

High-quality draft genome sequences of seven *Ralstonia* spp. isolated from temperate forest soils

A. Li Han Chan,¹ Mallory J. Choudoir,² Achala Narayanan,³ Damayanti Rodriguez-Ramos,⁴ Grace Pold,⁵ Andrew F. Billings,¹ Kristen M. DeAngelis¹

AUTHOR AFFILIATIONS See affiliation list on p. 4.

ABSTRACT We report seven highquality draft genomes of *Ralstonia* spp. isolated from the Harvard Forest Long-Term Warming Experimental plots: four *de novo* hybrid assemblies and three *de novo* long-read assemblies. The genomes have a minimum estimated completeness of 92.6% and an average GC content of 63.45%.

KEYWORDS *Ralstonia*, forest soils

Ralstonia are free-living soil bacteria known for pathogenicity (1) and terrestrial carbon cycling (2). To better understand *Ralstonia* and how they respond to climate stress, we sequenced and annotated seven *Ralstonia* isolated from experimentally heated and control soils in Prospect Hill at the Harvard Forest Long-Term Warming Experiment in Petersham, MA, USA (42.54°N, 72.18°W) (3). Soils were collected in April and June 2014 with 1/2-inch tubular soil corers to a depth of 10 cm, with the organic and mineral horizons split by eye (4). The respective isolation media and atmosphere are described in Table 1. For whole-genome sequencing, isolates were streaked from glycerol stocks onto 10% tryptic soy agar or Reasoner's 2 agar (pH 6). Single colonies were picked to grow in the same liquid media as the agar on a shaking incubator at 30°C until the stationary phase. Genomic DNA extraction was performed on the pellets using the methods listed in Table 1. An additional RNase A treatment was performed for GP95 according to reference 5. We validated the average DNA fragment size to be 30–50 kb on 0.5% agarose gel with Quick-Load 1 kb Extend DNA ladder (New England Biolabs, Ipswich, MA, USA). For long-read sequencing using Oxford Nanopore Technologies (Oxford, UK), we followed the ligation sequencing library protocol compatible with the consumables listed in Table 1. No size selection or shearing was performed prior to library preparation. Six to seven isolates were multiplexed for a single sequencing run for 48–72 hours, with barcoded allocation turned on in the MinKNOW interface (Oxford Nanopore Technologies), except GP95, which was sequenced on a single flow cell. Fast5 or Pod5 files were base called using the high accuracy model with base callers listed in Table 1. The reads were subsampled to a minimum of 40× target coverage and controlled for quality and length using “–min_length 1000 –min_mean_q 85” on Filtlong version 0.2.1 (6). We used Flye version 2.9.1 (7) to generate the *de novo* assembly, Racon version 1.4.3 (8) and Minimap2 version 2.24 (9) to generate the consensus sequence, and Medaka version 1.7.2 (Oxford Nanopore Technologies) to generate the final polished genome draft sequence.

For *de novo* hybrid assembled isolates, short read libraries were constructed using Illumina (San Diego, CA, USA) DNA Prep kit and IDT (Coralville, IA, USA) 10 bp UDI indices and sequenced on NextSeq 2000 (Illumina) to produce 2 × 151 bp reads. Demultiplexing, quality control, and adapter trimming were performed with bcl-convert version 3.9.3 (Illumina). *De novo* hybrid assemblies were generated using Unicycler version 0.5.0 (14). Default parameters were used except where otherwise noted.

Editor Simon Roux, DOE Joint Genome Institute, Berkeley, California, USA

Address correspondence to Kristen M. DeAngelis, , deangelis@umass.edu.

The authors declare no conflict of interest.

See the funding table on p. 4.

Received 14 January 2025

Accepted 21 March 2025

Published 14 April 2025

Copyright © 2025 Chan 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 Provenance with relevant citations and genome assembly statistics of each *Ralstonia* spp. isolate

