

Genomes of three root-associated bacteria isolated from *Sorghum bicolor* L. Moench under arid-land conditions

Ciara Garcia,¹ Duke Pauli,^{1,2} Dave Baltrus,^{1,3} A. Elizabeth Arnold^{1,2,3,4}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Diverse Proteobacteria and Actinobacteria affiliate with plants, including strains associated with plant growth and resilience. Here, we share annotated genome sequences for three root-associated bacteria isolated from *Sorghum bicolor* L. Moench in Arizona: *Sphingomonas* sp. and *Roseateles* sp. (Proteobacteria), and *Promicromonospora* sp. (Actinobacteria).

KEYWORDS agriculture, sorghum, rhizosphere, epiphytic, plant-associated bacteria

Members of the Proteobacteria and Actinobacteria comprise a large proportion of most plant microbiomes (1, 2). Sorghum, a cereal grain that can be grown in arid regions, is an emerging model crop for stress resilience. Root-associated microbial communities differ among sorghum genotypes, aligning with cultivar-specific productivity (3). Here, we present three genome sequences from root-associated bacteria isolated from sorghum in arid-land agriculture conditions. This work contributes to insights linking crop performance and genomic features of the microbiome.

Isolates were originally obtained in 2021–2022 from roots of high-performing (*Sphingomonas* sp. and *Promicromonospora* sp.) and low-performing sorghum (*Roseateles* sp.) collected at the Maricopa Agricultural Center in Arizona, USA (33°04'37"N, 111°58'26"W) (3). Isolates from this culture collection were streaked from glycerol stocks stored at –80°C (3) onto Luria broth Agar and grown at 27°C in aerobic conditions for 48–72 h. Isolated colonies were grown in 4 mL of Luria Broth for 40 h at 150 rpm to reach ca. 10⁸ CFU. Cultures were harvested entirely for DNA extraction with the Wizard Genomic DNA Isolation Kit (Promega, Wisconsin, USA; manufacturer's protocol). Optional incubation with lysozyme ensured efficient lysis for *Promicromonospora*. DNA was eluted in 40 µL of elution buffer (10 mM Tris, pH 8.5, 0.1 mM ethylenediaminetetraacetic acid [EDTA]). DNA quantity and quality were assessed with a NanoDrop 2000 (Thermo Scientific). Genomic DNA was diluted to 50 ng/µL in nuclease-free water and sequenced by Plasmidsaurus (Eugene, OR, USA) via hybrid Oxford Nanopore Technology (ONT) long-read+Illumina short-read whole-genome sequencing. An amplification-free long-read sequencing library was prepared with the ONT Rapid Barcoding Kit 96 v14 without fragmentation or size selection and sequenced on the PromethION P24 using an R10.4.1 flow cell and ont-dorado-for-promethion v0.5.0 with super-accurate as the base-calling mode (<https://github.com/nanoporetech/dorado>). The 5% worst fastq reads were filtered with Filtlong v0.2.1 (see 4). Reads were downsampled to 250 Mb and a draft genome was assembled with Flye v2.9.1 with parameters for high-quality ONT reads, then polished via Medaka v1.8.0. An additional library was prepared from the same gDNAs with an ExpressPlex 96 Library Prep Kit (SeqWell) and sequenced on an Illumina NextSeq2000 with paired-end 2 × 150 bp chemistry. Unfiltered .fastq reads were used to polish the ONT .fna assembly via Polypolish v0.6.0 (5), yielding a .fasta file reflecting hybrid assembly. One closed circular contig was produced by each assembly. Genome

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

Address correspondence to A. Elizabeth Arnold, arnold@ag.arizona.edu.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 21 April 2025

Accepted 15 August 2025

Published 12 September 2025

Copyright © 2025 Garcia et al. 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/).

TABLE 1 Assembly and genomic characteristics for three *Sorghum* root-associated bacteria^a

Species	<i>Sphingomonas</i> sp.	<i>Roseateles</i> sp.	<i>Promicromonospora</i> sp.
Strain ID	MS122	MS654	MS192
GenBank accession	CP186695	CP186694	CP186693
Genome size (Mbp)	4.11	5.79	5.99
ONT raw reads	27,211	26,008	41,736
Illumina raw reads	322,167	1,127,045	643,642
N50 (kb)	14.28	19.95	20.46
Coverage from hybrid assembly	39x	36x	52x
CDS	3,869	4,891	5,271
tRNAs	46	63	51
GC Content	67.9	68.2	72.9
Completeness (%)	98.92	99.84	100.00
Contamination (%)	1.42	0.93	0.00

^aTaxonomy was inferred from TYGS. Completeness and contamination reflect values obtained from CheckM.

completeness and contamination were assessed with CheckM v1.2.2. Unless otherwise noted, default parameters were used for all software.

