



Complete Genome Sequences of Four Putatively Antibiotic-Producing Bacteria Isolated from Soil in Arkansas, USA

 Ruth Plymale,^a Griffin Hopkins,^a Taylor Johnson,^{a*} Taylor Savage,^{a§} Danielle Schaal^{a◇}

^aDepartment of Biology, Ouachita Baptist University, Arkadelphia, Arkansas, USA

ABSTRACT Soil bacteria can be a valuable source of antimicrobial compounds. Here, we report the complete genomes of four soil bacteria that were isolated by undergraduate microbiology students as part of a course-based research experience. These genomes were assembled using a hybrid approach combining paired-end Illumina reads with Oxford Nanopore Technologies MinION reads.

Soil ecosystems have historically been good sources of antibiotic-producing bacteria (1) and, although genome mining has shown promise for antibiotic discovery (2), the culture and screening of soil bacteria may still be valuable (3). These four strains were selected for their antagonistic activity against two or more ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species) pathogens (4) as part of a course-based microbiology research experience (5). Soil samples were taken at the global positioning system (GPS) coordinates indicated in Table 1. Strain isolation was achieved by diluting soil samples in 1 × phosphate-buffered saline (PBS), incubating each dilution aerobically on the isolation agar listed in Table 1 at 25°C for 5 days, and streaking individual colonies on tryptic soy agar (TSA) and growing them aerobically at 25°C for 2 days. Once these strains were selected for further study, each strain was grown on two TSA plates at 25°C for 2 days. One plate of each strain was shipped overnight to Genewiz (South Plainfield, NJ) for 16S rRNA sequencing of a colony from each plate, targeting the V1 through V9 regions; these sequences were analyzed for genus identification using NCBI BLAST with the 16S rRNA sequence database (6). The bacterial growth on the second plate for each strain was scraped into brain heart infusion broth with 25% glycerol and frozen at –80°C. Each strain was grown from frozen stock on TSA at 25°C for 2 days, and the plates were shipped overnight to the Microbial Genome Sequencing Center (MiGS) (Pittsburgh, PA), where a colony was removed from each plate for DNA extraction and sequencing. Genomic DNA was extracted using a combination of lysozyme and proteinase K enzymatic treatments before finishing with a Qiagen blood and tissue kit. The same genomic DNA isolation of each strain was used for both long- and short-read sequencing with Oxford Nanopore Technologies (ONT) and Illumina systems, respectively (7). Default parameters were used for all software. For ONT sequencing, libraries were prepared using kit SQK-LSK109 according to the manufacturer's specifications (no DNA size selection/shearing), sequencing was performed with a MinION R9 flow cell, and base calling was performed using Guppy v4.2.2 (GPU mode) (8). An Illumina Nextera kit was used with modifications, as described by Baym et al. (9), to prepare Illumina libraries, which were sequenced on a NextSeq 550 platform; bcl2fastq v2.20.0.422 was used for demultiplexing, quality control, and trimming of the Illumina paired-end reads (2 × 150 bp) (10), Porechop v0.2.4 for quality trimming and removal of adapters for ONT sequencing (11), and Unicycler v0.4.8 for hybrid *de novo* assembly, circularization, and rotation to locate *dnaA* at nucleotide 1 (12). Each chromosome or plasmid was assembled into a single contig. Sequencing and assembly details are given in Table 1.

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

Copyright © 2022 Plymale et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

Address correspondence to Ruth Plymale, plymaler@obu.edu.

*Present address: Taylor Johnson, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan, USA.

§Present address: Taylor Savage, University of Tennessee Health Science Center College of Dentistry, Memphis, Tennessee, USA.

◇Present address: Danielle Schaal, Department of Microbiology and Immunology, Louisiana State University Health Science Center, New Orleans, Louisiana, USA.

The authors declare no conflict of interest.

