

Complete genome sequences of five *Variovorax* strains isolated from the *Populus* rhizosphere and endosphere

Delaney G. Beals,¹ Alyssa A. Carrell,¹ Brisa Rodriguez,¹ William G. Alexander,¹ Christopher W. Schadt,¹ Leah H. Hochanadel,¹ Mircea Podar,¹ Dale A. Pelletier¹

AUTHOR AFFILIATION See affiliation list on p. 3.

ABSTRACT We report the complete genome sequences of five *Variovorax* strains isolated from the roots of *Populus* species in greenhouse and field settings. Genome sizes range from 7.44 to 7.71 Mbp and encode 6,077–6,879 predicted protein-coding genes. These genome assemblies expand the known genomic catalog of *Variovorax* in plant-associated environments.

KEYWORDS plant-microbe interface, endosphere, rhizosphere, poplar, *Variovorax*

The bacterial genus *Variovorax* (Burkholderiales, *Comamonadaceae*) comprises metabolically versatile bacteria inhabiting soils and plant roots (1–6). We report complete genome sequences of strains BK018, GV025, GV048, YR566, and WT075, isolated from *Populus*-associated habitats to support ecological and functional diversity studies within this genus.

Strains were isolated from *Populus* root-associated compartments using dilution plating on R2A agar incubated at 28°C for 2 days. BK018 originated from a field-grown *P. deltoides* tree in Bellville, GA (32.1348° N, 81.9657° W), and YR566 from a *P. deltoides* tree in Jerusalem, NC (35.8381° N, 80.4856° W). WT075 was recovered from the rhizosphere soil of a *P. trichocarpa* tree in Naches, WA (46.7024° N, 120.6601° W). GV025 and GV048 were obtained from surface-sterilized roots of *P. trichocarpa* plants grown in a greenhouse (Oak Ridge National Laboratory), inoculated with soil collected from a *P. trichocarpa* common garden in Corvallis, OR (44.5880° N, 123.1935° W). Taxonomic identification preceded genome sequencing by 16S rRNA gene amplification (primers 27F/1492R), Sanger sequencing (Eurofins Genomics, Louisville, KY, USA), and BLASTn comparison against the NCBI rRNA database, showing ≥99% identity to *Variovorax* species.

Frozen stocks were revived on R2A agar, and single colonies were grown in R2A broth at 28°C overnight for DNA extraction. Genomic DNA was isolated using the DNeasy Blood and Tissue Kit (Qiagen) and eluted in 0.1× TE buffer. Libraries for WT075 were prepared in-house using the Oxford Nanopore Technologies (ONT) Ligation Sequencing Kit without shearing or size selection and sequenced on an R10.4.1 MinION flow cell. Libraries for the remaining strains were prepared by Plasmidsaurus (Louisville, KY, USA) using the ONT Ligation Sequencing Kit and sequenced on GridION R10.4.1 flow cells with V14 chemistry. High molecular weight DNA was tagmented without additional shearing or size selection, following the amplification-free ligation protocol. Base calling for WT075 used Dorado v0.7.1 (ONT) with the high-accuracy calling model; for the remaining strains, Guppy v6.4.6 (ONT) was used. Each strain yielded 87k–104k reads (N50: 15.2–18.6 kb). Reads filtered with Filtrlong v0.2.1 (minimum Q ≥ 10, top 95% retained) were assembled with Tricycler v0.5.4 (7) using Flye v2.9.4 (8), Raven v1.5.3 (9), and Miniasm v0.3 (10, 11) and polished with Minipolish v0.1.3 (12) and Medaka v2.0.1 (ONT). Default parameters were used unless noted.

Editor Irene L. G. Newton, Indiana University, Bloomington, Bloomington, Indiana, USA

Address correspondence to Dale A. Pelletier, pelletierda@ornl.gov.

The authors declare no conflict of interest.

Received 2 September 2025

Accepted 13 November 2025

Published 8 December 2025

Copyright © 2025 Beals 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/).