The most closely related type strains identified via the Type Strain Genome Server (TYGS) v400 (6) were used to calculate average nucleotide identity (ANI) via EZ Bio Cloud's orthologous ANI tool (<https://www.ezbiocloud.net/tools/ani>). We detected 84.7% ANI between *Sphingomonas* sp. (MS122) and *Sphingomonas koreensis* NBRC16723, 86.3% ANI between *Roseateles* sp. (MS654) and *Roseateles chitinovorans* HWN-4, and 86.1% ANI between *Promicromonospora* sp. (MS192) and *Promicromonospora thailandica* DSM26652. Assembled genomes were uploaded to NCBI (Table 1) and annotated with PGAP v6.10 (7). Annotation with antiSMASH v8.0 (8) indicated that *Promicromonospora* sp. and *Roseateles* sp. have genes predicted to encode for acyl amino acids and siderophores. *Promicromonospora* sp. also has genes predicted to encode for ectoine, an extremolyte that confers resistance to salt and temperature stress (9).

ACKNOWLEDGMENTS

We thank the University of Arizona School of Plant Sciences and College of Agriculture, Life and Environmental Sciences for funding. Additional support was provided by the Robert L. Gilbertson Mycological Herbarium and the United States Department of Agriculture – National Institute of Food and Agriculture (AEA: USDA NIFA ARZT-1361340-H25-242; ARZT-1259370-S25-200). CDG was supported by a Graduate Research Fellowship from the National Science Foundation. We thank C. Dehler and M. Lee for laboratory assistance.

AUTHOR AFFILIATIONS

¹School of Plant Sciences, University of Arizona, Tucson, Arizona, USA

²Center for Agroecosystem Research in the Desert (ARID), University of Arizona, Tucson, Arizona, USA

³BIO5 Institute, University of Arizona, Tucson, Arizona, USA

⁴Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona, USA

AUTHOR ORCID*s*

Ciara Garcia  <https://orcid.org/0009-0004-4538-4519>

Duke Pauli  <http://orcid.org/0000-0002-8292-2388>

A. Elizabeth Arnold  <http://orcid.org/0000-0002-7013-4026>

FUNDING

Funder	Grant(s)	Author(s)
National Institute of Food and Agriculture	ARZT-1361340-H25-242, ARZT-1259370-S25-200	A. Elizabeth Arnold

AUTHOR CONTRIBUTIONS

A. Elizabeth Arnold, Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review and editing.

DATA AVAILABILITY

Assembled genomes are listed in Table 1. Raw reads have been deposited in the Sequence Read Archive under the BioProject accession [PRJNA1252778](https://doi.org/10.1093/bioinformatics/btad125) under accessions [SRX28466841](https://doi.org/10.1093/bioinformatics/btad125), [SRX28466843](https://doi.org/10.1093/bioinformatics/btad125), and [SRX28466845](https://doi.org/10.1093/bioinformatics/btad125) for ONT reads, and under accessions [SRX28466842](https://doi.org/10.1093/bioinformatics/btad125), [SRX28466844](https://doi.org/10.1093/bioinformatics/btad125), and [SRX28466846](https://doi.org/10.1093/bioinformatics/btad125) for Illumina reads.

REFERENCES

- Schlaeppli K, Bulgarelli D. 2015. The plant microbiome at work. *Mol Plant Microbe Interact* 28:212–217. <https://doi.org/10.1094/MPMI-10-14-0334-FI>
- Compant S, Samad A, Faist H, Sessitsch A. 2019. A review on the plant microbiome: ecology, functions, and emerging trends in microbial application. *J Adv Res* 19:29–37. <https://doi.org/10.1016/j.jare.2019.03.004>
- Garcia C, Pauli D, Plecki C, Alnasser H, Rozzi B, Calleja S, Arnold AE. 2024. The root endophytic microbiome shifts under drought in high-performing sorghum. *Phytobiomes Journal* 8:282–296. <https://doi.org/10.1094/PBIOMES-09-23-0095-R>
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>
- Wick RR, Holt KE. 2022. Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 18:e1009802. <https://doi.org/10.1371/journal.pcbi.1009802>
- 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>
- 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>
- Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJN, van Wezel GP, Medema MH, Weber T. 2023. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* 51:W46–W50. <https://doi.org/10.1093/nar/gkad344>
- Lentzen G, Schwarz T. 2006. Extremolytes: natural compounds from extremophiles for versatile applications. *Appl Microbiol Biotechnol* 72:623–634. <https://doi.org/10.1007/s00253-006-0553-9>