

Complete genome sequences of four *Ochrobactrum pituitosum* strains and three *Pseudochrobactrum saccharolyticum* strains isolated from *Caenorhabditis elegans* gut microbiomes

Sandhya Lakshmi Narayanan,¹ Carsten Fortmann-Grote,¹ Paul B. Rainey^{1,2}

AUTHOR AFFILIATIONS See affiliation list on p. 6.

ABSTRACT We report complete genomes of four strains of *Ochrobactrum pituitosum* and three strains of *Pseudochrobactrum saccharolyticum* isolated from gut microbiomes of *Caenorhabditis elegans* colonized with microbial communities extracted from garden compost. Genomes were assembled from Oxford Nanopore long read data and polished with Illumina short read data.

KEYWORDS *Ochrobactrum*, *Pseudochrobactrum*, assembly, annotation, environmental sample, *C. elegans*, compost

We announce closed genome assemblies and annotations of four *Ochrobactrum pituitosum* strains (NCBI: txid571256) and three *Pseudochrobactrum saccharolyticum* strains (NCBI: txid354352). Both taxa are part of the Family Brucellaceae (NCBI: txid118882).

Strains were isolated from the gut microbiome of *Caenorhabditis elegans* strain N2 (Caenorhabditis Genomic Center) (1). Three microbial communities were extracted from garden compost, frozen at -80°C , re-thawed, and acclimatized on plates (M9 buffer, 1.5% agar) containing 4 cm² cellulose paper at 20°C. After 2 weeks, community-degraded paper was suspended in buffer, and 100 μL of the slurry was transferred to an identical piece of paper on fresh agar. Approximately 100 nematodes were added to the paper, with nematodes feeding leading to the colonization of the gut by bacteria.

After three 2-week transfers, nematodes were collected from the paper, surface-sterilized, and pulverized for 5 min using a tissue lyser (TissueLyser II, Qiagen) for microbiome extraction. Extracted samples were serially diluted and plated on TSA agar plates and incubated at 30°C for 24–48 h. Single colonies were purified by re-streaking colonies on TSA plates and stored in 70% glycerol saline aliquots at -80°C . Samples were thawed and re-grown (overnight) in shaking 6 mL LB broth at 28°C. Genomic DNA was extracted following the method described in a previous study (2) and purified using the Genomic DNA Clean & Concentrate kit–10 (Zymo Research). DNA was quantified using the Invitrogen Qubit dsDNA Broad Range (BR) assay kit.

Long-read sequencing (Oxford Nanopore Technology) was performed by Eurofins (Germany). Short-read sequencing was performed on a NextSeq500 platform (MPI for Evolutionary Biology) at 150 bp read length, paired-end sequencing, and 60-fold coverage. In-house library preparation used a Tn5 transposase enzyme library (3). Short-read raw data were quality-controlled using MultiQC v1.11 (4). Here, and in the following, default parameters were applied unless otherwise noted.

Raw nanopore sequencing (5) was filtered for low quality and short reads using Filtrlong v0.2.1 (6) before being assembled using Flye v2.9.3 (7); parameters were optimized for bacterial genomes (7). Contigs were polished using Medaka v1.8 (8). Assembled genomes were annotated using Bakta v1.8.2 (9). Quality of the assembled

Editor Zhenjiang Zech Xu, Nanchang University, Nanchang, Jiangxi, China

Address correspondence to Carsten Fortmann-Grote, grotec@evolbio.mpg.de.

The authors declare no conflict of interest.

Received 11 September 2025

Accepted 14 February 2026

Published 21 April 2026

Copyright © 2026 Narayanan 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/).