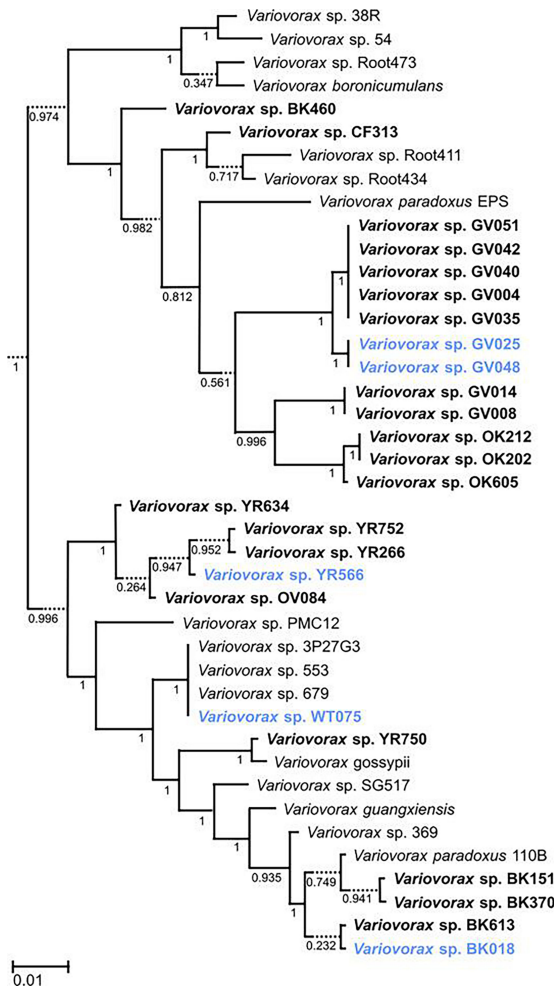


FIG 1 Maximum-likelihood phylogenetic tree based on 49 core universal bacterial proteins computed using SpeciesTree v2.2.0 in KBase. The five newly sequenced strains are shown in blue. *Variovorax* spp. strains previously isolated from *Populus* trees are shown in bold. Branch support values and scale bar indicate substitutions per site.

TABLE 1 Genome statistics for newly sequenced *Variovorax* strains

Strain	Genome size (bp)	GC (%)	N50	Predicted genes	Completeness (%)/contamination (%)	Contig 1 length (bp)	Contig 1 coverage	Contig 2 length (bp)	Contig 2 coverage
BK018	7,593,393	67.7	6,331,524	6,618 (6,611 unique)	100/0.47	6,331,524	65.6	1,261,869	63.8
GV025	7,489,868	69.3	5,711,485	6,085 (6,077 unique)	100/0.53	5,711,485	62.8	1,778,383	72
GV048	7,489,859	69.3	5,711,476	6,085 (6,078 unique)	100/0.53	5,711,476	66.5	1,778,383	66.3
WT075	7,447,132	67.1	7,447,132	6,517 (6,470 unique)	100/0.62	7,447,132	40.7		
YR566	7,705,554	66.1	7,249,798	6,879 (6,852 unique)	100/0.47	7,249,798	65.1	455,756	56.5

WT075 was assembled into a single circular chromosome, while all others each yielded two circular contigs. Sequencing coverage ranged from 40.7× to 72.0× (Table 1). Genome sizes ranged from 7.44 to 7.71 Mbp with GC contents between 66.1% and 69.3%. Prokka v1.14.6 (13) predicted 6,077–6,879 protein-coding genes per genome (14). Circularity was confirmed by identifying overlapping ends and trimming with Tricycler; contigs were not rotated to a specific gene.

Taxonomic classification (GTDB-Tk v2.3.2 [(15)], ANI comparisons (FastANI v0.1.3 [16]), and phylogenetic placement (SpeciesTree v2.2.0 [17], Fig. 1) were performed in KBase (18). BK018 was identified as *Variovorax paradoxus* A and WT075 as *Variovorax* sp003952185 (ANI ≥98.6%). GV025 and GV048 shared 99.99% ANI and grouped with *Variovorax* sp. OK212 (91.6% ANI), while YR566 was classified as *Variovorax* sp900106655 (95.8% ANI).

ACKNOWLEDGMENTS

This research was sponsored by the Genomic Science Program, U.S. Department of Energy, Office of Science, Biological and Environmental Research, as part of the Plant Microbe Interfaces Scientific Focus Area (<http://pmi.ornl.gov>). Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