<i>Ralstonia</i> spp. isolate name	AB22-23	AB28-3	AB36-13C	GP103	GP71	GP101	GP95
Soil horizon	Mineral	Mineral	Mineral	Organic	Organic	Organic	Organic
Warming treatment	Control	Warm	Warm	Warm	Warm	Warm	Control
Isolation media	1% nutrient broth (10)	1% nutrient broth (10)	1% nutrient broth (10)	VL55 plant polymer (11)	VL55 plant polymer (11)	VL55 plant polymer (11)	VL55 plant polymer (11)
Isolation atmosphere	5% hydrogen, 5% CO ₂ , and 90% nitrogen	Aerobic	Aerobic-COY fluctuating (4)	Aerobic	Aerobic	Aerobic	Aerobic
DNA extraction method	Qiagen DNeasy Blood & Tissue Kit (Germantown, MD, USA)	Qiagen DNeasy Blood & Tissue Kit (Germantown, MD, USA)	Qiagen DNeasy Blood & Tissue Kit (Germantown, MD, USA)	Qiagen DNeasy Blood & Tissue Kit (Germantown, MD, USA)	Qiagen DNeasy Blood & Tissue Kit (Germantown, MD, USA)	Qiagen DNeasy Blood & Tissue Kit (Germantown, MD, USA)	Modified CTAB-lysozyme extraction protocol (12, 13)
Sequencing facility	SeqCenter, Pittsburgh, PA, USA, and inhouse	SeqCenter, Pittsburgh, PA, USA, and inhouse	SeqCenter, Pittsburgh, PA, USA, and inhouse	SeqCenter, Pittsburgh, PA, USA, and inhouse	Inhouse ONT	Inhouse ONT	Inhouse ONT
Sequencing technology	Illumina NextSeq 2000 and ONT	Illumina NextSeq 2000 and ONT	Illumina NextSeq 2000 and ONT	Illumina NextSeq 2000 and ONT	ONT	ONT	ONT
ONT Flow Cell	FLO-MIN106	FLO-MIN106	FLO-MIN106	FLO-MIN106	FLO-MIN106	FLO-MIN106	FLO-MIN114
ONT Library	SQK-LSK109, EXP-NBD104	SQK-LSK109, EXP-NBD104	SQK-LSK109, EXP-NBD104	SQK-LSK109, EXP-NBD104	SQK-LSK109, EXP-NBD104	SQK-LSK109, EXP-NBD104	SQK-LSK114
Preparation Kit	Guppy version 6.5.7	Guppy version 6.5.7	Guppy version 6.5.7	Guppy version 6.5.7	Guppy version 6.5.7	Guppy version 6.5.7	Dorado version 0.7.3
Base caller	SAMN33771254	SAMN33771256	SAMN33771258	SAMN33771263	SAMN33771268	SAMN33771262	SAMN44697474
BioSample	SRX19837939	SRX19837949	SRX24035310	SRX19837956	SRX19837945	SRX19837955	SRX26711341
SRA accession no. (ONT)	SRX26601167	SRX26601168	SRX26601169	SRX26601170	N/A ^a	N/A ^a	N/A ^a
SRA accession no. (Illumina)	SRX26601167	SRX26601168	SRX26601169	SRX26601170	N/A ^a	N/A ^a	N/A ^a
WGS accession no.	JBKIK0000000000	JBKIJ0000000000	JBKII0000000000	JBKIG0000000000	JBKIF0000000000	JBKIH0000000000	JBKIE0000000000
Assembly year	2022	2022	2022	2022	2022	2021	2024
Assembly type	de novo hybrid	de novo hybrid	de novo hybrid	de novo hybrid	de novo long read	de novo long read	de novo long read
Reads Illumina (read pairs)	3,625,687	2,957,812	3,160,759	3,519,480	N/A	N/A	N/A
Number of raw ONT reads (bp)	149,100,000	220,000,000	246,700,000	220,000,000	2,100,000,000	5,645,342,330	6,011,438,671
Sequencing N ₅₀ for ONT (bp)	12,636	17,948	10,557	24,364	43,073	16,001	4,038
Filtered reads ONT (bp)	149,072,979	488,529,099	246,663,274	554,542,900	240,016,725	216,041,805	236,625,424
Genome size (bp)	5,298,867	5,483,600	5,513,759	5,515,304	5,432,770	5,513,656	5,346,046

(Continued on next page)

TABLE 1 Provenance with relevant citations and genome assembly statistics of each *Ralstonia* spp. isolate (Continued)

<i>Ralstonia</i> spp. isolate name	AB22-23	AB28-3	AB36-13C	GP103	GP71	GP101	GP95
Coverage	184.9	151.1	160.9	179.0	44.2	39.2	44.3
Assembly N ₅₀ values (bp)	1,386,635	3,507,490	3,537,649	3,537,649	3,537,633	3,537,632	3,533,349
GC content (%)	63.57	63.61	63.58	63.58	62.63	63.58	63.63
No. of contigs	7	6	5	6	4	5	6
Contamination (%)	0.47	0.47	0.47	0.47	0.47	0.47	0
Completeness (%)	92.6	99.94	99.94	99.94	99.06	98.92	99.94
Most closely related genome	<i>Ralstonia thomasii</i> (99.72%)	<i>Ralstonia thomasii</i> (99.73%)	<i>Ralstonia thomasii</i> (99.73%)	<i>Ralstonia thomasii</i> (99.73%)	<i>Ralstonia thomasii</i> (99.79%)	<i>Ralstonia thomasii</i> (99.72%)	<i>Ralstonia thomasii</i> (99.73%)
according to whole genome sequence in NCBI (average nucleotide identity)							
23S rRNA genes count	21	21	21	21	20	21	21
16S rRNA genes count	17	19	19	19	19	16	19
5S rRNA genes count	3	3	3	3	3	3	3
trRNA count	64	78	78	78	65	86	78

^aN/A, not applicable as these strains were assembled with ONT reads only.

Genomes were identified as *Ralstonia* with average nucleotide identity by NCBI (15). The genome sizes for these isolates range from 5,298,867 to 5,515,304 bp, with GC content ranging from 62.63% to 63.63%. Genome contamination and completeness were estimated with CheckM version 1.0.18 (16). Genomes were not circularized and were annotated with RASTtk version 1.073 (17).