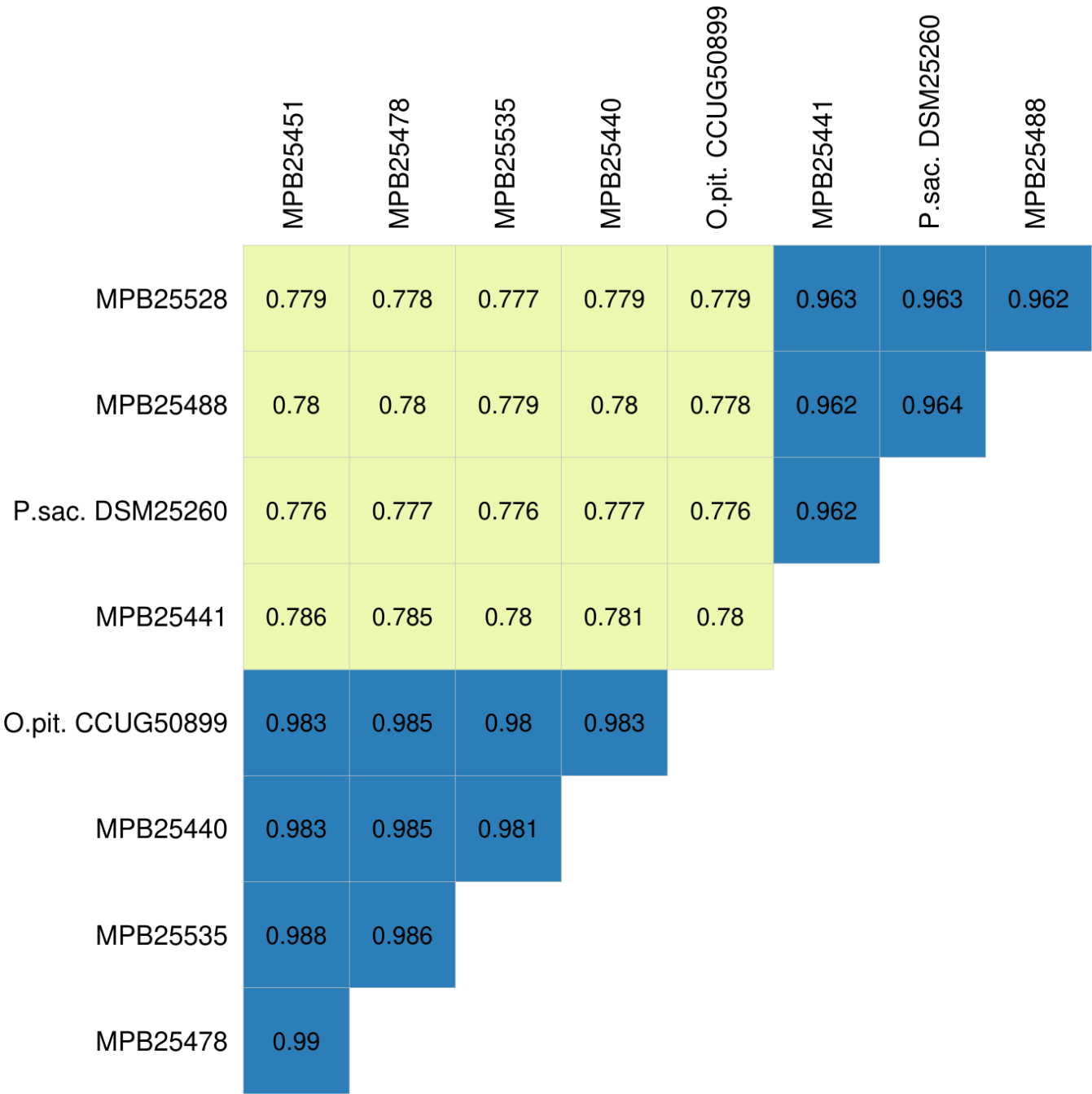


FIG 1 Clustered heatmap of ANI similarity values for seven announced strains and type strains *P. saccharolyticum* DSM25260 and *O. pituitosum* CCUG50899.

genomes was assessed using QUAST v5.2 (10), CheckM2 v1.0.1 (11), and Mash v2.3 (12). As a final QC step, sequence-cleaned reads were mapped onto the assembly using minimap2 v2.24 (13). Finally, long-read assemblies were polished with our Illumina short reads using nextpolish v1.4.1 (14) and postprocessed with samtools v1.2.1 (15).

Assembled genomes were circularized with DNAPler (16) (default parameters) such that the start site coincides with the *dnaA* gene start codon. Average genome coverage was 84.0x with a GC content of 57.8%.

Genera were identified by 16S metabarcode sequencing (using primers 341F (CCTACGGGNGGCWGCAG) and 805R (GACTACHVGGGTATCTAATCC) to amplify the hypervariable V3-V4 region) and analyzed using phylos v1.52.0 (17). Isolates were