Received 20 July 2021

Accepted 4 December 2021

Published 6 January 2022

TABLE 1 Sequencing and assembly details for the isolated strains

Strain	GPS coordinates	Isolation agar	GenBank accession no.	Genome size (Mbp)	No. of contigs	SRA accession no. for ONT reads	No. of trimmed ONT reads	ONT read coverage (x)	N_{50} (Mbp)	SRA accession no. for Illumina reads	No. of Illumina read pairs	Illumina read coverage (x)	GC content (%)
<i>Bacillus halotolerans</i> strain MBH1	34.124260, -93.050649	TSA	CP070976	4.23	1	SRX11445848	192,595	338	4.23	SRX11445847	3,145,590	216	43.59
<i>Burkholderia</i> sp. strain LAS2	34.124167, -93.061667	TSA	CP071052, CP071053, CP071054, CP071055, CP071056	7.68	5	SRX11443407	97,395	190	4.46	SRX11443406	2,594,142	99	66.89
<i>Pseudomonas entomophila</i> strain Small	34.144506, -93.057681	Potato dextrose agar	CP070982	5.96	1	SRX11443507	194,084	220	5.96	SRX11443506	2,867,625	140	63.94
<i>Pseudomonas rhodesiae</i> strain AAMF24	39.3102, -93.1519	TSA	CP070980	5.772	1	SRX11444531	159,922	216	5.77	SRX11444530	2,987,467	151	60.49

The assembled genomes were annotated by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v5.1 (13).

Data availability. All sequences were deposited in the NCBI GenBank database under BioProject accession number [PRJNA673262](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA673262). The assembled genomes and ONT and Illumina reads are available under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

This publication was made possible by the Arkansas INBRE program, supported by a grant from the National Institute of General Medical Sciences (NIGMS) (grant P20 GM103429 from the National Institutes of Health).

We thank the Hubbard Center for Genome Studies and the Arkansas High Performance Computing Center for use of their computational resources.

REFERENCES

1. Williams ST, Vickers JC. 1986. The ecology of antibiotic production. *Microb Ecol* 12:43–52. <https://doi.org/10.1007/BF02153221>.
2. Foulston L. 2019. Genome mining and prospects for antibiotic discovery. *Curr Opin Microbiol* 51:1–8. <https://doi.org/10.1016/j.mib.2019.01.001>.
3. Baltz RH. 2008. Renaissance in antibacterial discovery from actinomycetes. *Curr Opin Pharmacol* 8:557–563. <https://doi.org/10.1016/j.coph.2008.04.008>.
4. Pendleton JN, Gorman SP, Gilmore BF. 2013. Clinical relevance of the ESKAPE pathogens. *Expert Rev Anti Infect Ther* 11:297–308. <https://doi.org/10.1586/eri.13.12>.
5. Davis E, Sloan T, Aurelius K, Barbour A, Bodey E, Clark B, Dennis C, Drown R, Fleming M, Humbert A, Glasgo E, Kerns T, Lingro K, McMillin M, Meyer A, Pope B, Stalevicz A, Steffen B, Steindl A, Williams C, Wimberley C, Zenas R, Butela K, Wildschutte H. 2017. Antibiotic discovery throughout the Small World Initiative: a molecular strategy to identify biosynthetic gene clusters involved in antagonistic activity. *Microbiologyopen* 6:e00435. <https://doi.org/10.1002/mbo3.435>.
6. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. 2008. NCBI BLAST: a better web interface. *Nucleic Acids Res* 36:W5–W9. <https://doi.org/10.1093/nar/gkn201>.
7. Spencer BL, Chatterjee A, Duerkop BA, Baker CJ, Doran KS. 2021. Complete genome sequence of neonatal clinical group B streptococcal isolate CJB111. *Microbiol Resour Announc* 10:e01268–20. <https://doi.org/10.1128/MRA.01268-20>.
8. Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol* 20:129. <https://doi.org/10.1186/s13059-019-1727-y>.
9. Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony R. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One* 10:e0128036. <https://doi.org/10.1371/journal.pone.0128036>.
10. Illumina. 2019. bcl2fastq conversion software v2.20. Illumina, San Diego, CA. https://support.illumina.com/content/dam/illumina-support/documents/documentation/software_documentation/bcl2fastq/bcl2fastq2-v2-20-software-guide-15051736-03.pdf.
11. Wick RR. 2017. Porechop. <http://github.com/rrwick/Porechop>.
12. 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>.
13. 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>.