



Draft Genome Sequences of *Pseudomonas* sp. Strains MWU12-2037 and MWU12-2345, Isolated from Peat and Sandy Bog Soils in the Cape Cod National Seashore, Massachusetts

 Ayman Anasi,^a  Scott Soby^{a,b}

^aBiomedical Sciences, College of Graduate Studies, Midwestern University, Glendale, Arizona, USA

^bCollege of Veterinary Medicine, Midwestern University, Glendale, Arizona, USA

ABSTRACT *Pseudomonas* sp. strains MWU12-2037 and MWU12-2345 were isolated from peat and sandy bog soils in wild cranberry bogs in the Cape Cod National Seashore (Massachusetts, USA) as part of a culture-dependent survey of relatively unexplored wetlands soil microbiomes. Both genomes exceeded 7 Mbp.

Despite the importance of wetlands bacteria in geochemical processes and ecosystem dynamics, their biological functions are not well understood. The genus *Pseudomonas* comprises a large number of species that together produce a long list of secondary metabolites that allow adaptation to many different environments (1, 2), but outside agricultural settings, little is known about how these bacteria influence their ecosystems. *Pseudomonas* sp. strains MWU12-2037 and MWU12-2345 were isolated from peat (42.070624 N, –70.210548 W) and sandy bog (42.064742 N, –70.117562 W) soils in wild cranberry bogs of the Cape Cod National Seashore. Soil cores were vortexed in sterile water, plated onto King's medium B (KMB) agar plates containing 50 $\mu\text{g mL}^{-1}$ each of cycloheximide and ampicillin, incubated at 26°C for 48 h, colony purified 3 \times , and stored at –80°C in 34% glycerol. The isolates were recovered from frozen storage on KMB agar; populations were inoculated into overnight KMB broth cultures for genomic DNA (gDNA) isolation using a DNeasy blood and tissue kit (Qiagen). Illumina-compatible gDNA libraries were generated from gDNA with a Kapa Biosystem Hyperplus library kit (KK8514). DNA was enzymatically sheared to \approx 500-bp fragments, end-repaired, and A-tailed. Illumina-compatible adapters with unique indexes (IDT 00989130v2) were individually ligated to each sample, cleaned using Kapa Biosciences pure beads (KK8002), and amplified with Kapa HiFi enzyme (KK2502). The libraries were analyzed for fragment size using the TapeStation system (Agilent Technologies), quantified by quantitative PCR (qPCR) using a Kapa library quantification kit (KK4835; Thermo Fisher Scientific; QuantStudio 5), and then multiplex-pooled and sequenced using an Illumina MiSeq device on a 2 \times 250-bp flow cell. The raw reads were assembled using Unicycler v0.4.8 (3) and polished using Pilon v1.23 (4) within the PATRIC (<http://patricbrc.org>) Comprehensive Genome Analysis v3.6.12 pipeline using default settings, except for the trim function, which was set to “true” for quality control and trimming using Trim Galore v0.4.0 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) (5, 6). The genome sequences were annotated using RASTtk (7) as part of the PATRIC pipeline. Using the Type (Strain) Genome Server (<https://tygs.dsmz.de/>), isolates were placed within the genus *Pseudomonas* with high confidence using the Genome BLAST Distance Phylogeny approach (8) and were found to be closely related (digital DNA-DNA hybridization [dDDH_{dd}] score = 86.6%).

MWU12-2037 had an assembled genome size of 7,222,596 bp distributed over 57 contigs, 61.34% G+C content from 1,956,500 reads, and a read length of 909,223,076 bp. The sequence coverage was 125 \times , with an N_{50} value of 422,677 bp. There were 4,934 protein-coding, 60 tRNA, and 4 rRNA genes. MWU12-2345 had an assembled genome size of 7,382,352 bp distributed over 57 contigs, 61.22% G+C content from 1,552,993 reads, and a

Editor David A. Baltrus, University of Arizona

Copyright © 2022 Anasi and Soby. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Scott Soby, ssobyx@midwestern.edu.

The authors declare no conflict of interest.

Received 26 May 2022

Accepted 23 June 2022

Published 12 July 2022

read length of 712,111,537 bp. The sequence coverage was 96 \times , with an N_{50} value of 335,205 bp. There were 5,063 protein-coding, 61 tRNA, and 5 rRNA genes.

Data availability. This whole-genome shotgun project for MWU12-2037 and MWU12-2345 has been deposited at DDBJ/ENA/GenBank under accession numbers [JALLIZ000000000](https://doi.org/10.1093/nar/gkw1017) and [JALLJA000000000](https://doi.org/10.1093/nar/gkw1017) and BioSample accession numbers [SAMN26856950](https://www.ncbi.nlm.nih.gov/biosample/SAMN26856950) and [SAMN26856949](https://www.ncbi.nlm.nih.gov/biosample/SAMN26856949), respectively. The versions described are [JALLIZ000000000.1](https://www.ncbi.nlm.nih.gov/assembly/JALLIZ000000000.1) and [JALLJA000000000.1](https://www.ncbi.nlm.nih.gov/assembly/JALLJA000000000.1), respectively. The BioProject accession number is [PRJNA691338](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA691338). The SRA accession numbers are [SRR18508470](https://www.ncbi.nlm.nih.gov/sra/SRR18508470) for MWU12-2037 and [SRR18508469](https://www.ncbi.nlm.nih.gov/sra/SRR18508469) for MWU12-2345. The RASTtk annotations are available under open license at Zenodo (<https://zenodo.org/record/6441768#.YI9mFtrMLD4> and <https://zenodo.org/record/6441878#.YI9mFNrMLD4>).

ACKNOWLEDGMENTS

This research was supported by the Office of Research and Sponsored Programs and by the Biomedical Sciences Program, College of Graduate Studies, Midwestern University. We thank Alisha Harrison for isolation of gDNA, and the Arizona State University Genomics Core Facility for library construction and Illumina sequencing. We gratefully acknowledge the generous cooperation of the Cape Cod National Seashore and the U.S. National Park Service. This work satisfies a Bacterial Genomics course requirement for A.A.

REFERENCES

1. Passarelli-Araujo H, Franco GR, Venancio TM. 2022. Network analysis of ten thousand genomes shed light on *Pseudomonas* diversity and classification. *Microbiol Res* 254:126919. <https://doi.org/10.1016/j.micres.2021.126919>.
2. Hesse C, Schulz F, Bull CT, Shaffer BT, Yan Q, Shapiro N, Hassan KA, Varghese N, Elbourne LDH, Paulsen IT, Kyrpides N, Woyke T, Loper JE. 2018. Genome-based evolutionary history of *Pseudomonas* spp. *Environ Microbiol* 20:2142–2159. <https://doi.org/10.1111/1462-2920.14130>.
3. 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>.
4. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
5. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia F, Yoo H, Stevens RL. 2017. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res* 45:D535–D542. <https://doi.org/10.1093/nar/gkw1017>.
6. Krueger F. 2015. Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (reduced representation bisulfite-seq) libraries. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
7. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5:8365. <https://doi.org/10.1038/srep08365>.
8. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10:2182. <https://doi.org/10.1038/s41467-019-10210-3>.