TABLE 1 Summary table

Sample	Property	Value
SAMEA118333800	Organism	<i>O. pituitosum</i>
	Taxon ID	71256
	Isolate/strain	MPB25535
	Short read exp	ERX14406277
	Short read run	ERR15002230
	Short read count	1,727,681
	Short read GC%	53.3
	Short read mean length (RI, R2)	150,150
	Long read exp	ERX14512245
	Long read run	ERR15107547
	Long read count	87,517
	Long read mean length	6113
	Long read GC%	55.0
	Assembly	GCA_965577835
	Total size	4,963,799
	Contigs	3
	GC%	53.4
	N50	13,839,640
	ANI (<i>O. pituitosum</i> CCUG50899)	98.06
	Annotation	ERZ26904280
	Annotated CDSs	4662
SAMEA118333801	Organism	<i>O. pituitosum</i>
	Taxon ID	71256
	Isolate/Strain	MPB25451
	Short read exp	ERX14406278
	Short read run	ERR15002231
	Short read count	1,790,809
	Short read mean length (R1, R2)	150,150
	Short read GC%	53.5
	Long read exp	ERX14512247
	Long read run	ERR15107549
	Long read count	50,100
	Long read mean length	4595
	Long read GC%	51.8
	Assembly	GCA_965579405
	Total size	5,097,185
	Contigs	3
	GC%	53.5
	N50	6,950,440
	ANI (<i>O. pituitosum</i> CCUG50899)	98.27
	Annotation	ERZ26904286
	Annotated CDSs	4825
SAMEA118333802	Organism	<i>O. pituitosum</i>
	Taxon ID	71256
	Isolate/Strain	MPB25440
	Short read exp	ERX14406279
	Short read run	ERR15002232
	Short read count	1,595,120
	Short read mean length (RI, R2)	150,150
	Short read GC%	53.5
	Long read exp	ERX14512249
	Long read run	ERR15107551

(Continued on next page)

TABLE 1 Summary table (Continued)

Sample	Property	Value
SAMEA118333803	Long read count	44,497
	Long read mean length	6424
	Long read GC%	53.5
	Assembly	GCA_965579335
	Total size	4,550,329
	Contigs	4
	GC%	53.5
	N50	7959060
	ANI (<i>O. pituitosum</i> CCUG50899)	98.40
	Annotation	ERZ26904287
	Annotated CDSs	4285
	Organism	<i>O. pituitosum</i>
	Taxon ID	71256
	Isolate/Strain	MPB25478
	Short read exp	ERX14512100
	Short read run	ERR15107402
	Short read count	1,630,005
	Short read mean length (R1, R2)	150,150
	Short read GC%	53.3
	Long read exp	ERX14512251
	Long read run	ERR15107553
	Long read count	74,814
	Long read mean length	3862
	Long read GC%	53.3
	Assembly	GCA_965579345
	Total size	4,852,914
	Contigs	4
	GC%	53.3
SAMEA118333804	N50	39,450,584
	ANI (<i>O. pituitosum</i> CCUG50899)	98.52
	Annotation	ERZ26904288
	Annotated CDSs	4552
	Organism	<i>P. saccharolyticum</i>
	Taxon ID	54352
	Isolate/Strain	MPB25528
	Short read exp	ERX14406336
	Short read run	ERR15002289
	Short read count	440,692
	Short read mean length (R1, R2)	150,150
	Short read GC%	51.5
	Long read exp	ERX14512252
	Long read run	ERR15107554
	Long read count	48,318
	Long read mean length	5292
	Long read GC%	56.3
	Assembly	GCA_965579315
	Total size	3,733,418
	Contigs	3
	GC%	51.5
	N50	12184706
	ANI (<i>P. saccharolyticum</i> DSM25260)	96.29
	Annotation	ERZ26904290

(Continued on next page)

TABLE 1 Summary table (Continued)

Sample	Property	Value
SAMEA118333805	Annotated CDSs	3348
	Organism	<i>P. saccharolyticum</i>
	Taxon ID	54352
	Isolate/Strain	MPB25488
	Short read exp	ERX14406304
	Short read run	ERR15002257
	Short read count	419,074
	Short read mean length (R1, R2)	150,150
	Short read GC%	51.4
	Long read exp	ERX14512253
	Long read run	ERR15107555
	Long read count	33,095
	Long read mean length	6,247
	Long read GC%	55.1
	Assembly	GCA_965579355
	Total size	3,753,639
	Contigs	2
	GC%	51.5
	N50	10,475,352
	ANI (<i>P. saccharolyticum</i> DSM25260)	96.35
	Annotation	ERZ26904289
	Annotated CDSs	3365
SAMEA118333806	Organism	<i>P. saccharolyticum</i>
	Taxon ID	54352
	Isolate/Strain	MPB25441
	Short read exp	ERX14406305
	Short read run	ERR15002258
	Short read count	478,435
	Short read mean length (R1, R2)	150,150
	Short read GC%	51.6
	Long read exp	ERX14512255
	Long read run	ERR15107557
	Long read count	149,774
	Long read mean length	5140
	Long read GC%	54.7
	Assembly	GCA_965579305
	Total size	3,839,354
	Contigs	2
	GC%	51.6
	N50	5,324,660
	ANI (<i>P. saccharolyticum</i> CCUG33852)	96.20
	Annotation	ERZ26904291
	Annotated CDSs	3473

inferred from a *gyrA* neighbor-joining phylogeny (Tamura-Nei distance) using ClustalΩ v1.2.4 (1, 18).

