org/10.1093/bioinformatics/bty648

Downloaded from https://journals.asm.org/journal/mra on 04 March 2024 by 129.82.94.12.

13. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. Gigascience 10:giab008. https://doi.org/10.1093/gigascience/giab008
14. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom 7:000685. https://doi.org/10.1099/mgen.0.000685
15. 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

Downloaded from https://journals.asm.org/journal/mra on 04 March 2024 by 129.82.94.12.