AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Bacteriology | Announcement

Draft genome sequence of *Planococcus sp*. SK3692

Sophie M. Kiehl,[1] **Riley Anderson,**[1] **Traci L. Kinkel,**[1] **Mark D. Stenglein**[1]

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT Whole-genome sequencing was performed for a *Planococcus sp*. isolate. This bacterium was of interest because of its vibrant orange pigmentation.

KEYWORDS bacteriology, genomics, *Planococcus*

This bacterium was isolated in January 2021 as part of the Small World Initiative (1) from the brushed fur of a dog that had recently been in the Pacific Ocean at Camarillo, CA (34.2164N 119.0376W). The sample collection protocol (#5131) was reviewed by Colorado State University's Institutional Animal Care & Use Committee Waiver Subcommittee and determined to be an exempt activity. Fur was placed in phosphate-buffered saline and a dilution series was plated on tryptone soy agar (TSA). Plates were incubated for 5 days at 30°C aerobically. An orange-pigmented colony was subsequently restreaked on TSA until a pure isolate was obtained.

An overnight tryptone soy broth culture was inoculated with a single colony and grown aerobically at 30°C. DNA was extracted from the culture using the QiaAmp Fast DNA Tissue Kit (Qiagen; Valencia, USA) following the manufacturer's protocol with a lysozyme pre-incubation step. This DNA was input to Illumina and nanopore sequencing. For Illumina library preparation, DNA was fragmented using a Covaris M220 ultrasonicator to an average fragment length of ~400 bp. DNA size distributions were assessed using Agilent 2200 Tapestation. Illumina libraries were prepared using a sparQ DNA library prep kit (Quantabio; Beverly, USA) following the manufacturer's protocol. Libraries were size-selected to a range of 300–600 bp using a BluePippin instrument (Sage Science; Beverly, USA). Libraries were sequenced on an Illumina MiSeq instrument using a v2 500 cycle sequencing kit to generate 2 × 250 paired-end reads.

Cutadapt 3.5 (2) removed adapter sequences, reads shorter than 80 bases, and bases with quality score less than 30. This reduced 1,688,050 Illumina reads to 1,636,878.

Nanopore libraries were prepared using the Rapid Barcoding Kit SQK-RBK004 [Oxford Nanopore Technologies (ONT); Oxford, UK] following the manufacturer's protocol using unfragmented input DNA. Libraries were sequenced on a MinION device using one ONT Flongle R9.4.1 flow cell (FLO-FLG001). Basecalling was performed with guppy 6.3.4 (3). The 6045 nanopore reads had an average length of 6312 bp, and no reads were filtered by length or quality (4).

Long and short reads were assembled using SPAdes 3.15.5, with contigs scaffolded using long reads (5). In general, software used default parameters; full command lines are available in the github repository linked below. Assembly statistics were calculated using Quast 5.2.0 (6). The assembly contained six scaffolds longer than 400 bp (lengths 1.8 Mb, 1.5 Mb, 234 kb, 48 kb, 12 kb, 7 kb). A BLASTN 2.13.0 (7) search of scaffolds against the NCBI nucleotide database returned *Planococcus sp*. sequences with alignment identities of ~90%, with the largest three scaffolds aligning to chromosomes and the shorter three aligning to plasmid sequences. The assembly was 3.5 Mb long and had an N_{50} of 1.8 Mb and a GC content of 47.5%. The average coverage depths

Editor Leighton Pritchard, University of Strathclyde, Glasgow, United Kingdom

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 21 June 2023
Accepted 30 November 2023
Published 20 December 2023

of 108× and 10.5× for Illumina and nanopore reads, respectively, were calculated by mapping with minimap2 v2.2.4 and using samtools v1.17 (8, 9).

The GenBank assembly was annotated using PGAP 6.4 (10), which identified 3,563 genes (3,450 protein-coding). NCBI's automated taxonomic assignment verification, based on average nucleotide identity, corroborated the *Planoccocus* genus assignment (11).

ACKNOWLEDGMENTS

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

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

AUTHOR AFFILIATION

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

Sophie M. Kiehl, Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Riley Anderson, Conceptualization, Investigation, Writing – review and editing | Traci L. Kinkel, Conceptualization, Resources, Supervision, Writing – review and editing | Mark D. Stenglein, Conceptualization, Resources, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Assembly ASM2889048v2 has been deposited in DDBJ/ENA/GenBank under BioProject accession PRJNA929260, BioSample accession SAMN32954226. Raw reads have been deposited in SRA under accession SRS16684895. The accessions for the Illumina and Nanopore reads are, respectively, SRX19283240 and SRX19283241. The analysis code is available at GitHub.

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). The FASEB Journal 28:618. https://doi.org/10.1096/fasebj.28.1_supplement.618.41
2. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet j 17:10. https://doi.org/10.14806/ej.17.1.200
3. Perešíni P, Boža V, Brejová B, Vinař T. 2021. Nanopore base calling on the edge. Bioinform 37:4661–4667. https://doi.org/10.1093/bioinformatics/btab528
4. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinform 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149
5. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes *de Novo* assembler. Bioinform 70:e102. https://doi.org/10.1038/s41598-020-75855-3
6. 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
7. WheelerD, Bhagwat M. 2007. BLAST QuickStartComparative genomics. Vol. 1 and 2. Humana Press, Totowa, NJ.
8. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinform 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191
9. 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

10. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. Nucleic Acids Res 49:D1020–D1028. https://doi.org/10.1093/nar/gkaa1105

11. Ciufo S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. Int J Syst Evol Microbiol 68:2386–2392. https://doi.org/10.1099/ijsem.0.002809