Strains taxonomy was determined by ANI analysis (19, 20), including type strains *Ochrobactrum pituitosum* CCUG:50899 and *Pseudochrobactrum saccharolyticum* DSM:25620, using FastANI v0.1.3 (21) (Fig. 1). All samples, raw reads, assemblies, and annotations are deposited on the ENA, see Table 1.

ACKNOWLEDGMENTS

We thank the Max Planck Society for generous core support and acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG) Collaborative Research Center 1182 "Origin and Function of Metaorganisms" (grant no. SFB1182). We are also grateful to Pauline Buffard for short read sequencing data, and Ishan Bhatt for contributions to bioinformatic processing.

AUTHOR AFFILIATIONS

¹Department of Microbial Population Biology, Max Planck Institute for Evolutionary Biology, Plön, Germany

²Laboratory of Biophysics and Evolution, CBI, ESPCI Paris, Université PSL, CNRS, Paris, France

AUTHOR ORCIDS

Sandhya Lakshmi Narayanan  <http://orcid.org/0009-0007-3995-795X>

Carsten Fortmann-Grote  <http://orcid.org/0000-0002-2579-5546>

AUTHOR CONTRIBUTIONS

Sandhya Lakshmi Narayanan, Data curation, Formal analysis, Investigation, Software, Visualization, Writing – original draft, Writing – review and editing | Carsten Fortmann-Grote, Data curation, Formal analysis, Investigation, Software, Validation, Writing – original draft, Writing – review and editing | Paul B. Rainey, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

Samples and data were submitted to the ENA BioProject [PRJEB87311](https://www.ebi.ac.uk/ena/browser/view/PRJEB87311), see Table 1 for accessions. Code for gyrA and ANI analysis is deposited in <https://archive.softwareheritage.org/swh:1:dir:067262a26b6b51b384fe7955736135216cbbca5d>.

REFERENCES

1. Buffard P. 2023. A simple path to complexity: horizontal gene transfer in microbial communities. PhD
2. Theodosiou L, Farr AD, Rainey PB. 2023. Barcoding populations of *Pseudomonas fluorescens* SBW25. *J Mol Evol* 91:254–262. <https://doi.org/10.1007/s00239-023-10103-6>
3. Picelli S, Björklund AK, Reinius B, Sagasser S, Winberg G, Sandberg R. 2014. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res* 24:2033–2040. <https://doi.org/10.1101/gr.177881.114>
4. Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32:3047–3048. <https://doi.org/10.1093/bioinformatics/btw354>
5. Fortmann-Grote C. 2025. Raw and processed ont long read sequencing data for "complete genome sequences of four *Ochrobactrum pituitosum* strains and three *Pseudochrobactrum saccharolyticum* strains isolated from *Caenorhabditis elegans* gut microbiomes. Available from: <https://doi.org/10.5281/ZENODO.17973826>
6. R. Wick. 2022. Rrwick/filtlong. <https://github.com/rrwick/Filtlong>.
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. Oxford Nanopore technologies Ltd. nanoporetech/medaka. 2026. Available from: <https://github.com/nanoporetech/medaka>
9. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification: find out more about bakta, the motivation, challenges and applications, here. *Microb Genom* 7:11. <https://doi.org/10.1099/mgen.0.000685>
10. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
11. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. 2023. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods* 20:1203–1212. <https://doi.org/10.1038/s41592-023-01940-w>
12. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. 2016. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* 17:132. <https://doi.org/10.1186/s13059-016-0997-x>
13. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
14. Hu J, Fan J, Sun Z, Liu S. 2020. NextPolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics* 36:2253–2255. <https://doi.org/10.1093/bioinformatics/btz891>
15. Bonfield JK, Marshall J, Danecek P, Li H, Ohan V, Whitwham A, Keane T, Davies RM. 2021. HTSLib: C library for reading/writing high-throughput sequencing data. *Gigascience* 10:giab007. <https://doi.org/10.1093/gigascience/giab007>
16. Bouras G, Grigson SR, Papudeshi B, Mallawaarachchi V, Roach MJ. 2024. Dnaapler: a tool to reorient circular microbial genomes. *JOSS* 9:5968. <https://doi.org/10.21105/joss.05968>

17. McMurdie PJ, Holmes S. 2013. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. <https://doi.org/10.1371/journal.pone.0061217>
18. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using clustal Omega. *Mol Syst Biol* 7:539. <https://doi.org/10.1038/msb.2011.75>
19. 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>
20. Rodriguez-R LM, Conrad RE, Viver T, Feistel DJ, Lindner BG, Venter SN, Orellana LH, Amann R, Rossello-Mora R, Konstantinidis KT. 2024. An ANI gap within bacterial species that advances the definitions of intra-species units. *mBio* 15:e0269623. <https://doi.org/10.1128/mbio.02696-23>
21. 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>