ACKNOWLEDGMENTS

We are grateful to Drs. Serita Frey and Mel Knorr for maintaining the infrastructure at Harvard Forest.

This project was conducted with support from the National Science Foundation under the grant IDs DEB-1456528, DEB-1749206, and DEB-1832210.

AUTHOR AFFILIATIONS

¹Department of Microbiology, University of Massachusetts Amherst, Amherst, Massachusetts, USA

²Department of Plant and Microbial Biology, North Carolina State University, Raleigh, North Carolina, USA

³Department of Plant and Microbial Biology, University of Minnesota, St. Paul, Minnesota, USA

⁴Department of Bacteriology, University of Wisconsin Madison, Madison, Wisconsin, USA

⁵Department of Soil and Environment, Swedish University of Agricultural Sciences, Uppsala, Sweden

AUTHOR ORCIDs

A. Li Han Chan  <http://orcid.org/0000-0002-7775-6207>

Mallory J. Choudoir  <http://orcid.org/0000-0002-9117-5150>

Achala Narayanan  <http://orcid.org/0000-0001-9574-6822>

Grace Pold  <http://orcid.org/0000-0002-7536-3944>

Kristen M. DeAngelis  <http://orcid.org/0000-0002-5585-4551>

FUNDING

Funder	Grant(s)	Author(s)
National Science Foundation	DEB-1749206, DEB-1456528, DEB-1832210	A. Li Han Chan Mallory J. Choudoir Achala Narayanan Damayanti Rodriguez-Ramos Grace Pold Andrew F. Billings Kristen M. DeAngelis

DATA AVAILABILITY

This project has been deposited in GenBank under the accession number [PRJNA944978](https://doi.org/10.26434/chemrxiv-2024-1007). The respective BioSample numbers, SRA accession numbers for Oxford Nanopore Technologies and Illumina sequencing (where applicable), and WGS accession numbers are listed in Table 1.

REFERENCES

- Ryan MP, Adley CC. 2014. *Ralstonia* spp.: emerging global opportunistic pathogens. *Eur J Clin Microbiol Infect Dis* 33:291–304. <https://doi.org/10.1007/s10096-013-1975-9>
- DeAngelis KM, Pold G, Topçuoğlu BD, van Diepen LTA, Varney RM, Blanchard JL, Melillo J, Frey SD. 2015. Long-term forest soil warming alters microbial communities in temperate forest soils. *Front Microbiol* 6:104. <https://doi.org/10.3389/fmicb.2015.00104>
- Peterjohn WT, Melillo JM, Steudler PA, Newkirk KM, Bowles FP, Aber JD. 1994. Responses of trace gas fluxes and N availability to experimentally

- elevated soil temperatures. *Ecol Appl* 4:617–625. <https://doi.org/10.2307/1941962>
4. Pold G, Billings AF, Blanchard JL, Burkhardt DB, Frey SD, Melillo JM, Schnabel J, van Diepen LTA, DeAngelis KM. 2016. Long-term warming alters carbohydrate degradation potential in temperate forest soils. *Appl Environ Microbiol* 82:6518–6530. <https://doi.org/10.1128/AEM.02012-16>
 5. Yoshinaga Y, Dalin E. 2016. RNase a cleanup of DNA samples. In Joint genome institute department of energy
 6. Wick R. 2021. Filtlong: quality filtering tool for long reads. Github. Available from: <https://github.com/rwick/Filtlong>. Retrieved 11 Mar 2025.
 7. 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>
 8. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 27:737–746. <https://doi.org/10.1101/gr.214270.116>
 9. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
 10. Janssen PH, Yates PS, Grinton BE, Taylor PM, Sait M. 2002. Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions *Acidobacteria*, *Actinobacteria*, *Proteobacteria*, and *Verrucomicrobia*. *Appl Environ Microbiol* 68:2391–2396. <https://doi.org/10.1128/AEM.68.5.2391-2396.2002>
 11. Eichorst SA, Kuske CR, Schmidt TM. 2011. Influence of plant polymers on the distribution and cultivation of bacteria in the phylum *Acidobacteria*. *Appl Environ Microbiol* 77:586–596. <https://doi.org/10.1128/AEM.01080-10>
 12. Larsen MH, Biermann K, Tandberg S, Hsu T, Jacobs WR Jr. 2007. Genetic manipulation of *Mycobacterium tuberculosis*. *Curr Protoc Microbiol* Chapter 10:Unit 10A.2. <https://doi.org/10.1002/9780471729259.mc10a02s6>
 13. Choudoir MJ, Narayanan A, Rodriguez-Ramos D, Simoes R, Efroni A, Sondrini A, DeAngelis KM. 2023. Genetic manipulation of *Mycobacterium tuberculosis*. *Curr Protoc Microbiol* 6:10A–2. <https://doi.org/10.1101/2023.03.16.532972>
 14. 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>
 15. Ciufu S, Kannan S, Sharma S, Badretin 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>
 16. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>
 17. 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>