AUTHOR AFFILIATION

¹Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

AUTHOR ORCIDs

Delaney G. Beals  <http://orcid.org/0000-0003-0627-4574>

Alyssa A. Carrell  <http://orcid.org/0000-0003-1142-4709>

Christopher W. Schadt  <https://orcid.org/0000-0001-8759-2448>

Mircea Podar  <https://orcid.org/0000-0003-2776-0205>

Dale A. Pelletier  <http://orcid.org/0000-0002-4321-7918>

DATA AVAILABILITY

Unprocessed sequencing reads have been deposited in GenBank under BioProject accession number [PRJNA1255178](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1255178). Individual genome assemblies have been deposited to GenBank under the accession numbers [CP197088](https://www.ncbi.nlm.nih.gov/assembly/CP197088)–[CP197089](https://www.ncbi.nlm.nih.gov/assembly/CP197089), [CP197086](https://www.ncbi.nlm.nih.gov/assembly/CP197086)–[CP197087](https://www.ncbi.nlm.nih.gov/assembly/CP197087), [CP197084](https://www.ncbi.nlm.nih.gov/assembly/CP197084)–[CP197085](https://www.ncbi.nlm.nih.gov/assembly/CP197085), [CP197083](https://www.ncbi.nlm.nih.gov/assembly/CP197083), and [CP197092](https://www.ncbi.nlm.nih.gov/assembly/CP197092)–[CP197093](https://www.ncbi.nlm.nih.gov/assembly/CP197093). A KBase Narrative containing all sequence data and genome analysis steps is also publicly accessible (14).

REFERENCES

- Carper DL, Weston DJ, Barde A, Timm CM, Lu T-Y, Burdick LH, Jawdy SS, Klingeman DM, Robeson MS, Veach AM, Cregger MA, Kalluri UC, Schadt CW, Podar M, Doktycz MJ, Pelletier DA. 2021. Cultivating the bacterial microbiota of *Populus* roots *mSystems* 6:e01306–20. <https://doi.org/10.1128/mSystems.01306-20>
- Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, Rauf P, Huettel B, Reinhardt R, Schmelzer E, Peplies J, Gloeckner FO, Amann R, Eickhorst T, Schulze-Lefert P. 2012. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* 488:91–95. <https://doi.org/10.1038/nature11336>
- Taye ZM, Helgason BL, Bell JK, Norris CE, Vail S, Robinson SJ, Parkin IAP, Arcand M, Mamet S, Links MG, Dowhy T, Siciliano S, Lamb EG. 2019. Core and differentially abundant bacterial taxa in the rhizosphere of field grown *Brassica napus* genotypes: implications for canola breeding. *Front Microbiol* 10:3007. <https://doi.org/10.3389/fmicb.2019.03007>
- Finkel OM, Salas-González I, Castrillo G, Conway JM, Law TF, Teixeira P, Wilson ED, Fitzpatrick CR, Jones CD, Dangel JL. 2020. A single bacterial genus maintains root growth in a complex microbiome. *Nature* 587:103–108. <https://doi.org/10.1038/s41586-020-2778-7>
- Lerner H, Öztürk B, Dohrmann AB, Thomas J, Marchal K, De Mot R, Dehaen W, Tebbe CC, Springael D. 2021. DNA-SIP and repeated isolation corroborate *Variovorax* as a key organism in maintaining the genetic memory for linuron biodegradation in an agricultural soil. *FEMS Microbiol Ecol* 97:fiab051. <https://doi.org/10.1093/femsec/fiab051>
- HaedarJR, Yoshimura A. 2024. New variochelins from soil-isolated *Variovorax* sp. H002. *Beilstein J Org Chem* 20:692–700. <https://doi.org/10.3762/bjoc.20.63>
- Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, Wyras KL, Holt KE. 2021. Tricycler: consensus long-read assemblies for bacterial genomes. *Genome Biol* 22:266. <https://doi.org/10.1186/s13059-021-02483-z>
- 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>
- Vaser R, Šikić M. 2021. Time- and memory-efficient genome assembly with Raven. *Nat Comput Sci* 1:332–336. <https://doi.org/10.1038/s43588-021-00073-4>

10. Li H. 2016. Minimap and miniasm: fast mapping and *de novo* assembly for noisy long sequences. *Bioinformatics* 32:2103–2110. <https://doi.org/10.1093/bioinformatics/btw152>
11. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
12. Wick RR, Holt KE. 2021. Benchmarking of long-read assemblers for prokaryote whole genome sequencing. *F1000Res* 8:2138. <https://doi.org/10.12688/f1000research.21782.4>
13. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
14. Beals DG, Pelletier DA. 2025. Complete genome sequences of five *Variovorax* strains isolated from the *Populus* rhizosphere and endosphere. DOE Systems Biology Knowledgebase. <https://doi.org/10.25982/206210.127/2566193>
15. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>
16. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>
17. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>
18. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, et al. 2018. KBase: the United States department of energy systems biology knowledgebase. *Nat Biotechnol* 36:566–569. <https://doi.org/10.1038/nbt.4163>